DOE/EE-1721

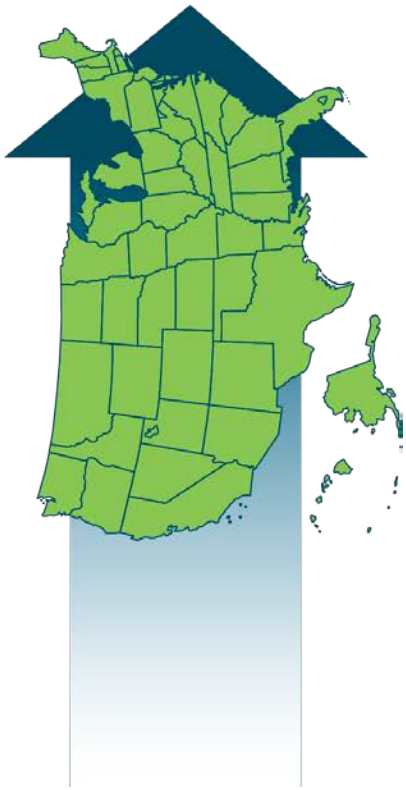
SEE Action

STATE & LOCAL ENERGY EFFICIENCY ACTION NETWORK

SEE Action Guide for States: Evaluation, Measurement, and Verification Frameworks—Guidance for Energy Efficiency Portfolios Funded by Utility Customers

Evaluation, Measurement, and Verification Working Group

January 2018



The State and Local Energy Efficiency Action Network is a state and local effort facilitated by the federal government that helps states, utilities, and other local stakeholders take energy efficiency to scale and achieve all cost-effective energy efficiency by 2020.

Learn more at www.seeaction.energy.gov

SEE Action Guide for States: Evaluation Measurement and Verification (EM&V) Frameworks—Guidance for Energy Efficiency Portfolios Funded by Utility Customers, is a product of the State and Local Energy Efficiency Action Network (SEE Action), facilitated by the U.S. Department of Energy and the U.S. Environmental Protection Agency. Content does not imply an endorsement by the individuals or organizations that are part of SEE Action working groups, and does not reflect the views, policies, or otherwise of the federal government.

This document was final as of January 12, 2018.

This report was funded by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under Lawrence Berkeley National Laboratory Contract No. DE-AC02-05CH1131. The EERE Project Manager is Michael Li.

If this document is referenced, it should be cited as:

Steven R. Schiller and Tom Eckman. 2017. *Evaluation Measurement and Verification (EM&V) Frameworks—Guidance for Energy Efficiency Portfolios Funded by Utility Customers.* Prepared by Lawrence Berkeley National Laboratory for the State and Local Energy Efficiency Action Network.

FOR MORE INFORMATION

Regarding *SEE Action Guide for States Evaluation Measurement and Verification (EM&V) Frameworks—Guidance for Energy Efficiency Portfolios Funded by Utility Customers,* please contact:

Michael Li
U.S. Department of Energy
michael.li@ee.doe.gov

Niko Dietsch
U.S. Environmental Protection Agency
dietsch.nikolaas@epa.gov

Regarding the State and Local Energy Efficiency Action Network, please contact:

Johanna Zetterberg
U.S. Department of Energy
johanna.zetterberg@ee.doe.gov

Acknowledgements

The authors wish to acknowledge the support and guidance of U.S. DOE project manager, Michael Li, who made this guide possible.

We also thank the following individuals for providing valuable input and extensive comments for a review draft of this guide. Although the authors benefited immensely from the input of these people, the content and any opinions stated are the responsibility of the authors and the listing of these individuals here is not intended to imply their agreement regarding any or all elements of this guide.

- Kevin Cooney, Navigant
- Joseph Dolengo, National Grid
- Sami Khawaja, The Cadmus Group, LLC
- Joe Loper, Itron
- Jane Peters, Research Into Action
- Valerie Richardson, DNV/GL
- Robert Stephenson, Vermont Energy Investment Corporation
- Mary Sutter, Grounded Research and Consulting, LLC

The National Renewable Energy Laboratory provided technical editing, formatting, and graphics services.

Acronyms and Abbreviations

ACEEE	American Council for an Energy-Efficient Economy
AEA	American Evaluation Association
AESP	Association of Energy Services Professionals
AMI	advanced metering infrastructure
ASHRAE	American Society of Heating, Refrigerating and Air-Conditioning Engineers
BPA	Bonneville Power Administration
CALMAC	California Measurement Advisory Council
CaSPM	California Standard Practice Manual
CE	cost-effectiveness
CO_2	carbon dioxide
DSM	demand-side management
DOE	United States Department of Energy
EEAC	Energy Efficiency Advisory Committee (Delaware)
EERE	Office of Energy Efficiency and Renewable Energy
EEPS	energy efficiency performance standard
EERS	energy efficiency resource standard
EMC	Evaluation Management Committee (Massachusetts)
EM&V	evaluation, measurement, and verification
ESCO	energy service company
ESPC	energy savings performance contract
EUL	effective useful life
EVO	Efficiency Valuation Organization
FEMP	Federal Energy Management Project
HPUC	Hawaii Public Utilities Commission
IEPEC	International Energy Program Evaluation Conference
IPL	Indianapolis Power and Light
IRP	integrated resource planning
IPMVP	International Performance Measurement and Verification Protocol
kW	kilowatt
kWh	kilowatt-hour

M&V	measurement and verification
M&V 2.0	advanced measurement and verification data collection and analytics
MEA	Maryland Energy Administration
MMBtu	Million British Thermal Units
NEI	non-energy impacts
NOx	nitrogen oxides
NSPM	National Standard Practice Manual
NTG	net-to-gross
PSE	Puget Sound Energy
RVF	resource value framework
RTO	regional transmission organization
SEE Action	State and Local Energy Efficiency Action Network
SO_X	sulfur oxides
SWE	statewide evaluator (Pennsylvania)
T&D	transmission and distribution
TRM	technical reference manual
UMP	U.S. Department of Energy's Uniform Methods Project
WUTC	Washington State Utilities and Transportation Commission

Table of Contents

List of Tables

List of Figures

About This Document

This guide addresses developing evaluation, measurement, and verification (EM&V) or "evaluation" framework documents for energy efficiency portfolios, specifically those associated with programs funded by utility customers. An EM&V framework is a primary guiding document that defines EM&V objectives, processes, and activities that constitute a jurisdiction's EM&V infrastructure. Thus, at a minimum, frameworks set forth a jurisdiction's fundamental evaluation goals, principles, metrics, and definitions; summarize budgets, schedules, and reporting expectations; indicate policies that define allowable EM&V baselines and methods for assessing efficiency actions and their cost-effectiveness; and define the roles and responsibilities of various entities involved in EM&V. The EM&V frameworks serve two primary purposes: Supporting consistent, documented, and comparable EM&V within a jurisdiction; and providing—for all stakeholders—an understanding of how EM&V is conducted within a jurisdiction, which can reduce stakeholder concerns regarding EM&V results. In addition to efficiency portfolios, EM&V frameworks also can be applicable to other types of distributed energy resources, such as demand response.

This guide focuses on topics that can be documented in EM&V frameworks. It also discusses processes for developing frameworks, emphasizes collaborative efforts, describes some example frameworks, and provides background on EM&V concepts and methods. The intended audience for this guide is those parties involved in creating, reviewing, and possibly approving an EM&V framework, including state-utility regulators, administrators of energy efficiency programs (including publicly owned and investor-owned utilities and government and non-governmental organizations), efficiency program implementers, evaluation consultants, and other stakeholders, such as energy efficiency industry representatives and consumer advocates. All of these groups have a direct interest in the policies that guide efficiency portfolios and programs and the processes that are used to evaluate them.

Using This Document

This guide is intended to serve as a reference for users across a wide range of EM&V experience. Those users who are less familiar with efficiency programs and their evaluation might find it beneficial to read the first two sections. Sections 1 and 2 provide an overview of efficiency EM&V context and the basic components of EM&V infrastructures as described in frameworks. Additionally, Appendix A and Appendix B provide more details about EM&V methods and references to standard industry resources. Users with more experience with efficiency programs and evaluation might wish to proceed directly to descriptions of framework contents and development processes provided in Section 3 and Section 4 (including the process flow chart and checklist located at the end of Section 4) and the references to framework examples in Section 5.

Executive Summary

This guide describes creating and updating frameworks for evaluation, measurement, and verification (EM&V) of utility customer–funded energy efficiency programs. An EM&V framework codifies how a jurisdiction's EM&V will be conducted, and typically addresses topics such as EM&V principles, objectives, budgets, schedules, reporting requirements, and the roles and responsibilities of various entities involved in EM&V. Frameworks are intended to be functional documents that essentially define a jurisdiction's EM&V infrastructure, and thus are useful for those people involved in any aspect of the evaluation process.

Table ES-1 lists the topics commonly covered in framework documents and discussed in this guide. The table is organized into five categories of topics, three of which are recommended for inclusion in all EM&V frameworks. Although certain fundamental topics should be addressed in all framework documents, the full range of topics they address can vary from jurisdiction to jurisdiction. Additionally, the objectives principles, and EM&V methods described in EM&V frameworks, as well as the formats, can vary from jurisdiction to jurisdiction.

FRAMEWORKS SUPPORT CONFIDENCE IN EVALUATION RESULTS

EM&V helps stakeholders understand both how much savings occurred (and for whom) and why those savings occurred—essentially showing what works and why. When stakeholders are engaged in developing framework documents, frameworks can aid these stakeholders in understanding, appreciation, and support of the evaluation process and results.

Table ES.1. EM&V Framework Document Topics

Foundational Topics –Recommended for Inclusion in All Frameworks			Optional Topics –Recommended for Consideration	
Fundamental Topics/Issues	**EM&V Scope Topics**	**Impact Evaluation Approach Topics**	**Other Evaluations Topics**	**Logistics Topics**
Definitions	Metrics	EM&V methods that will/can be used	Considerations for specific market segments, for example, low-income and hard-to-reach	Data management approaches, including confidentiality considerations
Efficiency portfolio description and goals	Evaluation reports and other outputs	How baselines are defined		Report formats, including websites and public access
Evaluation objectives	Scale of the evaluation effort (e.g., budget)	Sampling	Cost-effectiveness studies	
Evaluation principles	Timing of the evaluation cycles and reporting	Consideration of interactive effects and persistence	Process evaluations	
How evaluated savings estimates are applied—retrospective or prospectively	Roles and responsibilities	Expectations for metric certainty (reliability)	Market evaluations	Dispute resolution

An evaluation framework document tends to be "fixed" for several years because its focus is on a jurisdiction's evaluation infrastructure, rather than on specific EM&V activities which can change from year to year as programs and efficiency measures change and technology enables new methods. Frameworks should be reviewed and

revised periodically, however, because both programs and evaluation needs evolve. When comprehensively prepared, the framework also sets the expectations for the content and scope of the other evaluation documents such as program EM&V plans, project and energy efficiency measure specific analyses, technical reference manuals (TRMs) and evaluation reports. Figure ES-1 shows a hierarchy of the evaluation planning documents typically covered within a framework.

Figure ES.1. Energy efficiency evaluation planning document hierarchy

The authors reviewed multiple frameworks developed around the United States and gathered input from other experts to prepare this guide as a reference for users across a wide range of EM&V experience. The next two sections provide introductions to energy efficiency and energy efficiency EM&V (Section 1), and the components of a jurisdiction's EM&V infrastructure, the framework document, and other typical EM&V planning and reporting documents (Section 2). Subsequent sections provide descriptions of content options for framework documents (Section 3) and approaches for developing these documents (Section 4). Section 4 includes a figure that lays out the task categories and major tasks associated with developing a framework as well as process and content checklists. Section 5 points to some example framework documents that have been developed for specific efficiency portfolio jurisdictions or administrators. A references section is provided along with appendices on EM&V basics and resources.

1. Introduction and Background

1.1. Efficiency Program Context

Energy efficiency activity in the United States largely has been driven by building codes, appliance and equipment efficiency standards, contractors' conventional and energy service performance contracts, and programs paid for by utility customers. The programs paid for with utility-customer funds cover technologies and practices designed to improve both natural gas and electricity efficiency in all market sectors (e.g., residential, commercial) and represent a significant portion of the funding for efficiency investments.[1] These programs are administered in different states by utilities, state agencies, non-profits, and for-profit entities and are overseen by regulators or other decision makers such as public utility commissions, boards of publicly owned utilities, and the governing bodies of cooperative utilities. Utility customer–funded energy efficiency programs are the focus of this document and are simply referred to as efficiency programs or portfolios throughout the document. Much of the content, however, is also applicable to evaluation frameworks that can be used with other utility customer–funded distributed energy resource programs, such as those associated with demand response, distributed generation, and distributed energy storage.

1.2. Efficiency Evaluation, Measurement, and Verification Context

Efficiency portfolios can be considered to have three interrelated and cyclical components—planning, implementation, and evaluation. This is conceptually presented in Figure 1.1. Although this guide discusses the development of a framework for efficiency portfolio evaluation, measurement, and verification (EM&V), it is important to realize that EM&V is not an end in itself, but rather is a support function for planning, implementing, and providing oversight for successful efficiency actions. That is, evaluation is part of a continuous improvement process and not just activities undertaken to document savings. Thus, the concept of EM&V frameworks is presented and is discussed in the context of integration of evaluation into and in support of portfolio planning and implementation.

EM&V FUNDAMENTALS

Appendix A provides some background information on efficiency EM&V fundamentals, metrics, and methods for impact evaluation. This appendix is intended as introductory primer for those less familiar with efficiency EM&V. Appendix B lists some additional EM&V resources available online.

Figure 1.1. The elements of efficiency program planning, implementation, and evaluation cycle (Schiller 2012)

[1] See related reports at https://emp.lbl.gov/projects/utility-customer-funded.

Efficiency EM&V can have three primary objectives, as shown in Figure 1.2.

- **Document the benefits and costs (impacts)** of a program and determine whether the program (or portfolio of programs) met its goals.
- **Understand why program-induced effects occurred** and identify ways to improve current and future programs, including supporting continuous improvement—such as updating measure and program-impact estimates for potential studies and technical reference manuals (TRMs).
- **Support energy demand forecasting and resource planning** by understanding the historical and future resource contributions of efficiency as compared to other energy resources.

Figure 1.2. Evaluation objectives (Schiller 2012)

Each jurisdiction, as it develops its own efficiency portfolio, should consider one or more of the above-listed EM&V objectives as it establishes its EM&V infrastructure (i.e., methods and approaches, budgets, timelines) that balances best practices with costs, timing considerations, and acceptable levels of risk—specifically with regard to determining efficiency impacts (see text box below).

EM&V should do the following with regard to best practices.

- Be integral to a typically cyclic planning-implementation-evaluation process. Therefore, evaluation planning should be part of the program planning process, including the alignment of implementation and evaluation budgets and schedules. This is done so that evaluation information can be provided in a cost-effective and timely manner.
- Support the efficiency programs being evaluated and, in general, the success of the programs and related energy and policy goals, by providing appropriate documentation of progress toward the goals, as well as feedback required by program administrators and implementers to continuously improve the programs and plan future efforts.
- Utilize industry-standard EM&V methods, analytical tools, and data-collection methods to the furthest extent possible.
- Be based on budgets and resources adequate to support—over the entire evaluation cycle—the evaluation goals and the level of quality (certainty) expected in the evaluation results.

WHAT'S AN ACCEPTABLE LEVEL OF RISK?

Relative risk is an important concept for consideration when developing EM&V infrastructures—which inevitably involves trade-offs between risks and benefits, and influences how much to spend on evaluation. For efficiency, risk management is dominated by the inability to directly measure savings, which creates uncertainty. This of course is balanced by the uncertainties associated with other (e.g., supply-side) resources (e.g., performance risks, cost of construction, uncertainties associated with future fuel costs). This leads to a basic impact evaluation question: "How good is good enough?" or, less succinctly, "How certain does one have to be of the energy savings estimate that results from EM&V activities, and is that level of certainty properly balanced against the amount of effort (e.g., resources, time, money) it takes to obtain that level of certainty?"

Tolerance for uncertainty is driven by how much risk is associated with getting the wrong answer. With efficiency, for example, the risks include crediting too much or too little savings to the actions that have been taken as part of an efficiency program. This can lead to expending too many resources on ineffective actions (or the opposite), or simply not obtaining a desired outcome (e.g., less energy consumption). Other risks are counterbalancing. These include spending too much on EM&V beyond the importance of reducing uncertainty (e.g., improving the confidence or precision of savings from a measure with small total potential). Potential overinvestments in EM&V to reduce relative risks can result in less investment in efficiency resources that are then replaced with other energy resources that have different—perhaps greater—risks associated with their performance, lifecycle costs, or both.

See Section 3.3.4 for a discussion of how EM&V frameworks might address uncertainty and risk. For a general discussion of risk management in an efficiency portfolio see Lemoine et al. 2008.

1.3. Evaluation (EM&V) Categories

Efficiency evaluation includes a range of assessment studies and other activities aimed at determining the effects of an efficiency program. There are four broad categories of efficiency evaluations: impact evaluations, process evaluations, market evaluations, and cost-effectiveness evaluations. Brief definitions of all four of these are provided below.

- **Impact evaluations** are assessments that can determine direct and indirect performance of an energy efficiency program. Impact evaluation involves near real-time or retrospective assessments of performance. Program impacts can include energy and demand savings and non-energy benefits (e.g., avoided emissions, job creation and local economic development, water savings). Impact evaluations are probably the most common efficiency evaluations conducted and they also support cost-effectiveness analyses.
- **Process evaluations** are formative, systematic assessments of efficiency programs, typically conducted retrospectively. They document program operations and identify and recommend improvements that are likely to increase a program's efficacy or effectiveness for acquiring efficiency resources, preferably while maintaining high levels of participant and market satisfaction. Process evaluations can target specific parts of a program's operation (e.g., the quality-control procedures) or an entire program.
- **Market evaluations** are assessments and characterizations of the structure or functioning of a market, the behavior of market participants, and market changes that result from one or more program efforts. Market assessments evaluations can include estimates of the current market role of efficiency (market baselines), as well as the potential role of efficiency in a market (potential studies). Market evaluations characterizations can indicate how the overall supply chain and market for efficiency products works and how they have been affected by a program(s). Market evaluations—perhaps overlapping with the impact evaluation category—also can be used to establish the overall change in the efficiency mix in markets resulting from both customer-funded programs and other forces driving efficiency improvements. Such studies inform both efficiency program planning and provide input to load forecasts and resource planning.

- **Cost-effectiveness assessments are** analyses that show the relationship between the value of a portfolio's (or project's, measure's, or program's) benefits and the costs incurred to achieve those benefits. The findings help determine whether to retain, revise, or eliminate program elements and provide feedback on whether efficiency is an effective investment as compared with energy supply options.

IMPACT EVALUATIONS
• Validate energy savings, load reductions, and cost-effectiveness • Assess characteristics of participants and non-participants (consumers and trade allies) with respect to their contribution to impacts • Provide input to load forecasts and resource planning
PROCESS EVALUATIONS
• Focus on program operations and implementation • Examine ways to improve program marketing and implementation
MARKET EVALUATIONS
• Systematically assess an entire market or any part of a market (e.g., customers, suppliers, channels of distribution, specific services or products) • Assess current and past baseline practices and technologies • Assess efficiency potential • Assess how a program influences markets and how changes in the market can influence program design, delivery, and strategy • Assess overall changes in the mix of efficient and inefficient products or practices
COST-EFFECTIVENESS ASSESSMENTS
• Assess benefits and costs from various perspectives

1.4. The EM&V Process

As shown in Figure 1.1, the efficiency program process consists of planning, implementing, and evaluating activities. Throughout this process, energy-savings values are characterized differently.

- *Projected savings* are values reported by a program implementer or administrator during the planning phase and before the efficiency activities are completed.
- *Claimed savings* are the values reported by a program implementer or administrator after the efficiency activities have been completed.
- *Evaluated savings* are the values reported by an evaluator after the efficiency activities and evaluations on have been completed.

Figure 1.3 depicts the relationship between the efficiency program phase and the taxonomy of savings.

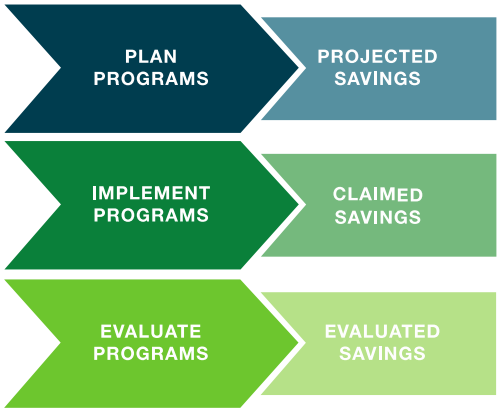


Figure 1.3. Reporting of impacts (such as energy savings) during planning, implementing, and evaluating phases of efficiency programs (Schiller 2012)

As shown in Figure 1.4, the evaluation phase of efficiency programs itself consists of planning, implementing, and reporting activities. These activities form the basis of the EM&V framework (discussed in the next section).

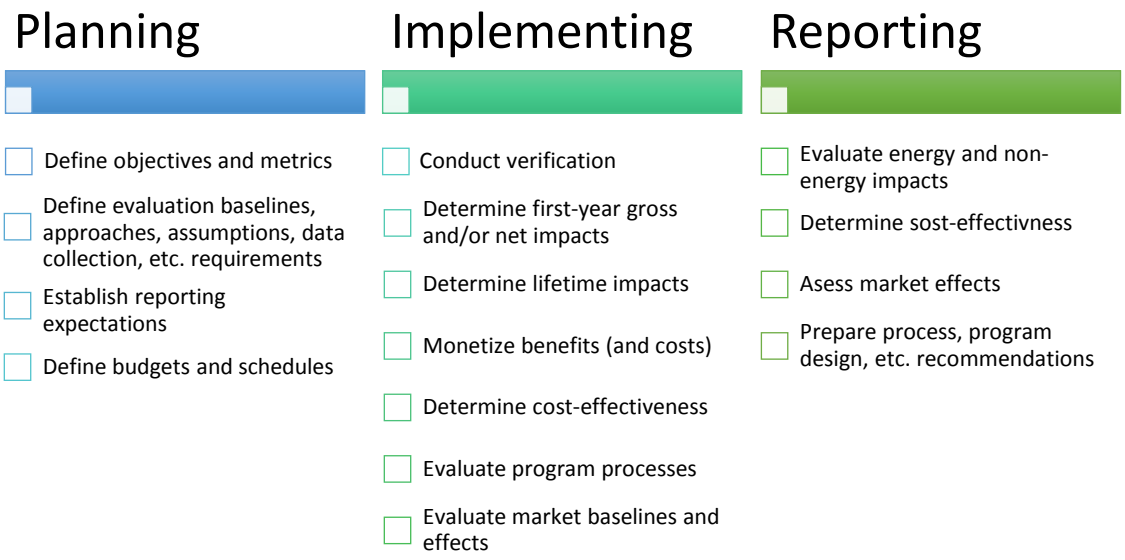


Figure 1.4. EM&V activities workflow for impact evaluation

2. EM&V Infrastructure, the Framework Document, and Related EM&V Plans and Reports

2.1. EM&V Infrastructure

In the context of EM&V for efficiency programs funded by utility customers, the term EM&V infrastructure is meant to describe the base or foundational components deployed to conduct EM&V in a jurisdiction. These components include policies, principles, metrics, EM&V processes and methods, products (e.g., evaluation reports), and roles and responsibilities of the entities involved in EM&V. For utility customer–funded efficiency programs, the list of infrastructure components is fairly consistent, although the approaches to EM&V described in these components and the associated resources vary significantly across jurisdictions.

FRAMEWORKS VERSUS PROTOCOLS

Within the efficiency community there is some overlap between the use of the terms EM&V Framework and EM&V Protocol. While there are no formal lines between the two, in this guide, Framework is defined as an overview document that covers ***general,*** and often policy, guidance on a wide range of why, what, who, how and when topics. EM&V protocols, on the other hand, prescribe ***details*** on the how of EM&V, i.e. how specific EM&V methods will be carried out, details on assumptions, etc. For example, a framework document might require or recommend the use of the Uniform Methods Project's (UMP) evaluation protocols. The UMP protocols include the specifics of sampling design and related accuracy requirements and the equations and assumptions to be used for energy use data analyses.

The concept of the EM&V Framework is for it to be a functional guidance document that is useful for the people who are engaged in the EM&V activities or using the output of the EM&V activities—from regulators to stakeholders. Thus, its exact form can vary from jurisdiction to jurisdiction as determined appropriate by the entity (entities) responsible for the EM&V (e.g., a regulator), hopefully with input from the framework's, and the EM&V output's, targeted users.

2.2. EM&V Framework

The EM&V framework is the document that explains and codifies the EM&V infrastructure with descriptions, requirements, and expectations associated with each covered component. An EM&V framework document is the primary, guiding record of the fundamental answers to the *why*, *what*, and *who* questions which form the basis of determining the appropriate EM&V infrastructure for an efficiency portfolio in a given jurisdiction. Frameworks also provide high-level[2] guidance and descriptions of *how* and *when* the EM&V will be conducted—including any impact, process, market, and cost-effectiveness evaluations. These elements are shown heuristically in Figure 2.1.

In effect, frameworks have two major elements. The first element is *broad* and sets the stage for evaluation by describing how evaluation works in the given jurisdiction. The other framework element is a *narrower* one that is directed at how evaluators should perform, and what stakeholders can expect evaluators to do and to report. This element can be explained as describing evaluation activities and setting specific requirements for evaluators to follow.

There are several ways in which the EM&V infrastructure can be described in a framework document. For example, the framework document can be organized by the participant roles (e.g., the role of regulators, portfolio administrators, program implementers), by outputs (e.g., metrics, reports), by resources (e.g., Technical Reference Manuals), or by rules/policies (e.g., regulations, protocols). This guide uses a functional description as the basic EM&V framework outline. These functional EM&V infrastructure topics are defined as follows.

- Fundamental topics and issues—including evaluation objectives and principles
- EM&V scope topics—including metrics, expected products (e.g., plans and reports), budget and timing, and roles and responsibilities requirements.

[2] The details of how EM&V is to be conducted—for example, specific methods, equations, assumptions—are defined in supporting documents, such as EM&V protocols and program-specific EM&V plans, which are defined elsewhere in this section.

- Evaluation methods—methods, approaches, and assumptions that will or can be used for impact, process, and market evaluations and cost-effectiveness analyses.
- Logistics—including topics such as data management and public access to evaluation information.

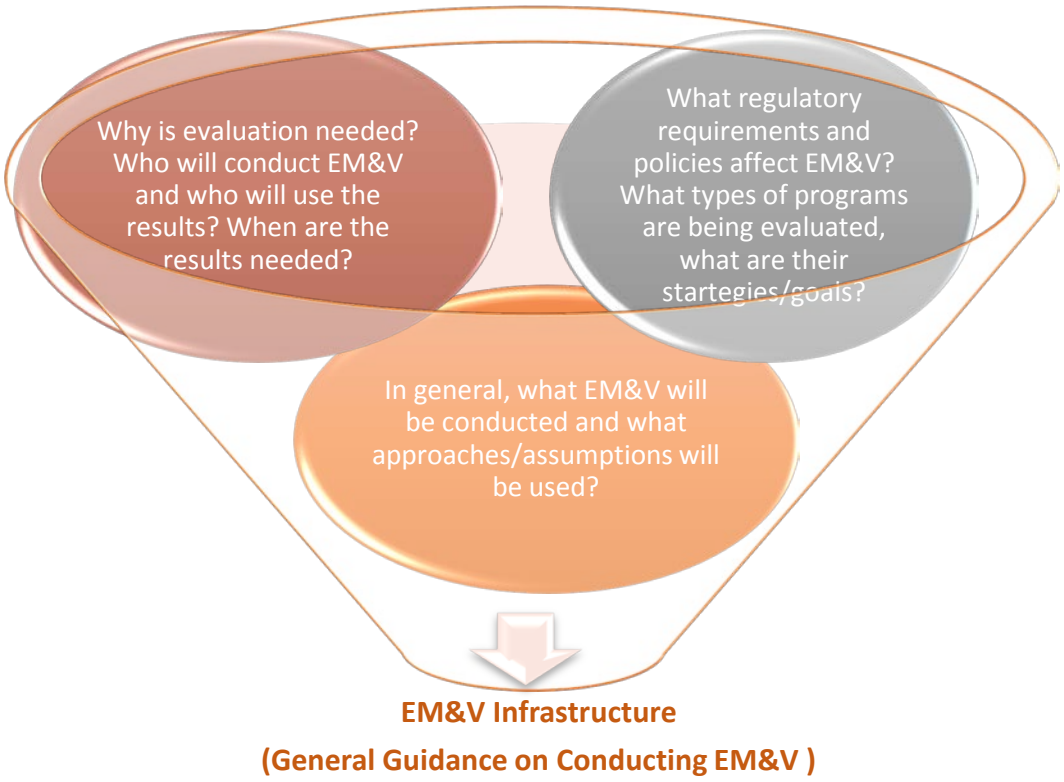


Figure 2.1. Heuristic presentation of EM&V framework contents

Section 3 discusses each of these topics in more detail. It also discusses which of these topics are (a) recommended for inclusion in *all* EM&V framework documents, and (b) which topics can be considered optional, but which jurisdictions probably will find useful to include when describing the EM&V infrastructure. What is included in the framework is up to those with responsibility for the EM&V efforts. The framework document is intended to support—not burden—successful EM&V, and provide confidence in the validity and value of the evaluation efforts.

As discussed in Section 4, the framework document can be prepared by entities such as a state regulator or portfolio administrator—usually with the support of an evaluation consultant. Frameworks, however, can have their greatest value when—irrespective of who has the lead in preparing the framework—stakeholders are engaged in its development. Stakeholder engagement is important because the framework is the principal document that they can focus on, It provides high-level input—the "forest" as compared to the "trees" of evaluation planning.

Regarding EM&V being a risk-mitigation strategy (see text box in Section 1), the framework also reduces risks associated with stakeholders disputing or not knowing the manner in which the EM&V will be conducted and the form and coverage of the EM&V reporting. If the framework is agreed to by at least the primary stakeholders prior to the performance of major evaluation activities, evaluation efforts can be well-supported and succeed in providing the results desired with minimal disagreement, at least over process-related issues.

For these reasons and to utilize the benefits of stakeholder input, it is recommended in this guide that collaborative approaches be used to develop frameworks. It also is suggested that regulators be involved in their development and approval. Regulator engagement in the development and approval of frameworks provides

program administrators, implementers, and evaluators with greater (but not absolute) assurance that both the process and results of EM&V activities will be accepted.

An evaluation framework document tends to be "fixed" for several years because its focus is on a jurisdiction's evaluation infrastructure, rather than on specific EM&V activities which can change from year to year as programs and efficiency measures change and technology enables new methods. Frameworks should be reviewed and revised periodically, however, as both programs and evaluation needs evolve. When comprehensively prepared, the framework also sets the expectations for the content and scope of the other evaluation documents such as program EM&V plans, project and energy efficiency measure specific analyses, technical reference manuals and evaluation reports. (These documents are discussed further in Section 2.3.)

Framework documents described in this guide can be just a few pages long—laying core principles, assignments, and activities that constitute the EM&V infrastructure—to hundreds of pages long, with coverage of multiple topics at varying degrees of detail. At least conceptually, about 20 to 50 pages can effectively cover the critical elements of a framework as defined in this guide.

How the evaluation infrastructure—its issues and principles—is defined in a framework for each jurisdiction greatly depends on the specific programmatic and regulatory context (including any performance mandates or administrator/utility shareholder financial performance incentives or penalties[3]) found within each jurisdiction, the objectives and scale of the efficiency activities being evaluated, and how EM&V results will be used. For example:

- One state might have very limited goals for energy efficiency and might not have performance incentives for its energy efficiency portfolio administrator. This state also could have a limited level of naturally occurring or mandated (via codes and standards) energy efficiency activity.
- Another state might have established aggressive, long-term energy-savings targets in legislation, developed a performance-based incentives scheme for program administrators, and have high energy costs as well as a need for very solid savings data for resource planning purposes. The high energy costs also might have resulted in a high level of natural and mandated energy efficiency activity.

Given the differences between these two hypothetical states, the first state's EM&V framework might only specify a limited level of EM&V, define permissive baselines, and focus on gross savings.[4] Conversely, the second state's EM&V framework might require very rigorous and more expensive EM&V—specify well-defined baselines—and use both net and gross savings metrics. Although these two hypothetical states' EM&V frameworks guidance on the topics of baselines, rigor, and reporting metrics are quite different, both frameworks address these fundamental topics so that all parties have clear expectations regarding their evaluation processes and results.

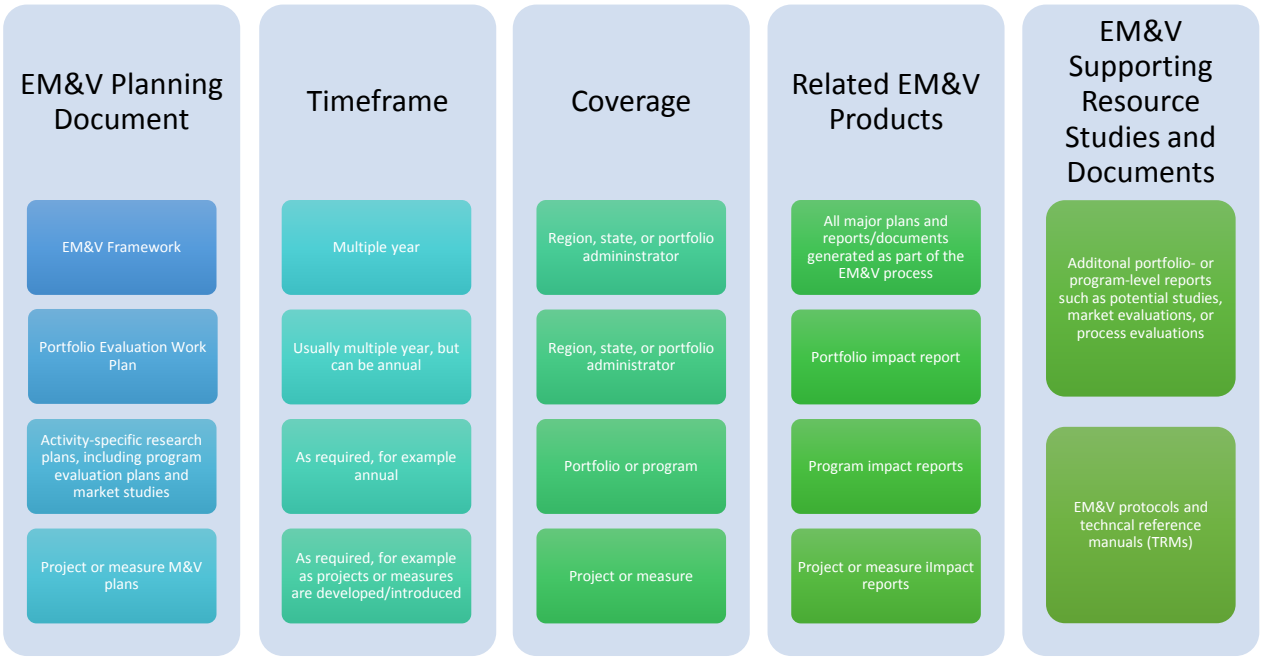
2.3. EM&V Planning Documents and Reports (Products)

This section describes the typical planning and reporting documents (or products) supporting or resulting from EM&V activities—and are all components of a jurisdiction's EM&V infrastructure. EM&V framework documents, which describe these plans and products, fit into a basic hierarchy and interrelationship of evaluation planning documents—the framework (at the top), portfolio EM&V plans, program EM&V plans, and project-specific M&V plans, with supporting EM&V studies and documents. The relationship of these documents to each other and to EM&V reports (portfolio, program, and project reports) is displayed in Table 2.1. Brief descriptions of each plan, report, and document type follow Table 2.1.

[3] The scale of the penalty or bonus can influence the level of rigor required in the evaluation as well as the level of attention paid to the evaluation activities.

[4] Gross and net savings are two of the common metrics used to describe the energy savings associated with efficiency activities, with the difference between the two mostly associated with attribution of savings to a particular program. See Appendix A for discussion of gross and net savings metrics as well as the resource document: *Chapter 23: Estimating Net Savings: Common Practices*. The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures (Violette and Rathbun 2014).

Table 2.1. Energy Efficiency Evaluation Plans and Related Documents



2.3.1. EM&V Planning Documents

Entities such as states, regulatory commissions, and utilities can establish and document their evaluation requirements and expectations in a series of planning documents. Table 2.1 outlines the hierarchy of these documents and indicates their typical applicability time frame and coverage level (e.g., state or utility program administrator, program or portfolio of programs, or individual projects). The increasing level of detail found in this hierarchy (from top row to bottom row in Table 2.1) of documents provides indications of the appropriate place for each stakeholder to provide input in the overall evaluation planning effort. For example, all stakeholders might be interested in an EM&V framework but, on the other end of the spectrum, only program implementers and evaluators usually are concerned with the details of a program- or project-specific plan.

Other than the EM&V Framework, these planning documents are described next with the recognition that, as with the EM&V frameworks, their scope and detail can vary from jurisdiction to jurisdiction.

- **Portfolio evaluation work plans** indicate which major evaluation activities will be conducted during the portfolio cycle (typically one, two, or three years) and describes them at a high level, including budget and allocation of resources and effort between programs, measures, market sectors. They tend to list all the planned, major evaluation activities with a schedule for when they are to be conducted. Examples of such major evaluation activities are the impact, process, and market evaluations, as well as preparation of supporting studies, such as updates to technical reference manuals, and potential studies. This EM&V plan, for example, might indicate (1) which programs would have rigorous impact evaluations in each year and which would only have installation or documentation verification reviews, (2) which programs will undergo process evaluations, and (3) which and when market studies can occur.
- **Evaluation Activity-Specific Research[5] Work Plans** are created for the major EM&V activities or studies planned in a given evaluation cycle prior. Such plans can focus on just determining gross or net impacts associated with programs or also asses the validity of the program theory for achieving these savings, as

[5] These are called "research" plans perhaps because of the history of evaluation activities conducted for a wide range of subjects well beyond efficiency as associated with research using experimental and quasi-experimental methods. The term research still applies, although this is very applied research.

might be described in a program's logic model.[6] Examples of these plans are (1) program-specific impact evaluation plans that go into substantial detail on what evaluation methods will be used, their schedules and budgets, the data that will be collected, and results that will be reported; (2) process, market effects, market baseline, and potential study plans that similarly provide sufficient detail to guide the actual implementation of the activity; and (3) research or evaluations for use as input to TRMs.

- **Project- or Measure-Specific M&V Work Plans.** Project-specific plans might be required for projects or measures that are part of a custom program and that are selected for analysis and inspection. These are the plans that describe the specific activities that would be conducted at a single site when one or more of the four International Performance Measurement and Verification Protocol (IPMVP)[7] measurement and verification options are applied, or just for inspections if a deemed savings method[8] is to be used.

2.3.2. EM&V Reports and Supporting Studies and Resource Documents

Evaluators report evaluation results and, as appropriate, provide input and work with (1) regulators, to assess whether goals have been met; (2) program administrators, to implement recommendations for current or future program improvements; and (3) resource planners, to understand the historical role and future role of energy efficiency as an energy resource (for example, as input to potential studies or updated TRMs). Reporting also provides information to energy consumers and the general public, who are funding these efforts (e.g., through charges on their utility bills), on what has been achieved with their investments.

Correlated with the evaluation planning documents described above, the following are the typical types of impact evaluation reports.[9]

- **Project-Specific M&V Reports.** These reports document the impacts determined for a specific project, site, or measure and the methods used to determine the impacts. They tend be reviewed with the project administrators and implementers before they are finalized, and their results are made available on a limited basis to protect the confidentiality of the consumer information that is usually included.
- **Program Impact Evaluation Reports.** The results of conducting the evaluation activities described in each impact evaluation plan are documented in an impact evaluation report. The report documents the impacts, and perhaps cost-effectiveness, of a program, as well as the methods used to determine the impacts. Program administrators and implementers usually have opportunities to provide input on these reports. The final reports also often are publicly available because they do not contain customer-specific or other confidential information.
- **Portfolio Impact Evaluation Reports.** The results of carrying out the evaluation activities described in an EM&V portfolio plan are documented in a portfolio cycle (e.g., annual or biennial) evaluation report. It documents the energy impact metrics (e.g., gross and net energy and demand savings, first year, and lifetime) and usually cost effectiveness associated with the portfolio of programs as well as the methods used to determine the impacts. Program administrators and implementers also usually have opportunities to provide input on these reports. The final reports almost always are made publicly available with summaries or graphics provided in a manner that is accessible to laypersons and with guidance on context and interpretation of the evaluation findings.

---

[6] Logic modeling is a thought process that efficiency program managers and evaluators use to develop a plausible and sensible model of how a program will work under defined conditions to solve identified problems. The logic model can be the basis for presenting a convincing story of the program's expected performance–telling stakeholders and others the problems the program focuses on, how the program will address the problems, and what outcomes and metrics can be used to assess success. Source: http://energy.gov/eere/analysis/program-evaluation-program-logic. Evaluations using logic models are known as theory-based evaluations.

[7] See www.evo-world.org.

[8] Schiller et. al. (2017). https://www4.eere.energy.gov/seeaction/system/files/documents/TRM%20Guide_Final_6.21.17.pdf.

[9] Usually there are draft and final versions of these documents and the EM&V framework is a good vehicle for describing not only what is covered in reports, who prepares them, and when they are prepared but also who reviews the reports and what entity is responsible for approving a final version.

It is often the case that portfolio cycle evaluation reports are completed well after planning for the next program and evaluation cycles must be started. Therefore, it is often desirable to include provisions for the production of interim portfolio impact evaluation reports so that progress indicators can be provided and problems, if they exist, are identified before the end of the program cycle.

The above list indicates the typical reports associated with impact evaluations. Other reports often are prepared by evaluators as indicated in the last column of Table 2.1, however, and are briefly described below.

- **Technical Reference Manuals (TRMs).** TRMs[10] are a resource that contains energy efficiency measure information used in program planning, implementation, tracking, and reporting and evaluation of impacts associated with the subject measures. TRMs provide information primarily used for estimating the energy and demand savings of end-use energy efficiency measures associated with utility customer–funded programs. These almost always are publicly accessible documents. TRMs tend to be associated with the deemed savings and measurement and verification method approaches to impact evaluation. TRM measure characterizations often are informed by program- and measure-specific impact evaluation efforts, as described above.

- **EM&V Protocols**. EM&V protocols, which can be standalone documents or be contained within a TRM, address the detailed expectations or requirements for any aspect of the evaluation processes—such as sampling, use of control groups, and process evaluations. Jurisdiction-specific TRMs and evaluation protocols, if developed, tend to be based on more established industry-standard protocols and guides, such as those referenced in Appendix B.

- **Market Baseline and Effects Studies.** Market effects (impact) evaluations generally are designed to characterize the impact of programs or portfolios on the overall efficiency of specific markets (e.g., single-family or commercial office building lighting, industrial motor efficiency). Results from such evaluations can be used to support load forecasting, the assessment of remaining energy efficiency potential, program design, and resource planning. These evaluations focus on the total market rather than just on program participants. Consequently, their results could provide a much more accurate view of the both the current market status and trends.

  Some specific metrics reported in these studies are both total product (e.g., number of units) sales and total "efficient" product sales for products targeted by efficiency programs. This is to determine, among other possible indicators, the cumulative impact of all forces acting on a market, not just individual efficiency programs, as well as baselines from which efficiency program savings can be determined. The results of these studies provide input into potential studies and TRMs, as well as utility-load forecasts and resource planning.

- **Potential Studies.** Another form of market study (although not formally an "evaluation") is the potential study. Potential studies are conducted before a portfolio is implemented to assess future savings potentials for different efficiency technologies, strategies, or approaches in different customer markets. These studies also can assess customer needs and barriers to adoption of efficiency, as well as how best to address these barriers through program design. Potential studies indicate what can be expected in terms of savings from a program. Potential often is defined in terms of technical potential (what is technically feasible given commercially available products and services), economic potential (the level of savings that can be achieved assuming a certain level of cost effectiveness is required), and market potential (the portion of economic or technical potential that is estimated to be achievable given market forces such as consumer interest and contractor availability).[11]

- **Process Evaluations.** The goal of process evaluations is to produce better and more cost-effective programs—that is, support continuous improvement. Process evaluations meet this goal by assessing the

[10] For more information on TRMs see Schiller et. al. (2017).

[11] For more information on potential studies see National Action Plan for Energy Efficiency (2007), *Guide for Conducting Energy Efficiency Potential Studies*. Prepared by Philip Mosenthal and Jeffrey Loiter, Optimal Energy, Inc. https://www.epa.gov/sites/production/files/2015-08/documents/potential_guide_0.pdf.

processes that a program undergoes during implementation, documenting program goals and objectives from a variety of perspectives, and describing program strengths and weaknesses so that success is highlighted and improvements can be made in a timely manner. Thus, process evaluations examine the efficacy and effectiveness of program implementation procedures and systems. These evaluations usually consist of asking questions of those involved in the program, analyzing their answers, and comparing results to established best practices and to the program's logic model that describes how the various inputs/outputs/resources and activities in a program will lead to desired short-term, near-term, and long-term outcomes. Typical process evaluation results involve recommendations for changing a program's structure, implementation approaches, and goals.

JURISDICTION-SPECIFIC VERSUS "UNIVERSAL" FRAMEWORKS

Because of differences in approaches to EM&V and lack of significant "cross-border trading" of efficiency savings, it has not been necessary for regulatory commissions and administrators of each state's utility customer–funded efficiency programs to define a saved unit of energy in exactly the same way (e.g., some count net savings, others count gross savings; some include transmission and distribution losses, others do not; some require a very high level of confidence and precision, others do not). Thus, each jurisdiction (e.g., state) can, and many do, develop EM&V frameworks and protocols that are appropriate for their own situations, with the EM&V requirements for each jurisdiction linked to the needs of that jurisdiction.

There can be exceptions to this where efficiency and demand response projects can participate in regional programs. For example, the PJM Interconnection and ISO-New England (regional transmission organizations) impose specific EM&V requirements across state borders for their demand-response programs.

3. Contents of an EM&V Framework Document

This section presents and discusses topics (e.g., issues and principles) that are defined and included in a jurisdiction-specific framework document. As defined in this guide, frameworks provide an overview and guidance on EM&V infrastructure topics applicable to efficiency portfolios or programs that use utility customer (or public) funds and have a program administrator with some government (regulatory) oversight. Drawing from established EM&V principles, efficiency program administrators, evaluators, and agencies with responsibility for overseeing EM&V activities (e.g., state utility commissions, energy offices) can define their own policy-specific EM&V infrastructure, and then document it in a framework using one or more of the processes discussed in Section 4. The subsections of this section provide definition and discussion of the most common EM&V Framework topics, as listed in Table 3.1.

Although there are certain foundational topics that are applicable to all jurisdictions, not all of the topics discussed in this section must be covered in a framework document, as some simply might not be relevant for a given jurisdiction. Additionally, some topics might not be considered of sufficient importance to require coverage in the framework or not worth the effort involved in defining appropriate guidance and perhaps gaining stakeholder agreement on the topic. Further, it might be decided that an initial framework document only needs to cover certain major infrastructure topics, with subsequent versions getting into more detail and addressing other topics as the value of a framework is established.

In Table 3.1, the topics that are considered foundational and highly recommended to be included in all frameworks are topics in the first three columns—Fundamental Topics/Issues, EM&V Scope Topics, and Impact Evaluation Topics. Other topics in the last two columns are optional, but still recommended for inclusion when applicable.

Table 3.1. EM&V Framework Document Topics

Foundational Topics—Recommended for Inclusion in All Frameworks			Optional Topics—Recommended for Consideration	
Fundamental Topics/Issues	**EM&V Scope Topics**	**Impact Evaluation Approach Topics**	**Other Evaluations Topics**	**Logistics Topics**
Definitions	Metrics, including the use of net and gross savings metrics, which cost-effectiveness test(s) will be used, and whether site or source metrics will be used	EM&V methods that will/can be used	Considerations for specific market segments, such as low-income and hard-to-reach segments	Data-management approaches, including confidentiality considerations
Efficiency portfolio description and goals	Evaluation reports and other outputs	How baselines are defined	Cost-effectiveness studies	Report formats, including websites and public access
Evaluation objectives	Scale of the evaluation effort (e.g., budget)	Sampling design consideration of interactive effects and persistence expectations for metric certainty (reliability)	Process evaluations	Dispute resolution
Evaluation principles	Timing of the evaluation cycles and reporting (e.g., annual vs. program funding cycle)		Market evaluations including potential studies	
How evaluated savings estimates are applied—retrospective or prospectively and for what purpose	Roles and responsibilities, including what types or which entities will conduct the EM&V activities and which entities will have oversight responsibilities			

3.1. Fundamental Topics and Issues—Topics Recommended for Inclusion in All EM&V Frameworks

3.1.1. Definitions

Common terminology that is understood by all those involved in the efficiency portfolio design, implementation, and evaluation is important to ensuring that guidance indicated in a framework is clear and is not (or only is minimally) ambiguous. Agreed-upon definitions also are important for the discussions among stakeholders as the framework is developed so that discussions are not hindered by misunderstandings about terminology. This is particularly relevant for the efficiency field, as there is not universal agreement on the definition of a number of key terms. Thus, a glossary of key terms is considered an essential component of any framework.

National sources for definitions are listed below.

- The glossary in the *Energy Efficiency Program Impact Evaluation Guide*. State and Local Energy Efficiency Action Network. 2012. Prepared by Steven R. Schiller, Schiller Consulting, Inc., www.seeaction.energy.gov.
- *Energy Efficiency Program Typology and Data Metrics: Enabling Multi-State Analyses Through the Use of Common Terminology*. Ian M. Hoffman, Megan A. Billingsley, Steven R. Schiller, Charles A. Goldman, and Elizabeth Stuart. LBNL-6370E. 2013 (August 28). https://emp.lbl.gov/sites/default/files/lbnl-6370e.pdf.
- *NEEP EM&V Forum Glossary of Terms and Acronyms*. 2011. http://www.neep.org/emv-forum-glossary-terms-and-acronyms.

FRAMEWORK EXCERPTS

This section includes text boxes with examples of how different framework topics are addressed in EM&V frameworks found in the United States. These are examples pulled from publicly available frameworks and are not necessarily typical or recommended. Also, given the limited space of text boxes, the example excerpts were selected in part because they are succinct. Other frameworks' approaches to topics or additional text in the frameworks referenced can be from a few pages long to comprising entire chapters.

3.1.2. Efficiency Portfolio Description and Goals

A starting point for defining characteristics of an evaluation effort is to generally describe the portfolio of programs to be evaluated, the overall portfolio goals, and (usually briefly or in reference to) the relevant polices or regulations that drive the portfolio. This would include indicating what types of programs are to be evaluated (e.g., incentive programs, direct install programs, market transformation programs), their markets, and the scale (e.g., budgets) and time frame of the programs. Typically, this is a brief part of any framework and simply can include reference to other portfolio focused documents.

Although a general overview of the portfolio of programs to be evaluated is helpful in a framework, of particular value is for the EM&V framework to include descriptions of the portfolio's goals—the EM&V, at least in part, should be supporting these goals and assessing metrics associated with these goals. Some typical portfolio goals—beyond energy and demand savings, cost effectiveness, and maximizing energy or peak savings within portfolio budgets—include:

- Maximizing leverage of portfolio dollars in creating private investment in energy-efficient products and services;
- Deferring specific resources (e.g., peaking or baseload power plants, transmission or distribution investments), which imply specific value to the timing and location of efficiency resources;
- Reducing pollution;
- Including an energy efficiency resource standard with cumulative, versus just annual, energy-savings goals;
- Supporting emerging technologies;
- Expanding the efficiency programs in anticipation of increasing energy efficiency–impacts goals (or vice-versa, expectation for a reduction in efficiency program activity);
- Increasing local economic development (e.g., jobs);
- Maximizing participation in programs (market penetration), perhaps specifically emphasizing low-income or disadvantaged community participation; and
- Satisfying consumers.

In addition to generally describing the portfolio of programs to be evaluated and the portfolio goals, the EM&V framework also can indicate whether the utility or other program administrators are eligible to receive a financial incentive (bonus) for managing a successful portfolio—and whether there is a penalty for not meeting the

established goals. If program administrators are eligible for financial incentives, then once the evaluation results are reported EM&V frameworks often include greater specificity and detail as a strategy to reduce the potential for stakeholder and program administrator disagreements over methodological issues and results of the evaluation.

3.1.3. Evaluation Principles

Evaluation principles define professional values that are to be embodied in an evaluation framework and, in turn, help guide evaluators in the developing their evaluation plans and deliverables. Such principles are considered an essential element of best-practice EM&V frameworks. Although each jurisdiction should tailor its evaluation principles to the policies and regulations that drive its energy efficiency programs, such principles should represent sound technical, economic, and regulatory practices and would be expected to be consistent with the expectations of a wide range of stakeholders. The following are some examples of specific principles that can be used as starting points in a framework document.

EXAMPLE EM&V PRINCIPLES—PUGET SOUND ENERGY FRAMEWORK

When choosing and planning evaluations, the following guiding principles are taken into consideration.

- Secondary research is leveraged as appropriate.
- Evaluation design undergoes expert review before and during planning and implementation.
- All key assumptions used by program planners are documented and verified in evaluations.
- The procurement process used to select evaluation contractors is timely, flexible, and transparent.
- Evaluation dollars and efforts are prioritized to focus on areas of largest savings or greatest uncertainty.
- Over time, to improve program delivery, evaluations are used to refine input assumptions used in savings estimation and resource analysis.

Source: Puget Sound Energy (2015)

- **Integral to the portfolio cycle.** The evaluation process should be integral to what is typically a cyclic planning-implementation-evaluation process. Therefore, evaluation planning should be part of the program planning process, so that the evaluation effort can support, in a timely manner, existing and future program implementation.
- **Adequate resources.** Evaluation budgets and resources should be adequate to support the evaluation scope, goals, and the level of quality (certainty) expected in the evaluation results over the entire time frame that program impacts are to be assessed. If budgets and resources are limited, then scope, timing, goals, and certainty expectations likely should be revisited.
- **Completeness and transparency.** Results and calculations should be coherently and completely compiled. Calculations should be well-documented and transparent, with reported levels of uncertainty. Key qualities of a good, transparent analysis include the following.
    - Describes the approaches and the variables used to determine energy savings;
    - Documents and states critical assumptions;
    - Presents documentation in a format that enables the reviewer to follow a connected path from assumptions to data collection, data analysis, and results; and
    - Reports levels and sources of uncertainty.
- **Relevance and balance in risk management, uncertainty, and costs.** The data, methods, and assumptions should be appropriate for the evaluated program. The level of effort expended in the evaluation process should be balanced with respect to the value of the savings, the uncertainty of their magnitude, and the risk of overestimated or underestimated savings levels. Impacts should be calculated at a level of uncertainty such that they are neither conservative nor optimistic, but provide the most likely values.
- **Consistency.** Evaluators working with the same data and using the same methods and assumptions should reach the same conclusions.

- **Independent evaluation.** Evaluators should be as free of bias as is reasonable and should not have a stake in the outcome of the evaluations with respect to the performance of the programs under consideration. Evaluation ethics can be a critical foundational element of an EM&V framework. As a resource for consideration in preparing such guidance in a framework, the American Evaluation Association (AEA) has a set of guiding ethical principles for evaluators. Available on AEA's website (www.eval.org), these principles are summarized here.
    - **Systematic inquiry.** Evaluators conduct systematic, data-based inquiries.
    - **Competence.** Evaluators perform competently for stakeholders.
    - **Integrity/honesty.** Evaluators display honesty and integrity in their own behavior, and attempt to ensure the honesty and integrity of the entire evaluation process.
    - **Respect.** Evaluators respect the security, dignity, and self-worth of respondents, program participants, clients, and other evaluation stakeholders.
    - **Responsibility** for general and public welfare. Evaluators articulate and take into account the diversity of general and public interests and values that could be related to the evaluation.

EVALUATOR AND ADMINISTRATOR RELATIONS

In addition to calling for minimal bias on the part of the evaluator, principles also can acknowledge that the relationship between the evaluator and the implementers and administrators—whose work is being evaluated—should be cooperative. This allows for information sharing, access to project sites, and for the results of the evaluator to be considered valid by the implementers and administrators and thus considered to be useful input for program improvement. There always will be some stress in the relationship, however, as (1) the evaluator cannot be unduly influenced by the implementer/administrator, or for that matter, by whoever hires the evaluator, including an entity such as a state regulator; and (2) the administrator/implementer will have a sense that the work is being judged by the evaluator, because the evaluator very well could have a significant impact on the compensation or penalties applied to the implementers and administrators. More on this subject is included in Section 3.2.5.

3.1.4. Objectives for Performing an Evaluation

Evaluation objectives described in an EM&V framework should focus on defining the intended use(s) of the information determined through evaluation activities and the intended audiences for such information. As mentioned in Section 1, efficiency evaluations tend to have one or more of three primary objectives:

- Document the impacts of a program, both benefits and costs, and determine whether the program (or portfolio of programs) met its goals;
- Provide an understanding of why program-induced effects occurred, and identify ways to improve current and future programs—supporting continuous improvement, including updating measure and program impact estimates for potential studies and TRMs; and
- Support energy demand forecasting and resource planning by comprehending the historical and future resource contributions of energy efficiency as compared to other energy resources.

EXAMPLES OF EM&V OBJECTIVES—ARKANSAS FRAMEWORK

The role of a program evaluation is to:

- **Quantify Results:** Document, measure, and estimate the energy and demand savings of a program and determine how well it has achieved its goals and managed its budget; and
- **Gain Understanding:** Determine why certain program effects occurred (or didn't occur) and identify ways to improve and refine current and future programs, and help select future programs.

Source: Public Service Commission of Arkansas (2016)

Therefore, a critical step in preparing an evaluation framework document is simply picking which of these objectives (if not all) are applicable, prioritizing them, and making them more specific to the subject portfolio(s). The following are some more-specific evaluation objectives, starting with impact evaluation objectives.

- Measure and document energy and coincident peak savings attributable to a program in a manner that is defensible in proceedings conducted to ensure that funds are properly and effectively spent.
- Measure and document avoided emissions.
- Provide specific data for demand forecasts and resource planning in an integrated resource planning (IRP) effort.
- Assess cost-effectiveness of programs and portfolio.
- Document program milestones, such as homes weatherized or people trained.
- Assess whether customer class equity requirements were met.
- Inform decisions regarding program administrator compensation, including any incentive payments (for regulated programs and performance-based programs). Assess whether there is a continuing need for each program in a portfolio.
- Provide ongoing feedback and guidance to the program administrator.

In practice, the selection of evaluation objectives is shaped by many situational factors, the most important of these are the program goals and how evaluation results will be used to support such goals. Therefore (as mentioned in Section 3.1.2) the EM&V framework should specify program goals that are sufficiently well-defined to enable the development of quantifiable evaluation metrics.

3.1.5. How Evaluated Savings Estimates Are Applied—Retrospective or Prospectively

Estimates of costs and savings from efficiency measures typically are made both prior to program implementation (i.e., projected savings) and after program implementation (i.e., claimed and evaluated savings). As would be expected, evaluated estimates of savings are considered a more accurate representation of actual savings than projected savings.[12] Thus, when an independent evaluator uses M&V and comparison group–based evaluation methods to determine impacts, the evaluated savings usually are given precedence and are the basis for reported values.

RETROACTIVE VERSUS PROSPECTIVE SAVINGS CALCULATION EXAMPLE—DELAWARE FRAMEWORK

Changes in deemed energy savings or other deemed assumptions that result from program evaluation shall not be applied retrospectively, but shall be applied to the program and portfolio prospectively in the next program cycle.

Changes to deemed savings assumptions shall be coordinated through the annual process of updating the Delaware TRM.

Source: Delaware (2017)

An issue arises, however, when the deemed savings method with deemed savings values is used to project, claim, and perhaps evaluate energy savings for an efficiency measure in a given program year (e.g., based on per-unit savings values in a TRM approved for that program year), *but an evaluation during that program year indicates that the TRM per-unit savings values are too high or too low for the subject measure.* The question thus becomes "Should the newly evaluated deemed savings value be adjusted retroactively for the current program year or only applied on a going-forward basis?" Consider the following example.

[12] Differences between claimed and evaluated savings also can be due to the timing of an evaluation, as the evaluated savings might no longer reflect the conditions that existed when a measure was planned or installed. If an evaluation is conducted a year after an industrial project is completed, for example, and operational conditions have changed over that year, then claimed savings legitimately could be different from evaluated savings. Although evaluation protocols should be designed to account for such changes, it is not always possible to do so if just a "snapshot" evaluation of impacts is completed.

- TRM per-unit deemed savings values developed in 2016 indicate that the savings from measures verified to have been installed in a 2017 program are 10,000 MWh.
- New data, however, indicates that the per-unit savings values in the TRM are high, and that the actual program savings are more likely to be 9,500 MWh, based on the same number of units verified to have been installed in 2017.
- Assuming the verification is correct and all the measures were installed, does the oversight body (e.g., a regulatory commission) credit the program with 10,000 MWh of savings or only 9,500 MWh for the subject program year?

There are accuracy and equity issues associated with the above options for how and when to apply updated TRM values. On one hand, the program administrator/implementer relied on an approved value for budgeting and savings estimates. For resource planners and other stakeholders, however, what matters most is the most-accurate indication of what occurred. Although no perfect solution exists, the policy on whether to apply the result of an EM&V study retroactively or prospectively[13] already might be determined by commission order or other regulatory guidance. Even if this is not the case, this issue is best decided upon and included in an evaluation framework document before it occurs during the portfolio cycle.

ALIGNING SAVINGS WITH APPLICATIONS—ONE OPTION

Figure 1.3 shows that each phase of efficiency development cycle—planning, implementation and evaluation—employs a different type of savings estimate. In considering whether evaluated savings should be applied retroactively or prospectively, it is useful to consider how these estimates will be used in each phase. In the planning phase, *projected* savings are used to establish program goals. Using projected savings—adjusted for actual program participation (i.e., *evaluated* savings)—to assess whether a program met its goals in the implementation phase seems appropriate. This holds program implementers accountable for actually installing the number of measures envisioned in the plans.

When *evaluated* savings are greater or less than *projected savings,* these revised estimates can be applied prospectively to set the program goals for the next year (or cycle) because these savings now represent the best estimate of impacts. Evaluated savings, as the best estimate of impacts, also can be applied both retroactively and prospectively to assess the actual energy and demand impacts of programs on the utility system (and emissions), to calculate lost revenues, cost-effectiveness, and utility load forecasting and resource planning.

Lastly, EM&V frameworks in jurisdictions that review and update their TRMs annually or within a program cycle should include evaluation schedules (or criteria for establishing evaluation schedules) that address the need to coordinate evaluations with TRM updates. If, for example, impact evaluation results are to be used to update values in a TRM, then those evaluations should be completed on a schedule that allows sufficient time for stakeholders to review the new value before inclusion in a TRM.

3.2. EM&V Scope—Topics Recommended for Inclusion in All EM&V Frameworks

3.2.1. Metrics

Metrics are indicators of the performance of a specific portfolio, program, project, or efficiency measure. They typically are thought of as associated with energy or costs savings, but can include a wide range of information that supports stakeholder needs—such as those of regulators and policymakers, utility system planners, and program designers. EM&V frameworks should include a description of which metrics are of interest and are to be assessed and reported; and should indicate their relative importance (i.e., priority in case of budget or other constraints). As a good practice, frameworks also should specify that—like any estimate—impact metrics are to be reported as expected values with an associated level of variability.

[13] Although this decision mostly is associated with the use of the deemed savings method, the issue can also be raised for other EM&V methods (M&V and comparison group EM&V), such as per the example provided in footnote 12.

The process for establishing metrics can include:

- Defining the critical functions or needs of stakeholders and the efficiency portfolio information required to support those functions,
- Identifying specific, quantifiable information outputs from evaluation activities, and
- Establishing granularity requirements for the metrics, typically time (e.g., hourly versus annual values) and locational (e.g., statewide versus local).[14]

Two important points about identifying specific, quantifiable metrics are that:

- They are most useful when they can be compared against established targets for the metrics (e.g., savings goals); and
- There is tendency—which perhaps should be avoided—to define metrics for what evaluators find easily quantifiable, even if other metrics are or can be equally or more important.

EXAMPLES OF METRICS—HAWAII FRAMEWORK

- Gross energy savings at generation, which is the primary metric counting toward the Energy Efficiency Performance Standard (EEPS) goals. First-year, lifecycle, and cumulative savings are assessed and reported.
- Free-ridership, which is used if needed to inform program design and to facilitate decisions necessary for limiting potential double-counting of savings from different EEPS activities.
- Persistence, measure life, and expected lifetime for each activity, which is employed in the calculation of first-year, lifecycle, and cumulative savings, costs, and benefits.
- Participant and non-participant spillover.
- Market effects.
- Implementation and administration costs.
- Co-benefits of energy efficiency savings, such as greenhouse gas reductions, job creation, and other benefits resulting from EEPS activities.
- Other metrics as recommended by the Technical Working Group or at the request of the commission.

Source: Public Utilities Commission of Hawaii (2012)

Table 3.2 provides some examples of metrics often associated with evaluation of efficiency portfolios.

[14] Time and location metrics usually are associated with energy and environmental (e.g., emissions) impacts.

Table 3.2. Example Efficiency Evaluation Metrics

Metric Category	Example Metrics—Major Categories	Example Metrics—Specifics
Energy Impacts As used in common practice, energy impacts are defined as those directly or indirectly associated with reductions in energy consumption, demand, or both.	• Gross energy savings, annual and lifetime • Net energy savings, annual and lifetime • Gross demand savings • Net demand savings • Utility system benefit	• Electricity savings: kilowatt-hour saved per year and per month • Demand savings (example 1): kilowatt-hour saved per month of each year of program, averaged over peak weekday hours • Demand savings (example 2): kilowatt-hour savings coincident with annual utility peak demand, reported for each year of the program • Avoided transmission and distribution costs • Natural gas savings: MMBtu saved per year and per month • Lifetime savings (savings that occur during the effective useful life of the efficiency measure): MWh or MMBtu saved during the measure's lifetime, in years. • Energy price effects • Reliability effects
Non-Energy Impacts Non-energy impacts are the wide variety of positive and negative effects beyond energy and capacity savings.	• Participant benefits • Societal benefits	• Avoided emissions (example 1): metric tons of CO_2 and SO_X avoided during each year of the program • Avoided emissions (example 2): metric tons of NO_X avoided during ozone season during the months of each year of the program • Decreased customer energy or other (e.g., water) costs • Increased property values • Improved comfort/indoor air quality • Reduced equipment operations and maintenance

Metric Category	Example Metrics—Major Categories	Example Metrics—Specifics
		(O&M) costs because of longer-lived measure relative to baseline • Jobs and local economic development
Cost-Effectiveness	• Cost benefit ratios such as defined in National Standard Practice Manual (NSPM)[15] • Cost of saved energy metrics[16]	• Jurisdiction-specific Resource Values Tests from NSPM • Levelized total costs of efficiency • Participant, societal, ratepayer, and administrator benefits and costs • Utility system benefits and costs • Societal benefits and costs
Market Transformation/ Adoption Market transformation is a reduction in market barriers resulting from a market intervention.	• Market effects that are likely to last after the intervention has been withdrawn, reduced, or changed.	• Total market or jurisdiction-wide unit energy consumption reductions • Supply chain adoption and growth (including reduced cost of energy-efficient products and energy efficiency services) • Participant and non-participant spillover[17] • Increased availability and purchase activity for energy-efficient products and energy-efficiency services, and reduced time for sales/installation). • Greater consumer awareness and confidence in efficiency benefits

[15] National Standard Practice Manual for Assessing Cost-Effectiveness of Energy Efficiency Resources (Edition 1, Spring 2017). https://nationalefficiencyscreening.org/national-standard-practice-manual/.

[16] For information on these metrics see: https://emp.lbl.gov/projects/what-it-costs-save-energy.

[17] Spillover is reductions in energy consumption or demand caused by the presence of an efficiency program, beyond the program-related gross savings of the participants and without direct financial or technical assistance from the program. There can be participant and non-participant spillover.

Interim Performance/Milestones Interim performance and milestone indicators are associated with the implementation of and compliance with efficiency program implementation.	• Marketing and outreach metrics • Consumer awareness and support • Workforce education and training metrics	• Number of participants (consumers/building owners, vendors, contractors, architects/engineers) • Characteristics of participants and non-participants • Penetration of programs into the eligible market (i.e., how many customers are participating, by size or other characteristics) and the average savings per participant • Initiation of program implementation policies and expenditures as well as evaluation activities

The granularity of most metrics, such as the locational and time-specific impacts of efficiency impacts on electricity demand, has become increasingly important. Therefore, it might be necessary in an EM&V framework to address metric time[18] and location (e.g., statewide, utility system wide, or specific to particular nodes on an electric grid) granularity reporting expectations. In the case of electricity, the time granularity of evaluation analyses relates to whether 15-minute, hourly, monthly, seasonal, annual, or lifetime data collection and savings reporting are required. For demand savings, the choice of a definition of the metric is particularly important (e.g., annual average, peak summer, coincident peak).

The "granularity decision" should be determined by how the information from the evaluation is to be used. Annual savings data generally only are useful for an overview of the program impacts. More detailed data, in particular lifetime impacts, usually are required for cost-effectiveness analyses, demand forecasting, and resource planning. For avoided emissions, annual values are typical; for certain programs, however, such as smog or greenhouse gas–emission mitigation programs, there are specific seasons or periods of interest either due to their coincidence with air pollution problems or variations in the power system emissions profile.

[18] For information on the time value of efficiency see: https://emp.lbl.gov/publications/time-varying-value-electric-energy.

3.2.2. Evaluation Reports and Other Outputs

Section 2.3 includes an annotated list of the typical plans and reports, or products, that are produced in the course of the EM&V activities. The list can be the starting point for defining which products (deliverables), are to be completed. Not included in the Section 2.3 lists are communications that are associated with typical project management activates—for example, the status reports prepared by an evaluator to indicate the status, progress, and issues associated with implementation of the evaluation; these are not usually described in a framework document.

The framework document can indicate, perhaps in just a table, the major, expected products. Also included in this portion of the framework (or elsewhere), should be an indication of which entities (e.g., implementer, evaluator, administrator, regulator) will prepare the major products, who will review and approve them, and schedules. More detailed frameworks also can indicate descriptions of the products with expectations for what they will cover, their purposes, and expected level of documentation. The framework also can include outlines for the plans and reports (*see* Section 3.5.2).

EXAMPLE CONTENTS DESCRIPTION FOR ANNUAL SAVINGS REPORT—MARYLAND FRAMEWORK

The annual net savings analysis reporting should include (at a minimum):

- A summary of methods, noting any material changes from the previous year in primary data-collection instruments and processes;
- Key algorithms or assumptions;
- Consolidated tables of values and results, including gross savings, NTG ratios, and net savings;
- Discussion of findings, including any anomalous results or significant year-to-year changes;
- Sources of uncertainty, including efforts to avoid, mitigate, or adjust for potential sources of bias, relative confidence, and precision associated with primary data collection, and disposition of omitted data;
- Recommendations pertaining to evaluation and program design or implementation; and
- Links or citations to anchor documents containing background discussion of methods and assumptions.

Source: Maryland (2017)

3.2.3. Scale of the Evaluation Effort (Budget)

A critical element of a framework document is defining an overall budget or providing budget guidance (i.e., setting criteria for selecting the funding level) for evaluation activities—impact, process, market, and cost-effectiveness evaluations. Establishing a budget for an evaluation requires consideration of all of the aspects of the evaluation process and balancing the trade-offs between (a) the costs, quality, and timeliness expectations for the evaluation activities (as discussed elsewhere in this guide as an overall theme of the guidance to be provided in the framework document); (b) the relative certainty associated with various best-practices, impact evaluation methods, and approaches (*see* Section 3.3); and (c) the value of the information generated by the efforts.

Budgets also can be influenced (increased) to accommodate support efforts aimed at assessing and reducing evaluation error, such as to pay for additional short-term metering, training of staff, or testing of questionnaires and recording forms to reduce data-collection errors. When sampling is employed, which is quite common, the determination of the appropriate sample sizes, based on precision and confidence expectations for the resulting values, also can be a major factor in setting an evaluation budget.

Table 3.3 provides some guidance with respect to the impact of several key issues on EM&V budget. In general, on a unit-of-saved-energy basis, costs are inversely proportional to the magnitude of the savings (i.e., larger projects have lower per-unit evaluation costs) and are directly proportional to uncertainty of predicted savings (i.e., projects with greater uncertainty in the predicted savings warrant greater EM&V costs).

Table 3.3. Issues that Affect Impact EM&V Budgets: General Guidance

Budget Affecting Issue	Commentary	Tends to Increase EM&V Budgets (As Percent of Program Budget)	Tends to Decrease EM&V Budgets (As Percent of Program Budget)
Program Goals and Budget	How large is the program in terms of budget and savings goals? Larger programs tend to have larger evaluations but smaller evaluation costs as a percentage of program expenditures.	Programs with large budgets and large goals	Programs with small budgets and small goals
Impact (e.g., Savings) Uncertainty	What is the level of uncertainty associated with the expected impacts of a program, and what is the risk that the program poses in the context of achieving (or not achieving) portfolio goals.	More uncertainty and greater impact on overall portfolio impacts	Less uncertainty and lesser impact on overall portfolio impacts
Program History of Known Impacts	Is it a new program with uncertain impacts or an established program with well-understood impacts?	New programs without historic, documented impacts	Established programs with a history of well-documented impacts
Use of Deemed Savings Method	Is it adequate to simply verify that the individual projects in a program were installed (and perhaps operating correctly) because of the extensive use of deemed savings values? Or, on the other end of the cost spectrum, are rigorous field inspections, data collection, and analyses on all—or a large sample of—projects in a program required?	Programs using deemed savings methods, if there is a substantial need to update inputs for deemed values (i.e., hours of use) Programs using ongoing measurement and verification and comparison group methods	Programs using deemed savings methods, if there is *not* a substantial need to update inputs for deemed values Programs using just short-term measurement and verification on a sample of projects
Regulatory Energy and Demand Savings or Emission-Reduction Goals	Is there a savings goal, perhaps associated with an energy efficiency resource standard, for which the impact evaluation will determine compliance?	Regulatory mandates to achieve defined energy or demand saving amounts or emission reductions	No regulatory mandates
Administrator Performance-Based Incentives	Are there cost-recovery or lost-revenue recovery financial impacts for utilities, or does the program administrator have a performance-based incentive, to be decided upon by a regulatory body using the impact evaluation results?	Administrator has performance-based incentive(s)	No administrator performance-based incentives

Conceptually, the trade-off process is shown in Figure 3-1. The goal is to find the balance point between increasing incremental investments in evaluation (costs) and decreasing incremental value in the evaluation information. That is, deciding how to derive the most value from the evaluation activities at the lowest cost and in the timeliest manner, or put even another way, deciding "How good is good enough?"

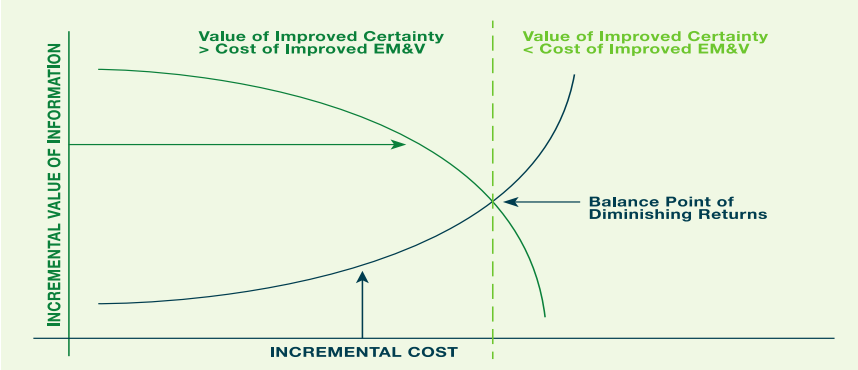


Figure 3.1. Incremental value of information versus incremental cost of evaluation (Schiller 2012)

Section 1 introduced the concept of EM&V being a form of risk management. Thus, part of evaluation planning is deciding what level of risk is acceptable and determining the requirements for accuracy and a corresponding budget in the framework. To reach this balance point, ideally there should be a process whereby evaluation methods and approaches are defined, budgeted, and then refined in an iterative process. In practice, however, evaluation budgets typically are set based on impressions of what is appropriate or examples from the past evaluations or other jurisdictions. Values usually are presented as an annual dollar amount or as a percentage of overall program budgets. If presented as guidance, a range could be indicated (e.g., 3% to 6% of portfolio budget) and allocated across programs and evaluation types, such as impacts, process, or market studies, as needed, with authority for final budgeting assigned to a regulator or administrator. Although it is difficult to generalize, common practice suggests that a reasonable spending range for evaluation (impact, process, and market) is 3% to 6% of a portfolio budget.[19] This should be considered rough guidance, however, because evaluation needs and the relative EM&V roles of program administrators and independent third-party evaluators (and thus how the budget is categorized between program and evaluation expenses) vary significantly between different states and different program administrators.

EXAMPLE BUDGET GUIDANCE—MASSACHUSETTS FRAMEWORK

The EM&V budget available to the research areas for the 2016–2018 Plan is projected to be in line with historical program budget levels. Twenty percent of each sector's available evaluation budget is allocated to the Cross-Cutting research area. The remaining evaluation budget in the residential and low-income sector is allocated to the Residential research area; the remaining evaluation budget in the C&I sector is allocated to the Non-Residential research area. Total evaluation budgets for the 2016-2018 Plan term are expected to be $18.7 million for gas programs and $41.3 million for electric programs.

Source: Massachusetts Department of Public Utilities (2015)

3.2.4. Timing of the Evaluation Cycles and Reporting

The evaluation time frame has several possible major components, some or all of which can—and probably should—be defined in a framework document.

- The period over which the evaluation activities will take place—
  - Will evaluation activities and reporting be based on an annual schedule or be tied to the efficiency portfolio cycle—which might be multi-year?
  - Will the evaluation assess persistence of impacts over the expected lifetimes of the implemented efficiency measures?

[19] Schiller 2012.

- How often each program will be evaluated (e.g., every program year or once every few years with results for years in between based on the most recent evaluation)?
- What is the reporting schedule? That is, when are reports and support document due? For example, when, after the end of a program year, is the evaluation due to be completed and publicly reported in a manner that supports timey feedback and the next program/portfolio planning cycle.

Timing is one of the three major components of the trade-offs mentioned throughout this guide. The other components being rigor and budget—very simply, it can be said that the more budget or less rigor could result in quicker results, and vice-versa. Beyond budget and rigor, however, there are several other factors that should be considered when establishing a reporting schedule in an EM&V framework document. These include:

- The length of the portfolio cycle (e.g., one, two, or three years);
- The desire or requirements to have early feedback for program implementers, and whether and when information is needed for a next portfolio cycle;
- Program lifecycle stage (evaluating a first-time program or a long-established program);
- Evaluation data collection time lags (e.g., sufficient post-installation billing records to estimate savings);
- Whether to estimate persistence of savings or to conduct ongoing verifications and analyses of persistence;
- Regulatory and management oversight requirements including time required to select and contract with the evaluator;
- Contractual requirements for reporting savings for "pay for performance" programs; and
- Timing requirements to use the evaluation results to update energy and demand savings as well as measure life estimates for specific efficiency measures, such as used for updating a TRM.

Within a framework document, each of these topics can be considered in depth or very simply, as is done in the Maryland Framework[20] with this very general sentence for net savings evaluations (that provides for a great deal of flexibility, but not a lot of guidance): "evaluations should be conducted with sufficient frequency to reflect changing markets, technologies, program designs, and other conditions." Conversely, also within the 2016 Maryland Framework is a table with a relatively detailed evaluation schedule, including major evaluation milestones(*see* Figure 3.2).

An ideal evaluation schedule begins before the start of the program's implementation (to collect any baseline data and define the EM&V work plan, including data-collection needs) and continues for some time after the program is completed to analyze persistence of savings. The actual timing of evaluation activities, however, is influenced by several—often both practical or competing—considerations. Thus, the first timing consideration to address in a framework probably is *when* to start the evaluation efforts. Programs and portfolios tend to get into a regular cycle; for example, the start of each program year could be on January 1, but the evaluation process might not get started until the spring.

[20] Maryland 2016, page 4-1. The 2017 version of the Maryland Framework does not present the schedule in a figure format. Currently, both versions are not available online but can be requested from Maryland Public Service Commission staff.

One approach in this situation is to accept the late start of an evaluation for the current year and conduct a more limited effort than perhaps desired, but then move the evaluation cycle and the program/portfolio cycle into better alignment within one or two years. This is an approach that was taken in Maryland for 2015 and 2016, as shown in Figure 3.2.

EXAMPLE EM&V FREQUENCY GUIDANCE—NEW YORK FRAMEWORK

Early program EM&V efforts should focus on process-related issues to serve as an early-warning system, especially for new initiatives. This approach can be used to determine whether the program is operating smoothly and is responsive to participants/market needs, and to identify opportunities to make improvements that can reduce costs and increase program effectiveness. This guidance does not establish a rigid timetable for process evaluation and impact evaluations. Generally, evaluations focusing on verifying program-level energy savings cannot be completed until a sufficient number of projects have been completed and post-installation operations can be observed. A typical program evaluation timetable includes a process evaluation in program year 1, and an impact evaluation in program year 2 or 3, with an emphasis on obtaining results as soon as is reasonably possible. For programs that have undergone recent full-scale evaluation, repeating such full-scale evaluation might not be necessary every few years, but rather the frequency and scope of the EM&V activity should be based on results of evaluations to date, information gaps, and other areas requiring analysis. Given the dynamic and transitional nature of current program offerings, shorter and more-targeted EM&V activities can be valuable for investigating issues that a typical program evaluation would not cover. Targeted evaluations can be initiated and completed at any time and could serve to better align and serve the planning process for the current suite of clean energy programs. Concurrent evaluation is also an option that can accelerate M&V data collection, and is a credible technique for assessing decision making closer to its time of occurrence while still vivid in the respondent's mind, and enables more expeditious feedback regarding program performance.

Source: New York (2016)

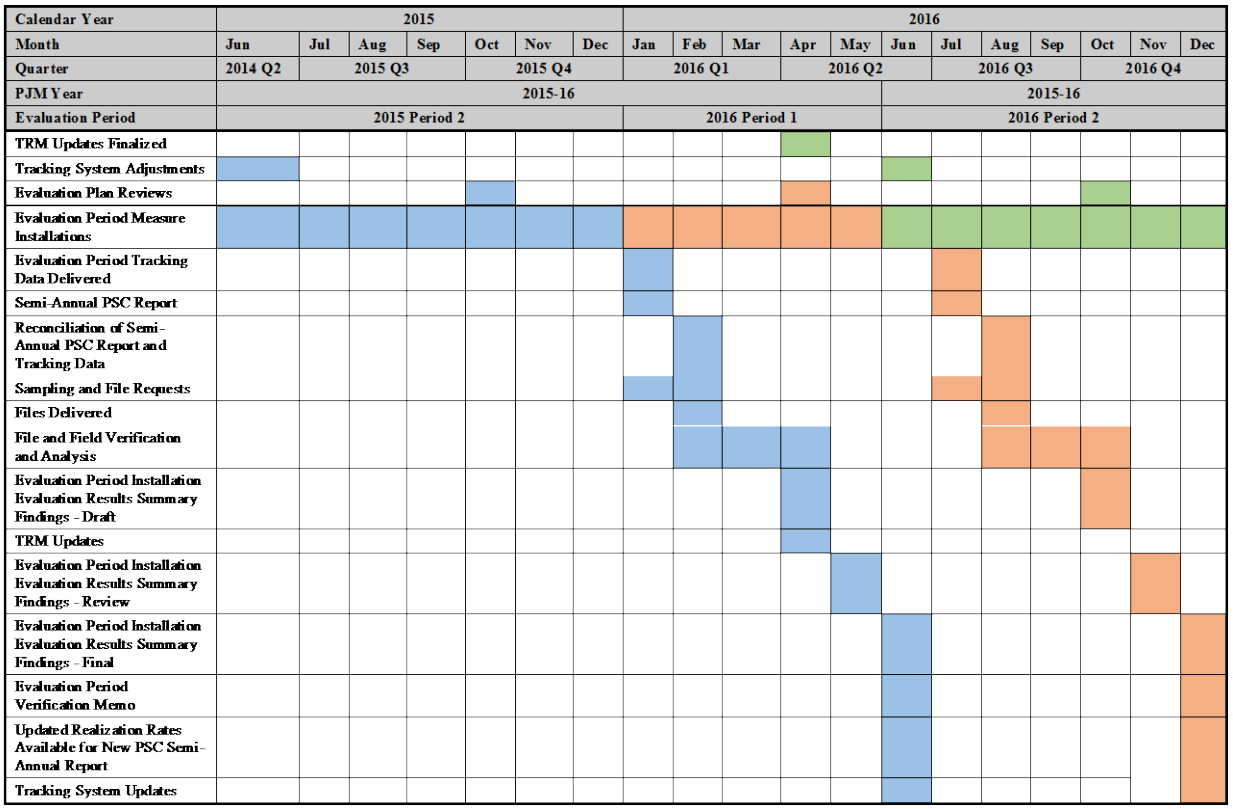


Figure 3.2. EmPOWER Maryland evaluation schedule (Maryland 2016)

3.2.5. Roles and Responsibilities

This component of the framework document should address who will conduct the evaluations and ideally define high-level descriptions of the roles and responsibilities associated with evaluation. Although evaluation is relevant to many kinds of efficiency portfolios, the importance of defining the role and independence of the evaluator is primarily an issue for programs that are funded with public or energy-consumer funds. As with the rest of this guide, this subsection addresses this situation in which a government regulator oversees a program administrator's conduct of efficiency activities and wants some level of independent assessment of program/portfolio impacts.

As discussed in Section 1, with impact evaluation—particularly with regard to determining energy and demand savings—there are two possible sets of results reported after a project or program is implemented: claimed savings and evaluated savings. Staff or consultant evaluators (e.g., engineers, analysts, econometricians), working with the program implementers and administrators, almost always prepare these claimed savings reports. Their role is to directly support the program implementation and prepare required internal and external (e.g., regulatory) reporting.

Conversely, evaluated savings are only required if some entity—such as a government regulatory agency—wants an independent third-party evaluator to either conduct a new impact evaluation or double-check the claimed savings that are provided by the implementer or administrator. This leads to several resulting questions that can and should be addressed in the EM&V framework document.

- Who is responsible for which evaluation activities?

- What are the qualification requirements for the third-party evaluator?[21]
- Will the evaluator, in preparing evaluated savings, simply confirm the work done by the implementer or administrator in preparing the claimed savings (verification)? Or will the independent third-party evaluator conduct some data collection and impact evaluation analyses?
- What are the relative EM&V–related responsibilities (such as providing data or reviewing evaluation reports) of implementers, administrators, evaluators, regulators, and other stakeholders?
- Is an independent third-party evaluator required and, if so, what is meant by "independent" and "third-party," and who selects and retains such an evaluator?

Regarding the last question listed above, there are no formal definitions of independent or third-party evaluator in the efficiency industry and the hiring entity could be the regulator or the administrator, or perhaps another entity. In general practice, however, "independent third party" is considered to mean that the evaluator has no financial stake in the evaluation results (e.g., magnitude of savings) and that its organization, its contracts, and its business relationships do not create bias in favor of or opposed to the interests of the administrator, implementers, program participants, or other specific stakeholder groups. In practice, different states' regulatory bodies have taken different approaches to defining the requirements for evaluators who are asked to review the claimed savings and prepare evaluated savings reports, and to who hires that evaluator—although, in many jurisdictions, the evaluator reports to a regulatory entity and not the program administrator or implementer, or the regulator has the final approval of evaluation reports.

EXAMPLE ROLES AND RESPONSIBILITIES—HAWAII FRAMEWORK

- The [Hawaii Public Utilities] Commission retains the right to review and approve all EEPS [Energy Efficiency Performance Standard] EM&V activities and adjust any EM&V results and any information resulting from EM&V.
- Contributing entities should provide the EEPS Reporting Contractor and EEPS EM&V Contractor with estimates of program savings goals and designs, EEPS EM&V plans, and implementation results as available. The EM&V plans and results will be reviewed by the EEPS EM&V Contractor.
- The EEPS EM&V Contractor will provide EM&V assistance to the Commission to support the EEPS.
- The EEPS Technical Working Group may provide review and feedback on all aspects of EEPS EM&V, as needed, including but not limited to, all documents, plans, and assumptions regarding ex ante and ex-post EEPS EM&V, acceptable methods and approaches, prioritization of EEPS EM&V resources, and compilation of EEPS portfolio-wide results.

Source: Public Utilities Commission of Hawaii (2012)

Irrespective of how the relationships are determined or who hires whom, the objective is for all parties to the evaluation to believe that the reported results are based on valid, unbiased information that is sufficiently reliable to serve as the basis for informed decisions.

Having all the evaluation data collection and analyses conducted independently by a third party, at least in theory, provides the greatest level of due diligence and integrity for evaluated savings values. Such analyses do add costs, however, and it is not uncommon for there to be repetition in the determination of savings—between the claimed and evaluated savings determination efforts. Also, because a common objective of evaluation is to help with program improvement, a totally independent approach does not directly favor a tight working relationship between the evaluator and the implementer/administrator. Thus, the selection of an evaluator can require balancing evaluation independence (so that the evaluation is considered objective) with the desire to have the

[21] There are no licensing requirements for efficiency evaluators. However, there is a wide range of training opportunities, including the Efficiency Valuation Organization's Certified Measurement and Verification Professional program (http://evo-world.org/en/products-services-mainmenu-en/certification-mainmenu-en), and classes offered at the International Energy Program Evaluation Conferences (https://www.iepec.org/) and by the Association of Energy Services Professionals (http://www.aesp.org). Additionally, the U.S. Department of Energy has initiated an EM&V certification effort (*see* Northeast Energy Efficiency Partnership 2016).

evaluator close enough to the process that the evaluation provides ongoing and early feedback without the implementer feeling inappropriately "defensive."

One way to look at the relative roles of the different entities involved in preparing claimed and evaluated savings (as well as project savings) is to consider the roles as associated with either oversight or administrator activities. These can be generally defined as follows.

- **Oversight activities.** Activities that are under the purview of the entity responsible for all efficiency programs and the associated EM&V implemented by program administrators in the subject jurisdiction (e.g., state or utility service territory). Oversight activities usually include coordination with the government/regulatory authority. They also might include feedback or guidance to and from stakeholders (including administrators and implementers) about the evaluation plans and implementation as well as the process of approving reported results.
- **Administrator activities.** Activities undertaken by the program administrators during the process of developing, implementing, and conducting EM&V activities pertinent to their implementation of efficiency programs. These EM&V activities also might be known as the *primary evaluation activities*.

During the development of a jurisdiction's EM&V framework it is essential that these roles and responsibilities be determined. One example is how Maryland defined these roles in its framework via a simple matrix, included as a table as part of the framework document.[22] This is only an example, however, not necessarily a recommended allocation of duties.

3.3. Impact Evaluation Methods—Topics Recommended for Inclusion in All EM&V Frameworks

This section discusses aspects of the impact evaluation approaches—such as methods, specific approaches, and baselines—that should be addressed in a framework document. Note that these topics can become quite technical, thus professional input is valuable for assisting decision making by all the stakeholders as framework topics and content are discussed.

Again, as touched on throughout this guide, providing direction within a framework often is about finding a balance between cost, accuracy, and timeliness. At the beginning of this section, an example is provided of two hypothetical states and how their requirements (why, what, when) can affect the selection of EM&V costs, rigor, and timing. Figure 3.3 graphically presents this concept of trade-offs at the beginning of this section because one of the most important components that should be addressed in all EM&V frameworks is guidance on required or allowable evaluation methods and the reliability (accuracy, rigor) of the evaluations—and that guidance is heavily driven by these trade-offs.

[22] Maryland 2017. See Appendix B, June 24, 2009, Consensus Report on the EmPOWER Maryland EM&V Process.

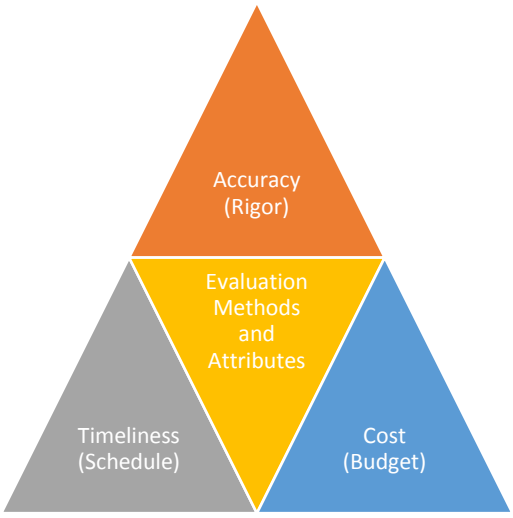


Figure 3.3. Using trade-offs to select evaluation attributes and approaches

3.3.1. EM&V Methods that Will/Can Be Used

This portion of the framework defines the EM&V methods and related approaches that can be used or are expected to be used for assessing the efficiency portfolio.

As with other sections, the level of detail of EM&V method and approach specification can vary based on needs and interests of those compiling the framework. This could be where the framework authors decide which general methods are appropriate or inappropriate (perhaps because they are not considered rigorous enough) for particular programs in the efficiency portfolio. One particular method issue that can be addressed is when the deemed savings method (e.g., using deemed savings values, factors, variables, and calculations from an approved TRM) is allowable; because the deemed savings value approach often is perceived as a low-cost method it can be the go-to method for most programs if not specified otherwise.[23]

EXAMPLE OF ALLOWABLE APPROACHES—HAWAII FRAMEWORK

Allowable approaches to [Energy Efficiency Performance Standard] EEPS EM&V shall be based on best practices, as articulated in nationally-recognized documents and protocols at the time the EEPS EM&V is undertaken. A range of approaches may be used. Rigor and precision levels for EEPS EM&V shall be determined within EEPS EM&V workplans, and shall balance best practices with the value of information, uncertainty, and resource availability.

Source: Public Utilities Commission of Hawaii (2012)

The framework document, however, might simply indicate that all industry standard methods can be used with the actual method selection delegated to the third-party evaluator, or perhaps the evaluator and an advisory or oversight group. For example, this portion of a framework document could indicate that Uniform Methods Project (UMP) protocols, and International Performance Measurement and Verification Protocol (IPMVP) methods should be applied and then other protocols and guidelines be applied as needed.[24] With this more minimal guidance, the evaluator prepares evaluation planning documents (*see* Section 2.3.1) and these indicate specific EM&V methods with some detail on the approaches to be used, as well as the details of how the selected method was selected

[23] Even in the case where only deemed savings are to be used, verification of the number of measures installed and the integrity of program tracking systems must occur. This could require a sampling design and decisions regarding how to verify (e.g., surveys, site visits). As noted in the budget section, although the deemed savings method is perceived as low cost, it does still require efforts to ensure the deemed savings values, factors, and other elements are updated and accurate.

[24] See appendices A and B for descriptions and references to UMP and IPMVP.

and is applied. This is a fairly typical way to approach EM&V planning, because—as discussed in Appendix A—evaluators have a variety of EM&V methods and related approaches at their disposal, and they need to match them to the program type and status (e.g., pilot or established), jurisdiction needs, regulatory requirements, and resources available.

M&V 2.0

In the continuing effort to improve EM&V methods, an approach for potentially reducing costs, uncertainty, and delays in obtaining results is the use of real-time or continuous M&V based on short-interval (hourly, daily) metering this approach is known as M&V 2.0. M&V 2.0 has been defined as "[t]he leveraging of smart grid investments, advances in interval meter data, nonintrusive load monitoring, and equipment-embedded sensors and controls to provide new tools with potential to reduce the cost of M&V, produce more timely results with higher confidence and transparency, and thereby increase the acceptance of the savings calculations." (Granderson et al. 2015). In practice, M&V 2.0 describes recent advances in metering (e.g., advanced metering infrastructure [AMI]) and monitoring (e.g., wireless sensors, smart thermostats), data availability and analytical tools (e.g., machine learning, interactive visualization, cloud-based analytical platforms) associated with documenting the energy and demand savings from specific energy efficiency measures or projects based on consumption data. EM&V frameworks could reference M&V 2.0 as a viable approach for determining impacts.

Not all of these methods and approaches will be appropriate in every situation, and each comes with its own unique combination of costs, timing impacts, data requirements, and resulting rigor. Thus, an EM&V framework document can just provide criteria for selecting EM&V methods. For example, large behavior-based programs[25] might be required to use randomized control groups to determine impacts, whereas programs with well-defined and understood prescriptive measure might be allowed to use deemed savings values from an approved TRM.

Alternatively, the framework document can indicate the stakeholders' perceptions of the trade-offs between the methods and related specific approaches. For example, Figure 3.4, for the impact evaluation activities mentioned in Appendix A, Section A-3, indicates one firm's heuristic estimate of the relative accuracy, timeliness, and cost effectiveness of these different methods and approaches to documenting impacts, which is illustrative of the trade-offs that should be considered when selecting methods. Having information available on trade-offs between evaluation methods can help stakeholders both understand and provide input into the selection of evaluation methods.

EXAMPLE OF BASELINE GUIDANCE—TEXAS FRAMEWORK

Baselines for preparation of TRM deemed savings values or deemed savings calculations or for other evaluation activities shall be defined by the EM&V contractor and commission staff shall review and approve them. When common practice baselines are defined for determining gross energy and/or demand savings for a measure or program, common practice may be documented by market studies. Baselines shall be defined by measure category as follows (deviations from these specifications may be made with justification and approval of commission staff).

(A) Baseline is existing conditions for the estimated remaining lifetime of existing equipment for early replacement of functional equipment still within its current useful life. Baseline is applicable code, standard or common practice for remaining lifetime of the measure past the estimated remaining lifetime of existing equipment.

[25] Behavior-based programs are those that utilize strategies intended to affect consumer energy use behaviors to achieve energy and peak demand savings. Examples include programs that compare participant energy use with their neighbors', providing real-time information and feedback about energy use, and participant goal setting and reward points per unit of energy saved.

(B) Baseline is applicable code, standard, or common practice for replacement of functional equipment beyond its current useful life.

(C) Baseline is applicable code, standard, or common practice for unplanned replacements of failed equipment.

(D) Baseline is applicable code, standard, or common practice for new construction or major tenant improvements.

Source—Public Utilities Commission of Texas (2017)

EVALUATION METHODOLOGY	ACCURATE	TIMELY	COST EFFICIENT
Tracking System Review	○	●	●
Engineering Review	◒	●	◒
Customer Surveys	◒	◒	◒
Onsite Verification	◒	◒	◒
End-Use Metering	●	○	○
Calibrated Building Energy Simulation Modeling	●	◒	◒
Site-Specific Energy Simulation Modeling	●	◒	○
Billing Analysis	◒	●	◒
Survey-Based Approach	◒	◒	○

Figure 3.4. Impact evaluation activities comparison (Malinick et al. 2017)

3.3.2. Baseline Definitions

Impacts, such as energy and demand savings, caused by an efficiency activity typically are defined as the difference between:

- The numerical value of the metric of interest (e.g., kWh of energy use) with the efficiency activity in place, and
- What the value of the metric would have been absent that activity during the same period and under the same operating conditions.

As discussed in Appendix A, and more extensively in the industry references listed in Appendix B, what would have happened without the efficiency activity typically is called the baseline (or, more specifically, the "counterfactual scenario").

If there is a single technical element that is important to define in every EM&V framework document, it is establishing the definitions for the allowable baselines against which energy savings and other impacts are determined.

CONSISTENT BASELINES FOR DIFFERENT USES

Providing guidance on how baselines are selected in a framework document is important because defining baselines is a key (if not the key) challenge for determining the numerical values of any reported metrics, particularly energy and demand savings. If the conditions, equipment, and operations prior to the installation of an efficiency project or measure (known as pre-existing conditions) always were the baseline, then this determination would be relatively straightforward. However, many efficiency activities take place in a context of ongoing changes in consumer behavior, marketplace activities, regulation, technology, and project site operations—all of which can affect what has occurred in the absence of the efficiency activity. Specifying a baseline requires consideration of this context. When the efficiency activity involves new construction or new equipment installations subject to a building code or equipment standards, for example, those standards or code can affect the baseline definition.

In cases where quantitative efficiency goals (e.g., GWh, MW) are established, such as in a portfolio standard or emissions reduction program, the baseline assumed in the goals setting process can also affect baseline definitions for determining impacts. When impacts (e.g., energy savings, emission reductions) are to be determined via the EM&V process and compared against such goals there needs to be an alignment between the assumptions made for what was assumed would happen in the absence of the programs, i.e., the baseline used for establishing these goals, and baselines assumed in the evaluation. Otherwise the starting point for determination of accomplishment can be out of alignment with the evaluated impacts.

Table 3.4 provides a high-level summary of one approach for how applicable baselines can be defined in a framework for broad categories of programs. Alternatively, Table 3.5 has a more granular list of common efficiency activities and contextual situations with examples of the types of allowable baselines that could be defined in an EM&V Framework.

Table 3.4. Standard Practice Baseline Application for Common Program Categories (Schiller 2012)

PROGRAM CATEGORY FOR PURPOSES OF BASELINE DETERMINATION	EXISTING CONDITIONS BASELINE	CODES AND STANDARDS BASELINE	COMMON PRACTICE BASELINE
Early replacement or retrofit of functional equipment still within its current useful life Process improvements	X Existing conditions baseline for the remaining life of the replaced equipment or process	X C&S baseline for the time period after the remaining life of the replaced equipment	X Common practice baseline for the time period after the remaining life of the equipment
Replacement of functional equipment beyond its rated useful life		X	X
Unplanned replacement for (of) failed equipment		X	X
New construction and substantial existing building improvements		X	X
Non-equipment based programs (e.g., behavior-based and training programs)			X What people in a control group would be doing in the absence of the program

Table 3.5. Granular List of Common Efficiency Activities with Contextual Situations for Baseline Definition[26]

COMMON EFFICIENCY ACTIVITIES
• **Higher efficiency replacement**: Replacement of existing facility equipment or structural component (such as windows) with high-efficiency new equipment or component. • **Higher efficiency equipment in new installations**: Installation of high-efficiency equipment or structural components in new construction, major renovation, or other first installation of the equipment type that triggers a building energy code**.** • **Add-on efficiency**: Equipment or structural changes that can be added to facilities or equipment, such as insulation or controls. • **Operational or maintenance improvement:** Operational improvements such as adjusting set points or run times, or maintenance actions that improve efficiency without installation of new equipment affected by these improvements or actions. • **Combination EE measures installed as part of the same project:** Combinations of multiple EE measures (e.g., equipment replacement, operational improvement, add-on, new controls, building shell) that jointly affect the same systems. • **New construction or renovation at higher efficiency**: New construction or major renovation that triggers code, to produce a higher-efficiency performance building than required by code. • **New statewide equipment standards**: New statewide efficiency standards for manufacture or sale of particular types of energy-using equipment, setting a new mandatory minimum efficiency standard for a particular equipment type. • **Whole-building EE improvement**: Comprehensive assessment and improvements to building shell, equipment, or operations. • **Mass-market information and encouragement:** Provision of information and encouragement to adopt a wide variety of physical, operational, and behavioral efficiency improvements to large groups of customers. • **Building operations and maintenance training:** Provision of training to building operators on particular types of building operations and maintenance improvements.
CONTEXT FOR DEFINING BASELINES
• **Replace on failure:** Replace equipment at the end of its useful life with high-efficiency equipment. • **Early replacement:** Replace equipment prior to the end of its useful life with high-efficiency equipment. • Added to existing facility without concurrent equipment replacement (not triggering code). • Added to existing facility with concurrent equipment replacement (not triggering code). • Included with new construction/major renovation (triggering code).

3.3.3. Sampling Design

Selecting a sample of projects in an efficiency program to analyze and represent the program's entire population of projects is a common element of the evaluation process, although in some programs all of the projects can be selected for EM&V (a census). Examples of populations from which samples are selected and used in evaluation are participants (e.g., for process evaluation interviews, for determination of program-influenced savings), project sites (e.g., for on-site measurements and verification), and project claimed savings calculations (for engineering reviews). For individual projects' M&V, samples are also often selected; for example, one month could be selected

[26] From U.S. EPA Draft EM&V Guidance for Demand-Side Energy Efficiency, December 2016.

for metering to represent a whole year or a subset of motors (or light fixtures) that are retrofitted may be selected for inspection and metering to represent all the motors (or light fixtures) in a project.

Addressing sampling within a framework document is recommended even if it is just simply stating that sampling design is to be left to the professional judgement of the evaluator. This can be appropriate because sampling design can be complex and reliant on the specifics of a given program or project. Alternatively, and preferably, broad sampling criteria can be indicated.

SAMPLING GUIDANCE EXAMPLE—DELAWARE FRAMEWORK

Sampling approaches, sample-size targets, and confidence limits should provide the highest level of accuracy achievable balanced with the available resources. Large programs and programs that are important for reaching energy saving targets should have sampling approaches that reflect that importance. Low impact or smaller programs may have lower precision and confidence levels.

Source: Delaware (2017)

The most common criteria included in framework documents are those associated with the confidence and precision expected of samples (see Section 3.3.4 on certainty), but a framework document also can address process issues.

Some examples of process issues are listed below.

- Who will select the sample, the evaluator or the implementer?
- Are samples required to always be random, or can other approaches to sample selection be utilized?
- Is reporting of sample disposition required? What types of information should be reported, for example sample points omitted or replaced?
- What is the remedy if a sample is required to have a certain level of confidence and precision, but after selection and analysis of the sample it is determined that those requirements are not met? (Possible options are requiring a new sample to be drawn or correcting the sampling in future evaluations.)

3.3.4. Expectations for Metric Certainty (Reliability), Uncertainty, and Risk Assessment

Because of the counterfactual basis for efficiency evaluations, the indicated impacts from an efficiency evaluation will always be estimates. As discussed in one section of the Uniform Methods Project:[27]

> Uncertainty can be introduced at every stage of the evaluation, including the sampling, measurement, and adjustment. It is often difficult or impossible to quantify the effect of every potential source of error. Evaluation reports often limit uncertainty discussions to random error (especially sampling error and regression error), because there are well-understood methods for quantifying uncertainty due to random errors. However, a high-quality evaluation should include strategies for mitigating all major sources of uncertainty, and a high-quality report should discuss unquantifiable aspects of uncertainty so research consumers can fully assess the research rigor.

[27] Khawaja et al. 2013, p. 11-3.

UNCERTAINTY

Uncertainty is a measure of how "good" an estimate is, and refers to the amount or range of doubt surrounding a measured or calculated value. Without some measurement of uncertainty included in evaluation reports, it is difficult to judge an estimate's value as a basis for decision making and taking action based on the evaluation results. Any report of a program's energy savings, for instance, has a range of uncertainty surrounding the estimated values relative to the true values (which are not known). Causes of uncertainty are

- **Systematic error** (that is, not occurring by chance), such as non-coverage, non-response, self-selection, and some types of measurement errors.
- **Random error** (that is, occurring by chance), attributable to using a population sample rather than a census to develop the calculated or measured value. This error type also can be the result of some types of measurement error.

All EM&V frameworks documents should recognize and address the uncertainty associated with estimates by defining expectations for (a) how sources of uncertainty are to be addressed and reported, and (b) the certainty of reported results. This is important because additional investment (i.e., more budget, more time) in the estimation process can lead to increases in certainty, trade-offs between managing evaluation costs (and timing) and reductions in uncertainty are inevitably required. Providing some guidance on this subject in the framework document can help the evaluators sort through these tradeoffs to come to reasonable, and acceptable, levels of savings certainty as they design individual EM&V efforts (e.g., a program's impact evaluation).

In the framework document, very specific guidance can be provided. Framework documents such as the Puget Sound Energy example described in the text box, however, more typically provide general guidance and leave the specifics to be addressed in evaluation protocols. Examples of specific guidance are requirements for certainty and precision in any sampling decision—such as the often quoted 80/20 or 90/10 confidence/precision[28] levels, and whether these are to be applied at the measure, program, or portfolio level. More generally, and addressing both random and systematic errors, guidance can be included that results should or must be reported with indications of uncertainty. More specifically, because some error ranges are hard to or perhaps even impossible to quantify, a useful set of framework document requirements can be that:

- Certainty criteria be included in evaluation plans with justification for the level of certainty and quality assurance methods selected, and
- Evaluation results are to be presented with—
  - Quantitative or qualitative indications of the certainty associated with the results;
  - Discussion of the threats to their validity with consideration of potential biases;
  - Level of precision and confidence associated with any sample used and at what level (e.g., portfolio, program, overall, energy, demand); and
  - Quality-assurance systems and quality-control checks used to address uncertainty, including methods used to minimize bias.

EXAMPLE UNCERTAINTY GUIDANCE—PUGET SOUND ENERGY (PSE) FRAMEWORK

Uncertainty is defined for PSE's purposes as the range or interval of doubt surrounding a measured or calculated value within which the true value is expected to fall within some degree of confidence. EM&V resources will be deployed in a manner that provides the best value in terms of information that is required for oversight, market assessment, and program targeting, improvement, and planning. The level of investment put towards evaluation usually has a direct correlation to the amount of certainty achieved. One of the trade-offs in evaluations is thus between the costs expended and the uncertainty level. Results from an evaluation will be reported with the level of uncertainly or error rate defined and explained.

Evaluators are expected to control for systematic error through best practices and control random error by striving for a 90/10 confidence and precision level (using either a one-tailed or two-tailed test as appropriate) and requiring an 80/20 confidence level if sampling requirements can be shown to be unrealistic. Deviations from these specifications may be permitted with justification and review by the CRAG [Advisory Group].

Source: Puget Sound Energy (2015)

3.3.5. Consideration of Impact Persistence and Interactive Effects/Boundary Issues

This subsection briefly describes two other impact evaluation–related subjects, which should be addressed in a framework document. These topics relate to basic questions that an evaluator would ask as they prepare EM&V plans and reports for specific programs, projects, or measures, particularly associated with the scope and metrics that will be subject of the EM&V.

[28] The common factors for reporting random error-associated uncertainty are confidence and precision. Precision provides convenient shorthand for expressing the interval believed to contain the actual value. The confidence level is the probability that the interval actually contains the target quantity.

3.3.5.1 Savings Lifetimes and Savings Persistence

Energy and demand savings for a measure typically are estimated for one or more spans of time: (1) the first year, (2) a specified time horizon such as 10 years, or (3) the life of the measure. Although efficiency measure impacts (e.g., energy savings) often are calculated and reported as annual values, lifetime savings are essential for assessing the life-cycle benefits and cost-effectiveness of efficiency activities and for forecasting energy loads in resource planning. Lifetimes and persistence of energy savings are overlapping topics that often are not fully addressed in efficiency evaluations. In particular, the efficiency industry has not focused much effort on quantifying the lifetimes of savings and even less so on estimating savings persistence (or degradation) over the savings lifetime.

EFFECTIVE USEFUL LIFE (EUL)

A commonly used approach in the industry is to characterize measure lifetime as the EUL of a measure, defined as the median length of time (in years) that an energy efficiency measure is functional. An EUL is the number of years at which half of the measures remain in operation and half have expired. Conceptually, the EUL of an efficiency measure is a function of:

- Technical Equipment Life: The average number of years that a measure can operate
- Measure Persistence: The time that an energy-consuming measure lasts, considering business turnover, early retirement of installed equipment, and other reasons that measures might be removed, damaged, or discontinued.

Whether savings are to be reported as annual values, lifetime values, or both is a useful topic for defining in a framework. Furthermore, guidance on how effective useful life (EUL) (*see* text box) are defined and determined, as well as assumptions about what occurs at the end of a measure's EUL[29] are important topics that also should be covered in a framework document. One caution though is that if one is adopting guidance from other jurisdictions, the definitions and ways EULs are determined vary significantly across the country.[30] Additionally, the specific definitions and quanitifcation of lifetime impact metrics needed by energy resource planners, regulators, implementers, and others can vary; indicating both a challenge and an opportunity in the framework to define appropriate lifetime savings values that are of value to all users of the evaluation results.

[29] A range of assumptions can be made regarding what happens to energy savings (and emissions avoidance and other benefits) at the end of measure lifetime. One approach is that the energy use of the affected end use is assumed to revert back to the baseline efficiency at the end of the measure's life, so residual savings are zero. At the opposite end of the spectrum, it is assumed that efficient equipment and systems will be replaced with equipment or practices either equivalent to or more efficient than the original efficiency measure, so savings continue indefinitely (with or without incremental costs).

[30] Hoffman et al. 2015.

3.3.5.2 Boundary Issues and Interactive Effects

When evaluating impacts (e.g., energy, demand, emissions, economic development) it is important to properly define the project boundaries (i.e., the equipment, systems, facilities, or even markets and geographic regions that will be included in the analyses). Ideally, all primary effects (e.g., the intended savings) and secondary effects (unintended positive or negative effects—sometime called *interactive factors*) and all direct (at the project site) and indirect (at other sites) effects will be taken into account. From a practical point of view, with respect to energy and demand savings, this translates into deciding whether savings will be evaluated for specific pieces of equipment (where the "boundary" can include, for example, just motor savings or light bulb savings), the end-use system (such as the HVAC or the lighting system), whole facilities, or an entire energy supply and distribution "system." These aspects of EM&V should be raised in the framework and either addressed in the framework or alternatively left to being addressed on a program-by-program or even measure-by-measure basis in program EM&V or project M&V plans.[31]

INTERACTIVE EFFECTS

Interactive effects are increases or decreases in the use of electricity or other fuels that occur outside of the end uses targeted by a specific energy efficiency measure, project, or program. For example, reduction in lighting loads through an energy-efficient lighting retrofit can reduce buildings' air conditioning requirements and increase heating requirements because less heat is generated by energy-efficient lighting systems as compared with less-efficient lighting systems. Measures also can interact. For example, savings from the installation of weatherization measures will affect the savings associated with the installation of a higher-efficiency heat pump or furnace.

A larger "boundary issue," usually associated with efficiency measures that provide electricity saving is whether impacts (e.g., energy and demand savings) are to be reported on a site-basis or on the basis of savings at a grid-connected power plant (source-basis). The energy and demand savings at the power plant producing the electricity will be greater than the savings at the end-use (in the facility) due to transmission and distribution (T&D) losses. According to EIA data, national, annual T&D electricity losses average about 5% of the electricity that is transmitted in the United States.[32] Moreover, because savings from many energy efficiency measures occur during periods of high demand, the marginal T&D losses can be one to two times the average annual losses.[33] Thus, the savings at the power plant busbar can be meaningfully higher than the end-use savings. Whether T&D losses will be included in the boundary and reported as part of program impacts is something that should be defined in all EM&V framework documents for purposes of reporting clarity and transparency. Boundaries also are important for defining non-energy impacts (NEIs) as well. For example, if there are emission or job impacts to be assessed, it needs to be decided whether the impacts are local, statewide, national, or even international.

BOUNDARY EXAMPLE, INCLUDING T&D IMPACTS—DELAWARE FRAMEWORK

All transmission and distribution loss factors applied to customer or meter-level savings in order to estimate generation-level savings shall be based on estimates of marginal system line losses rather than average loss factors.

Source: Delaware (2017)

[31] Because the determination of interactive effects between measures (e.g., weatherization and high-efficiency furnaces) often involves the use of engineering calculations or sophisticated building or system simulation models, it most often is dealt with in protocols or the development of deemed savings values.

[32] http://www.eia.gov/tools/faqs/faq.cfm?id=105&t=3.

[33] Lazar and Baldwin (2011).

3.4. Approaches for Specific Program Categories and Process, Market, and Cost-Effectiveness Evaluations—Topics Recommended for Consideration

Prior sections herein focused on guidance topics mostly associated with impact evaluations. A framework document also can address other types of efficiency-evaluation activities as well as provide some guidance on how particular program categories should be addressed with regard to the evaluation activities. The topics described in this section are optional, but it is strongly recommended they be considered for inclusion in a framework.

3.4.1. Considerations for EM&V Focused on Specific Program Categories

Framework documents can include guidance for evaluating specific categories of efficiency programs that could have relatively unique attributes or objectives. Such guidance—even at a very general level—can assist in the development of portfolio and program-evaluation plans and budgets. Examples of program categories that might have unique evaluation criteria and examples of those criteria are listed below.

EXAMPLE EM&V FOCUS GUIDANCE—NEW YORK FRAMEWORK

In particular, program administrators should place a high priority on EM&V activities that target those programs, measures, or technologies that:

- Defer expensive infrastructure investments
- Are eligible for utility [Earnings Adjustment Mechanisms] EAMs
- Perform far above or below expectations
- Are implemented on a "test and learn" or pilot basis
- Have high energy savings variability
- Are based on a limited existing knowledge base
- Represent large contributions to the program administrator's overall portfolio savings.

Source: New York (2016)

- **Pilot programs**. Additional process and impact analyses might be desired given the need to determine whether these programs should be modified, expanded, or cancelled.
- **Low-income programs**. These programs can be implemented to provide a wide range of energy and non-energy benefits, thus additional evaluation activity could be focused on the non-energy benefits and perhaps less emphasis (i.e., rigor) on the energy benefits.
- **Hard-to-reach market programs**. These programs are targeted at market segments that have historically had low penetration of efficiency measures (e.g., small businesses) thus a focus of the evaluation could be on process and market analyses to assess which techniques are—or are no—increasing market penetration.
- **Emerging technology programs**. These programs are targeted at technologies that might not be widely available or that might still be evolving (e.g. advanced lighting controls, heat pump water heaters) therefore evaluations might focus on consumer acceptance, performance reliability, and cost trends, in addition to estimated savings.
- **Behavior-based programs.** Given the nature of these programs and the relatively limited (although increasing) experience to date with their implementation, more rigorous impact evaluations (such as randomized-control trials) could be required.

With respect to the last category, as an example, in the Arkansas guidance document,[34] the criteria for acceptable accuracy for evaluations of behavior programs is set at a more rigorous level than evaluations of conventional programs. This is due to the fact that average savings per participant typically associated with behavior-based programs are only a few percent of a customer's or building's total consumption, even though due to the large

[34] Arkansas Public Service Commission 2016. Page 75.

number of participating customers, total savings from these programs can be a significant part of a program administrator's portfolio.

3.4.2. Cost-Effectiveness, Process, and Market Evaluations

Within the framework document the full complement of evaluation activities that are to be conducted should be indicated. Beyond just a listing, information can be provided about the expected or required scope, budget, and timing of these activities. Section 3.2 discusses these issues for impact evaluations, which tend to be what most frameworks focus on. Sections 1 and 2 mention the other types of evaluations and this subsection provides some examples of topics associated with each that could be addressed in a framework:

3.4.2.1 Cost-Effectiveness Evaluations

The most common element addressed in an EM&V framework document with respect to cost-effectiveness (CE) is defining which CE tests will be used and cost and benefits elements that will be included in the tests (e.g. which participant costs and which, if any, non-energy benefits). Historically, efficiency has been assessed via tests as, or similar to those, defined in the California Standard Practice Manual (CaSPM),[35] or in some cases through integrated resource planning processes. Knowing which tests will be used, or whether IRP will be used, is the first step in defining what data will be required to determine cost-effectiveness.

A closely related topic that should be addressed in an EM&V framework is whether the selected cost-effectiveness test will be applied at the measure level, the program level, the sector level, or at the overall portfolio level. Although measure-level cost-effectiveness often is used for screening which measures are included in program offering, program-level cost-effectiveness typically is reported because it includes administration and marketing costs. Also, in many jurisdictions, the cost effectiveness of the entire portfolio of efficiency programs serves as the metric of most interest to regulators.

Knowing what data are required to assess the costs and benefit elements and at what level of aggregation the determination of cost-effectiveness will be made feeds into defining the EM&V activities and what data collection and analyses are required. The CaSPM describes data requirements for calculating cost-effectiveness under each of its tests. It is not very comprehensive with respect to data requirements, however, thus a second step for defining data requirements is delving into the specifics associated with each element of a CE test. Recently a new CE guide has been developed, the National Standard Practice Manual (NSPM),[36] which defines a resource value framework (RVF) that can be used to construct a jurisdiction's cost-effectiveness test(s) in a structured and documented manner that meets the specific interests and needs (as defined by the relevant policies) of the jurisdiction. Thus, using the RVF provides a way to not only define the jurisdiction-appropriate CE test(s), but also the test(s) data requirements, and thus can provide more information for the EM&V framework.

[35] California Public Utilities Commission (2001).
[36] National Efficiency Screening Project (2017).

3.4.2.2 Process Evaluations

The most common elements of a framework document with respect to process evaluations is simply to indicate whether they will be conducted and, if so, by whom and with what intent. For who conducts process evaluations, the two common entities are the administrator or implementers themselves (or consultants they hire), or less-commonly a third-party evaluator retained by an oversight body (e.g., a utility commission).

Other process evaluations topics that could be covered in a framework include the following.

- Which programs will have process evaluations or what are criteria for conducting process evaluations on particular programs (*see* text box)?
- What types of entities will be included in process evaluations, either to assess the process or to gather the input (e.g., administrators, contractors, vendors, participants, non-participants)?
- When and how often will the process evaluations take place?
- What are expected outcomes of process evaluations (e.g., recommendations for changing a program's structure, implementation approaches, and goals)?

POSSIBLE CRITERIA FOR WHEN TO CONDUCT A PROCESS EVALUATION

- The program is targeted for expansion and formative information is needed to support program design modifications for attracting higher participation levels
- Benefits are higher/lower than expected and/or are being achieved more quickly/slowly than expected
- There is limited program participation
- The program is a greater success than anticipated
- The program has a slow start-up
- Participants are reporting problems
- The program appears not to be cost-effective
- The program is built around a new concept that could be replicable for other populations, technologies, etc.

EXAMPLE PROCESS EVALUATION GUIDANCE—PUGET SOUND ENERGY FRAMEWORK

Process evaluations of the Company's Energy Efficiency Department programs will involve systematic assessments of programs or internal operations for the purposes of documenting program operations at the time of the examination, and identifying and recommending improvements to increase the program's efficiency or effectiveness for acquiring energy resources while maintaining high levels of participant satisfaction. The primary mechanisms used for process evaluations are data collection via surveys, questionnaires, and interviews to gather information and feedback from administrators, designers, participants (for example, facility operators or residential customers), implementation staff (including contractors, subcontractors, and field staff), and key policy makers. Other elements of a process evaluation can include creation or updating program theory and logic models, process mapping, workflow and productivity measurements, reviews, assessments, and testing of records, databases, program- related materials, and tools.

Source: Puget Sound Energy (2015)

3.4.2.3 Market Evaluations/Studies

As with the process evaluation, the most common elements of a framework document with respect to market studies simply indicate whether they will be conducted and, if so, by whom—typically consultants hired by the program administrator or an oversight body. Other information that can be useful in the framework with regard to market studies is defining budgets and time frames[37] as well as the expected outcomes and uses of the market studies evaluations—such as determining both total unit sales and total "efficient" unit sales for the products targeted by efficiency programs, for:

[37] Market studies can be expensive endeavors and take a significant amount of time to complete.

- Assessing cumulative impact of all forces acting on a market, to feed back into program planning, load forecasting, resource planning, and efficiency potential assessments;
- Defining common practice baselines and net to gross ratios; and
- Understanding consumer and supplier perspectives for improving program delivery efficacy.

3.5. Logistics—Topics Recommended for Consideration

There are several logistics-related topics that an EM&V framework could address. These are topics that often can be worked out as the portfolios are implemented and evaluations are designed and conducted. They are also topics that are often easier dealt with up front instead of waiting for issues to arise, however, such as in the cases of data confidentiality and disputes. Thus, the topics described in this section are optional, but are strongly recommended for consideration for inclusion in a framework.

3.5.1. Data Management Approaches, Including Confidentiality Considerations

Evaluations are based on data—often in very large amounts. Although data management easily could be defined as "down in the weeds," it can be a major attribute of a portfolio's implementation and evaluation, and a major cost. For stakeholders, the data questions addressed in the framework document tend to be associated with: "Will there be well-documented program-tracking systems with quality-assurance protocols?"; "Will consumer confidentiality be properly maintained by evaluators?"; and "What will be publicly accessible?"

EXAMPLE OF DATA MANAGEMENT GUIDANCE—TEXAS FRAMEWORK

For the purpose of analysis, the utility shall grant the EM&V contractor access to data maintained in the utilities' data tracking systems, including, but not limited to, the following proprietary customer information: customer identifying information, individual customer contracts, and load and usage data in accordance with § 25.272(g)(1)(A) of this title (relating to Code of Conduct for Electric Utilities and Their Affiliates). Such information shall be treated as confidential information.

(A) The utility shall maintain records for three (3) years that include the date, time, and nature of proprietary customer information released to the EM&V contractor.

(B) The EM&V contractor shall aggregate data in such a way as to protect customer, retail electric provider, and energy efficiency service provider proprietary information in any non-confidential reports or filings the EM&V contractor prepares.

(C) The EM&V contractor shall not utilize data provided or received under commission authority for any purposes outside the authorized scope of work the EM&V contractor performs for the commission.

(D) The EM&V contractor providing services under this section shall not release any information it receives related to the work performed unless directed to do so by the commission.

Source: Public Utilities Commission of Texas (2017)

Decision makers often do not need (or want) to see the detailed calculations or raw data inputs that drive evaluation results. But all parties usually want to know that reported results are built on a foundation of sound data inputs with good documentation and record keeping. Proper data-management strategies allow administrators, implementers, and evaluators—as well as oversight bodies—to delve into the underlying data, both to be able to review underlying assumptions and to combine the data in new ways as they see fit for current program reviews or future program developments.

The fundamentals of good data management are the same across industries. Within an efficiency portfolio, two areas of primary importance are the ability to compare results across time (longitudinal analysis) and ability to compare results by factors such as program type or delivery mechanism. Thus, within a framework document, although it very well might not be necessary to define major elements of data management (including

confidentiality), it can be worthwhile to indicate its importance and that data management procedures will be defined in a jurisdiction's evaluation protocols and each portfolio and program evaluation plan. Some examples of data management elements are:

- Format of data to be provided by tracking systems, compatibility and standardization;
- Access to data and summaries;
- Data confidentiality protection protocols; and
- Data quality assurance and control.

3.5.2. EM&V Plan and Report Formats, Including Websites and Public Access

Section 2.3 presents an annotated list of possible EM&V plans, reports, and other documents, and Section 3.2.2 covers specifying in the framework which evaluation plans, reports, and other documents will be completed as part of the evaluation efforts. Additionally, to specifying documents, the framework also can include outlines for the documents. Example outlines can be found in the *Energy Efficiency Program Impact Evaluation Guide* (Schiller 2012), the Northeast Energy Efficiency Partnership EM&V Forum Model EM&V Methods Standardized Reporting Forms,[38] and the California Public Utilities Commission Energy Division's Impact Evaluation Standard Reporting Guidelines.[39] Access to the reports and summary data/results can also be defined in the Framework, such as which evaluation information will be made public and presented on an efficiency portfolio website.

EXAMPLE EM&V PLAN CONTENT GUIDANCE—NEW YORK FRAMEWORK

- **Components of an EM&V plan.** An EM&V plan documents and demonstrates a commitment to transparent and credible EM&V activities and results. EM&V plans should include the components outlined in Appendix D. The details of EM&V plans will necessarily vary depending on the size, scope, and type of subject matter being evaluated. All EM&V plans, however, are expected to clearly explain how the resulting EM&V activities will be consistent with the core goals of providing reliable, timely, cost conscious and transparent results.
- **EM&V Plan Filing Requirements.** This Guidance stresses the importance of transparency of the EM&V activities and results. In support of this, program administrators are required to file all EM&V plans publicly through the Commission's Document Matter Management (DMM) System in Matter 16-02180, In the Matter of Clean Energy EM&V.

Source: New York (2016)

Examples of such websites (with links to evaluation documents) are:

- California—http://www.cpuc.ca.gov/energyefficiency/ and http://www.calmac.org
- Hawaii—https://hawaiienergy.com/about/information-reports
- Northeast Energy Efficiency Partnerships—http://neep.org/resources
- Northwest Energy Efficiency Alliance—http://neea.org/resource-center
- Massachusetts—http://ma-eeac.org
- Texas—http://www.texasefficiency.com/index.php/emv

[38] http://www.neep.org/initiatives/emv-forum/model-emv-methods-standardized-reporting-forms.
[39] http://www.energydataweb.com/cpucFiles/pdaDocs/1399/IESR_Guidelines_Memo_FINAL_11_30_2015.pdf.

3.5.3. Dispute Resolution

Disputes can arise from the process used to develop evaluation results (usually the impact evaluations) and the results themselves. Disputes are best addressed before they arise through parties understanding how the evaluations will be conducted (i.e., defined in the evaluation framework document and specific plans) and through good documentation and communication. Disputes do sometimes arise, however, and it can be best to define how they will be addressed before they occur. Most jurisdictions have their own approaches, with mechanisms for regular discussions, regulatory hearings, mediations, arbitration, or other solutions.

Even a few lines in an evaluation framework document defining the steps for dispute resolution can eliminate a great deal of difficulty should a dispute arise.

DISPUTE RESOLUTION EXAMPLE—MASSACHUSETTS FRAMEWORK

Although [Program Administrators] PAs and the EM&V Consultant will continue to work diligently to reach a consensus on evaluation issues, where there are areas of difference that may arise that cannot be resolved through consensus during the on-going interactive process between the EM&V Consultant and the PA evaluation staff, authority for decision-making will reside with the EM&V Consultant and the Council.

To enable the Program Administrators to fulfill their responsibility to report program savings to the Department with full confidence, an appeals process has been established, through which the PAs may bring decisions made by the EM&V Consultant or the Council for review and resolution. This process will be implemented through the formation of an evaluation appeals committee ("Appeals Committee") of the Council, whose responsibility in this area will be to hear the matter under dispute and rule so that the study may proceed in a timely way. In general, it is expected that this review process will be completed within 72 hours once an issue is elevated to the Appeals Committee. . . . The PAs shall be able to submit any such documents to the Department in conjunction with the filing of the Three-Year Plans, mid-term modifications, and term reports. The Department will be able to review the record of this decision in its review of Three-Year Plans, midterm modifications, plan-year reports, and term reports.

To date, the EM&V Consultant and PA Evaluation staff have been able to resolve all areas of differences without proceeding to the Appeals Committee. . . . This is a testament to the hard work and collaborative engagement of the PAs and the EM&V Consultant.

Source: Massachusetts Department of Public Utilities (2015)

4. Developing and Updating EM&V Frameworks

This section describes processes and tasks associated with developing or updating an EM&V framework. Figure 4.1 lays out task categories with high level task descriptions. Figure 4.1 includes task categories that are process related and task categories associated with framework document content. Table 4.1 provides a list of process-related EM&V framework development/updating activities and content options. The figure and table both are found at the end of this section.

4.1. Framework Development and Updating Process

The process-related tasks in Figure 4.1 are basic ones that usually are associated with any important document which is of interest to multiple stakeholders. These include defining the document's scope, determining who will participate in its development, deciding who will write the actual document (assumed to be one or more consultants—see below), building a schedule with milestones, and having an approval process for a final document (for example by a regulatory body).

When developing a jurisdiction's first framework, the prior sections of this guide and, in summary, Table 4.1, can guide what subjects to cover. Frameworks, once completed, tend to be static and remain unchanged for at least several years as fundamental topics such as EM&V objectives, roles and responsibilities, and data transparency expectations would not be expected to change often, unless there are major policy changes. Other aspects can be updated though, thus an updating process can be initiated when the oversight entity (e.g., regulator), or perhaps major stakeholders, see a need for updating or further clarification.

With respect to actually creating or updating a framework, the content and related issues can be addressed through a variety of mechanisms, such as having a framework document that is prepared

- By a consultant or an evaluator as an initial task of evaluation activities for a portfolio of programs (as was the case for the Connecticut Green Bank evaluation framework),
- Via a collaborative effort in which the group actively provides feedback as a consultant or other entity prepares the framework (as is the case in Washington State where each of the major investor-owned utilities develops a framework with an advisory group), or
- As part of a regulatory proceeding that results in the framework being adopted in a formal ruling (as is the case in Texas where the framework is part of the Commission's Energy Efficiency Rule).

Using one entity to prepare the framework with regular input and feedback from stakeholders participating in a collaborative effort (as discussed in Section 4.2) is strongly suggested. This is to (a) gather stakeholder input, because the end result of addressing the Section 3 (and perhaps other) topics is the evaluation framework that documents the evaluation infrastructure and expectations *for all stakeholders,* and (b) to avoid the pitfalls of writing documents by committee. For most, if not all, products (such as evaluation reports) if the who, what, and how are defined up front and at least tacitly agreed to by all parties, then there is reduced likelihood of disputes over the results, even if not everyone likes what the results indicate. In particular for EM&V, the framework document can define the funding and time requirements for reliable evaluations in a way that they are understood and balanced with the selected methods, information needs, and accuracy expectations. If this balancing is agreed to by at least the primary stakeholders prior to conducting major evaluation activities, then evaluation efforts can be well-supported and succeed in providing the results desired with minimal dispute, at least over process-related issues.

In addition to the recommendation that a collaborative-based process be used to develop framework documents, it is also recommended that there be engagement by regulatory staff (when regulated utility customer funds are involved). This engagement includes (a) from being one member of a collaborative to directly facilitating the collaborative and guiding the framework consultant through a contractual management process, and (b) if practical, the regulatory entity (e.g., public service commission) approving the framework document.

There are three reasons that engagement of regulatory staff is recommended.

- Regulators ultimately will be asked to rely on the guidance and information in the framework document by those submitting evaluation reports;
- Regulatory staff participation fosters timely, coordinated, and informed regulatory review of reported results; and
- Regulatory approval provides a sense of certainty that the framework effort was of value and will be utilized.

With respect to regulatory staff participation, particularly by those with evaluation expertise and experience in managing collaborative efforts, having such participation and even leadership also can help provide public oversight, allow all parties to be heard, provide consistency in focus, and ensure that data are made available to the extent possible. This oversight role also can be provided by neutral facilitators, but the additional participation of regulatory staff also helps with acceptance of results by the regulatory commissions.

4.2. Using a Collaborative Approach to Developing and Updating Framework Documents

As mentioned, it usually is best to develop frameworks with a collaborative (and open) process that includes program administrators, implementers, evaluators, and independent technical experts, as well as advocates. Collaboratives usually are organized by either a regulatory commission or the program administrator(s). Beyond the value of gaining stakeholder support, some other advantages of using a collaborative group to help support framework document development include the following.

ARKANSAS COLLABORATIVE EXAMPLE

Arkansas' Technical Reference Manual (TRM) contains many of the elements that can be described as an EM&V Framework. Arkansas' TRM was developed through a collaborative process which incorporates feedback from the seven investor-owned gas and electric utilities, stakeholders, program implementers and evaluators. This group is known as the Parties Working Collaboratively or PWC. The PWC develops and recommends the TRM updates which are then approved by the Arkansas Public Service Commission.

Source: Arkansas (2017)

- Supports a peer-review process to review and provide informed input into the development of the framework document.
- Establishes facilitation of ongoing stakeholder collaboration and coordinated program evaluation and planning over time and across all program administrators in a jurisdiction.
- Reduces uncertainty for program administrators regarding their evaluated savings (and their cost-effectiveness) and whether such savings will be challenged by regulators or intervenors in regulatory proceedings.
- Leverages existing knowledge across multiple stakeholders to work collectively with shared resources.

Of course, there also are barriers to the development of framework documents using collaboratives and to the development of frameworks in general. These general barriers tend to be the same as those found when evaluating efficiency programs—time and funding requirements and, in some jurisdictions, absence of a driving policy (such as an EERS that requires measurement for compliance with the standard). Another, often significant, barrier to frameworks development is a lack of consensus among stakeholders—particularly when multiple utility service territories are involved—on what to include in frameworks and the positions of different parties on topics such as requirements for certainty of results, budget priorities, importance of independent evaluators, and technical issues such as the proper definitions of baselines for different program activities. In these circumstances, there can be a tendency to adopt EM&V guidance that is just a "lowest common denominator" of what can be

agreed to by all parties. This outcome can and should be avoided, even if this results in not achieving complete consensus among all stakeholders.[40]

Strategies for surmounting these barriers generally start with an oversight body (e.g., the utility regulator) adopting clear policy guidance that prioritizes the efficiency program evaluation, encourages collaboration, and establishes multiyear funding agreements that provide the resources, structure, and stability to conduct and maintain ongoing EM&V activities. Such strategies usually require the support of key stakeholders who both inform the scope and ensure transparency in the development of the framework.

Although all collaboratives seek input from stakeholders, membership of collaboratives varies. Some useful membership guidelines for EM&V framework document collaboratives include the following.

- Members have at least some background in efficiency programs and their evaluation as well as the time and commitment to stay engaged throughout the entire framework development or updating process[41] and consider framework issues, drafts, and other factors.
- Collaborative members having defined roles and responsibilities, and collaborative membership that includes program administrators, implementers, evaluators, and independent technical experts as well as advocates and, as mentioned above, active regulatory staff participation for those frameworks involving regulated utility customer-funded programs.

For more information on efficiency collaboratives see the publication, *Energy Efficiency Collaboratives: Driving Ratepayer-Funded Efficiency through Regulatory Policies Working Group*.[42] Topics covered in that publication include: attributes of successful energy-efficiency collaboratives, collaborative scopes, method of decision making, role of a regulatory commission and relationship with a collaborative, and overarching collaborative principles. Other information associated specifically with collaboratives that helped define frameworks can be found in some of the examples included in Section 5 herein.

4.3. Developing Framework Document Content

The issues and topics presented in Section 3 can be considered for inclusion in the framework document in a linear sequence. In reality, however, many are interrelated and the overall planning process for addressing these topics/issues tends to be iterative. Often the first decision that needs to be made in developing an EM&V framework, and an area where there often is a significant amount of iteration (e.g., discussion among stakeholders), regarding defining a set of evaluation objectives and metrics, and addressing other fundamental topics such as principles, budgets, and timing. Multiple iterations of the EM&V framework document can then occur as stakeholders consider trade-offs between the objectives (and the associated metrics) and rigor versus cost versus schedule. From these decisions then flows what form the products (e.g., impact evaluations, process evaluations, cost-benefit analyses) will take, roles and responsibilities (particularly whether there will be a third-party evaluator, operating independently from the program administrators, and who will select and manage such an evaluator), and allowable or expected EM&V methods, assumptions (e.g., baselines, impact certainty), and logistics.

4.4. Using Consultants to Draft the Framework Document

The task of actually preparing the framework document is usually outsourced to a consulting firm. Although, in some cases, it can be prepared by regulatory commission or program administrator staff. Consultants are used because they can have specific expertise in the EM&V topics and could be seen as independent sources that help guide stakeholders in the balancing of conflicting interests and constraints (e.g., budget constraints), without

[40] For customer-funded efficiency programs operated by investor-owned utilities, the state regulatory commissions are the ultimate arbiter of what is acceptable, so stakeholders must recognize that the commissions will make the final decision of what is acceptable EM&V.
[41] Changes in personnel assigned to a collaborative working group can be a challenge if positions change with different people assigned to the working groups, even those from the same entity.
[42] State and Local Energy Efficiency Action Network (2015).

undue bias with respect to the outcome of evaluations. Additionally, as of the writing of this guide, Lawrence Berkeley National Laboratory has a technical assistance programs, funded by the U.S. DOE, that can provide technical support on EM&V framework topics for public agencies.[43]

Consultants that prepare EM&V frameworks tend to be the same firms that conduct program evaluations. There are probably on the order of 30 to 50 such firms in the United States, ranging in size from having just a single professional to large, multinational firms with perhaps a hundred individuals involved to one degree or another in evaluating efficiency portfolios. The areas of professional expertise in these firms are usually engineering, economics, statistics, and the social sciences—and teaming commonly is used to bring a range of expertise and experience to an evaluation effort. Often selected through a competitive process, resources for finding consultants are listed below.

- Posting of request for qualifications or proposals at one or both of these professional societies:
  - International Energy Program Evaluation Conference (IEPEC)—https://www.iepec.org/?page_id=2
  - Association of Energy Services Professionals (AESP)—http://www.aesp.org/?page=rfps
- The list maintained by the California Measurement Advisory Council of efficiency evaluators (CALMAC)—http://www.calmac.org/contractorcontact.asp

[43] Berkeley Lab, on behalf of the Department of Energy, upon request provides objective technical assistance to state regulatory commissions, state energy offices, tribes, and regional entities. See https://emp.lbl.gov/research/technical-assistance-states.

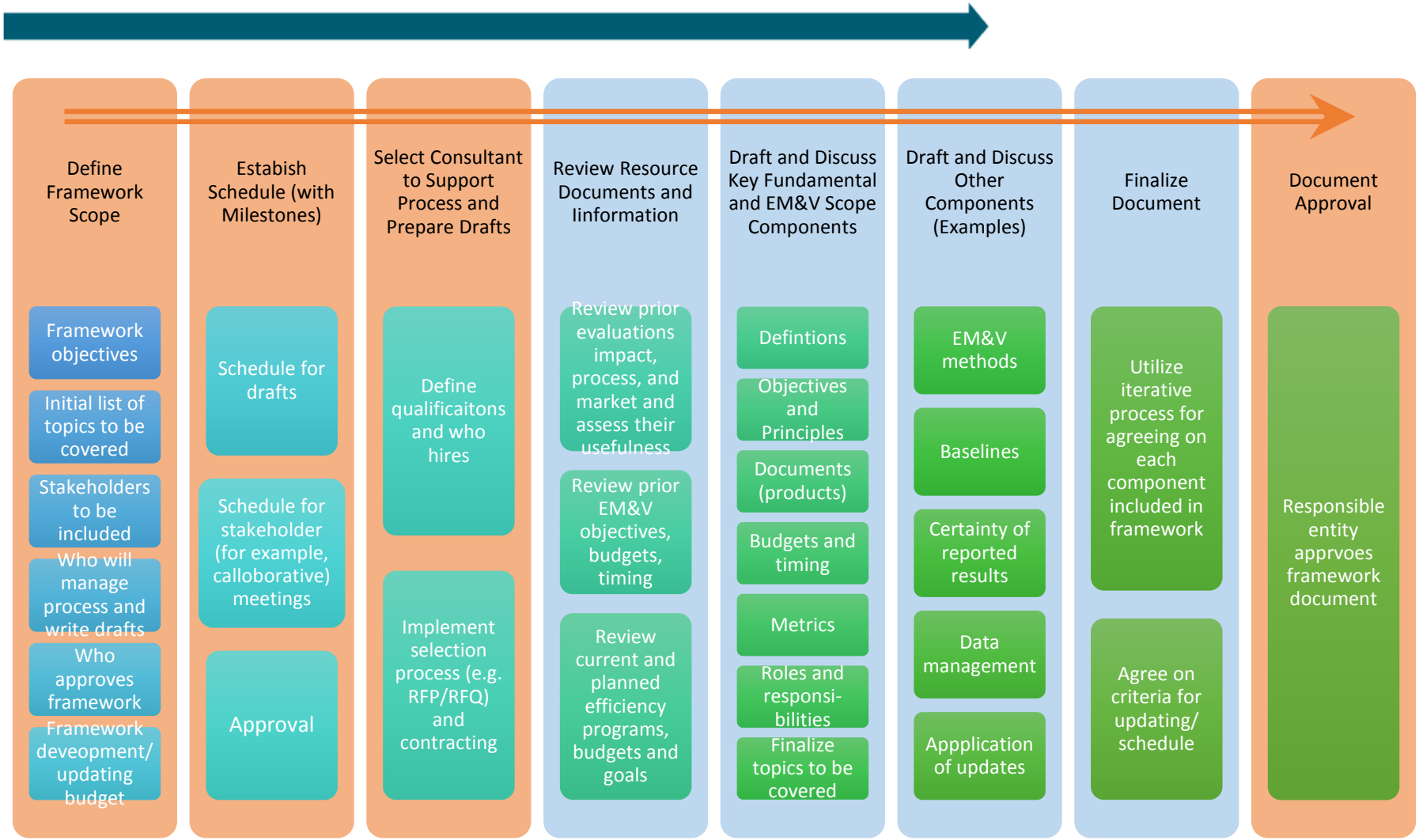


Figure 4.1. Example flowchart of framework development tasks

Note: Orange-shaded task categories are process related and blue-shaded task categories are associated with framework document content.

Table 4.1. Framework Development Process and Content Checklists

Framework Document Process Checklist

Define Framework Scope	Establish Milestones (Schedule)	Select Framework Consultant	Review Resources	Draft and Finalize Document	Document Approval
☐ Initial list of topics to be convered	☐ Drafts	☐ Define criteria and who hires	☐ Prior evaluaitons	☐ See content checklist below	☐ Regulator (or other entity) signoff
☐ Stakeholders to include	☐ Stakholder input/meetings	☐ Implement selection and contracting process	☐ Prior EM&V budgets, scopes, timing	☐ Iteriative process for agreeing on content	
☐ Who manages process	☐ Approval		☐ Current program programs	☐ Criteria/schedule for updates	
☐ Who approves Framework					
☐ Overall time frame and budget					

Framework Document Content Checklist

Fundamentals	**EM&V Scope**	**Impact Evaluation Approach**	**Approachs for Other Evaluaitons**	**Logistics**
☐ Definitions	☐ Metrics	☐ EM&V methods	☐ Considerations for specific market segments	☐ Data management
☐ Efficiency portfolio description	☐ Evaluation reports and other outputs	☐ Sampling design	☐ Process evaluations	☐ Report formats, websites, and public access
☐ Evaluation objectives	☐ Scale of the evaluation effort (e.g., budget)	☐ Baselines definition	☐ Market evaluations	☐ Dispute resolution
☐ Evaluation principles	☐ Timing of the evaluation	☐ Interactive effects and persistence	☐ Cost-effectiveness studies	
☐ Retroactive or prospective application of evaluation results	☐ Roles and responsibilities	☐ Expectations for metric certainty		

5. EM&V Framework Examples

This section provides brief descriptions and cites for several different entities' efficiency portfolio EM&V framework documents. These are provided as examples and not necessarily endorsements of the content of these frameworks or their indicated approaches to EM&V, as each jurisdiction determines what is best for its own requirements. Some of these are not specifically called out as frameworks by their authors, but they meet the criteria for framework documents as defined in this guide.

One of the selection criteria for presenting these frameworks as examples is that they are publicly available.[44] Another resource for information on how different states address EM&V is the American Council for an Energy-Efficient Economy's (ACEEE) EM&V page of their State and Local Policy Database.[45] Intended to be updated at least once per year, the ACEEE EM&V page indicates information about each state's EM&V methods and, for many of the states, links to related documents.

The topics covered in each of these example frameworks are illustrated in Table 5.1 (located at end of the section), using the topic categories described in Section 3. Table 5.1 is intended to indicate whether the topics listed are covered in some depth and are relatively easy to find in the referenced documents. In some cases, there might be topics not indicated as being covered, and for which there is some reference to the topic in the document, but perhaps only in passing or as related to other topics. Thus, Table 5.1 is intended as a broad indicator of coverage, not a comprehensive or definitive indicator.

5.1. Statewide Frameworks for EM&V of Investor-Owned Utility Efficiency Programs

5.1.1. Arkansas

Arkansas Public Service Commission (2017). Many of the elements of an EM&V framework are contained in Volume 1 of the Arkansas Technical Reference Manual. It was prepared by the Arkansas Commission's Independent Evaluation Monitor (a consultant team) on behalf of the "Parties Working Collaboratively," a stakeholder group. Volume 1 of the TRM document is divided into two sections: Section I—Overview of EM&V Terms, Methods, and Approaches; and Section II—EM&V Protocols. As explained in the document, the purpose of the EM&V protocols is to provide a common framework and set of reference points for conducting cost-effective program evaluations. These protocols describe the types of information that must be collected to conduct a comprehensive examination of a program's overall effectiveness, the recommended frequency for conducting these program evaluations, and the key metrics that must be reported during these evaluation activities. Also included is guidance on the role of EM&V as well as key definitions, recommendations regarding data capture, and EM&V reporting formats. In terms of topics addressed in the document, it was designed to address the specific topic areas called for coverage in the relevant commission energy efficiency order. The Arkansas document has a very useful index for finding topic coverage.

5.1.2. California

California Public Utilities Commission (2014). As defined in the California Measurement Advisory Council (CALMAC) online database[46]: "the [California] Evaluation Framework provides program evaluators, administrators, and others with a comprehensive set of guidelines for conducting evaluations of California's energy efficiency programs. The framework includes recommendations for conducting impact evaluations, including measurement and verification (M&V) efforts, as well as process, market effects, information/education/training program and non-energy benefits evaluations. It includes evaluation methodology descriptions and numerous references. Guidelines are also presented for evaluation sample design and statistical analysis and for assessing and reducing

[44] Detailed citations for each of these frameworks is included in the references section.
[45] http://database.aceee.org/state/evaluation-measurement-verification.
[46] http://www.calmac.org/search.asp.

the level of uncertainty of evaluation results. The framework includes a set of decision protocols for deciding what to evaluate and when to conduct the evaluations. It provides examples of how the evaluation structure can fit into a cycle of program portfolio planning, implementation, and evaluation." A companion document published in 2006 is the California Evaluators' Protocols,[47] which are intended as more detailed guidance for conduct of specific evaluations. The California framework is one of the best-known frameworks, however it is also one of the longer, if not the longest, frameworks (at nearly 500 pages) and is rather dated. A 2017 study was initiated by the California Public Utilities Commission and the investor-owned utilities to identify ways to improve the usefulness and usability of the California Evaluation Framework and ensure its applicability to meet evaluation needs of California given changing policy and industry environments.[48]

5.1.3. Delaware

Delaware Department of Natural Resources and Environmental Control (2017). Delaware's version of a framework is called the Evaluation, Measurement and Verification Procedures and Standards. These are contained in a department secretary's order. The order summarizes the intent of the procedures and standards in a manner that also summarizes the purposes of frameworks in general: "(1) develop and govern the overall approach to the evaluation of energy efficiency and demand response programs in Delaware; (2) standardize evaluation approaches for the assessment of energy efficiency and demand response programs; (3) provide specific guidance to Program Administrators, contractors and stakeholders for the evaluation of energy efficiency and demand response programs; and (4) ensure consistency between Program Administrators' energy efficiency evaluation plans, analysis, and reporting efforts." The Delaware document was prepared with input from the State's Energy Efficiency Advisory Committee (EEAC, a 13-person stakeholder group) and with a public input process. A somewhat unique aspect of the Delaware framework is fairly detailed descriptions of the responsibilities of the EEAC, Delaware Department of Natural Resources and Environmental Control, Program Administrators and Independent Evaluation Contractors.

5.1.4. Hawaii

Public Utilities Commission of Hawaii (2012). Hawaii has a framework for its Energy Efficiency Portfolio Standard (EEPS). Hawaii's EEPS is a law that requires continuous increases in energy savings over time. The framework was established as part of a Public Utilities Commission of Hawaii (HPUC) decision and covers a number of topics that define the goals and operation of EEPS programs, including EM&V. Although the EM&V discussed in the EEPS framework does not necessarily directly apply to the programs that use utility-customer funds (which are administered by a third-party), there is overlap in topics. The framework was developed by consultants to the HPUC, with input from parties to a formal HPUC docket. Within the EEPS framework there is a specific section on EM&V, but some topics discussed in this guide, and as indicted in Table 5.1, are addressed in other sections of the Hawaii EEPS framework. The Hawaii framework is one of the more succinct frameworks of the examples provided.

5.1.5. Maryland

Maryland (2017). This framework is called the "EmPOWER Energy Efficiency Programs Strategic Evaluation Guidance." "EmPOWER Maryland" is the name of the state's initiative to reduce energy consumption under six of Maryland's utilities and the Department of Housing and Community Development programs. The guidance is intended to compile the decisions made over time with respect to various programs' evaluation and cost-effectiveness analysis scopes, methods, assumptions and frequency, along with how evaluation and cost-effectiveness results are used for portfolio planning, informing program design, and tracking toward EmPOWER goals. The guide's objectives are defined as being to:

> 1) guide the planning, execution and reporting of evaluations; 2) record/inventory evaluation policy decisions made to date; 3) enhance transparency and clarity of methods and assumptions and processes for evaluating and reporting savings; 4) facilitate knowledge transfer and program continuity in the event of staff turnover for Program Administrators, Evaluators, [Public Service

[47] California Public Utilities Commission 2006.
[48] Malinick et al. 2017, p. 1.

Commission Technical Staff ("PSC Staff")], and other stakeholder organizations; and
5) acknowledge the many things that have been agreed on (NEEP 2016).

The document states that it is not intended to be comprehensive, but to supplement and complement orders from the Public Service Commission, various evaluation protocols, and the Mid-Atlantic TRM.[49] Also, as indicated in the document, "the Guidance represents a consensus of the EmPOWER Program Administrators, [PSC Staff], the Office of People's Counsel (OPC), the Maryland Energy Administration (MEA), the Statewide Evaluators, and the Independent Evaluator" (NEEP 2016).

5.1.6. Massachusetts

Massachusetts Department of Public Utilities (2015). Within the extensive (~2,600-page) Massachusetts Joint Statewide Three-Year Electric and Gas Energy Efficiency Plan (Plan) is a section on EM&V. Additionally, there is an extensive Strategic Evaluation Plan (SEP) included in the Plan as an appendix. Topics that, per this guide, can be considered part of a framework are covered throughout the whole SEP. Some unique contents of this framework are descriptions of strategic evaluation issues that can be addressed in future studies and brief work plans for evaluation topics that could be completed during an upcoming evaluation cycle. As noted within the document, the SEP was guided by a three-day Strategic Evaluation Planning Summit with program administrators, other stakeholders, and consultants. In Massachusetts, there is an Energy Efficiency Advisory Council led by the council chair, the Massachusetts Department of Energy Resources, and an Evaluation Management Committee (EMC)[50]—all involved in the development of quite comprehensive EM&V planning and framework documents.

5.1.7. New York

The example for New York is the "Evaluation, Measurement & Verification Guidance," which is intended to provide guidance to the state's efficiency program administrators (utilities and the New York State Energy Research and Development Authority) and evaluators for conducting EM&V activities associated with utility customer–funded clean energy programs. Prior to this guidance document, the New York Department of Public Service issued evaluation guidelines to support implementation and oversight of Energy Efficiency Portfolio Standard (EEPS) programs. These guidelines—the "New York Evaluation Plan Guidance for Energy Efficiency Portfolio Standard Program Administrators"—were developed with input from the former Evaluation Advisory Group. Then, in 2016, commission staff, in consultation with a Clean Energy Advisory Council, conducted a review of the "Evaluation Guidelines" and revised and reissued the new guidance document. The New York framework's coverage includes protocols for process impact and market evaluations.

5.1.8. Pennsylvania

Pennsylvania Public Utility Commission (2016). Many of the elements of an EM&V framework are contained in the document titled, "Evaluation Framework for Pennsylvania Act 129[51] Phase II Energy Efficiency and Conservation Programs. It was prepared by the Statewide Evaluator (SWE, a consultant team) for the commission, although this document is not part of a commission order and thus is not mandatory. The Evaluation Framework outlines the metrics, methodologies, and guidelines for measuring performance by detailing the processes that should be used to evaluate the Act 129 programs. This Pennsylvania framework also includes many topics that often are associated with TRMs and evaluation protocols. Consistent with the objectives of frameworks defined in this guide, the Pennsylvania document is a rulebook that establishes the Act 129 program evaluation process and communicates the expectations of the SWE to the utilities administering the programs and their evaluation contractors. As indicated in the document, utilities "that align their EM&V processes with the Evaluation Framework should expect less scrutiny from the SWE as part of the SWE audit activities." Also, as noted in the

[49] Northeast Energy Efficiency Partnerships (NEEP) 2016.
[50] The EMC serves as a steering committee for statewide evaluation issues, providing guidance and direction for each of the evaluation research areas.
[51] Seven Pennsylvania electric distribution companies implement efficiency and conservation programs to promote the goals and objectives of Pennsylvania's Act 129.

document, and consistent with framework objectives defined in this guide, it "sets the stage for discussions among" stakeholders and those engaged in the evaluation efforts.

5.1.9. Texas

Public Utilities Commission of Texas (2017). Within the long document known as the Texas Commission's Energy Efficiency Rule is a section (Section q) that provides guidance on the conduct of EM&V for the state's utility customer–funded efficiency programs. The rule and this EM&V framework were part of a rule-making process with parties to the rule having input on its content. This is another succinct example of an EM&V framework and although it covers a number of topics that the commission and stakeholders thought were most important and upon which agreement could be reached, the amount of detail is limited. Within other sections of the energy efficiency rule, however, there are other references to EM&V, such as in the definitions section.

5.2. Statewide Efficiency Financing-Based Program EM&V Framework—Connecticut Green Bank (2016)

The Green Bank is a quasi-public agency created by state legislation. It attracts and deploys private capital to accelerate the deployment of clean energy, including efficiency, in Connecticut. Although the Green Bank has some statutorily required auditing and reporting requirements, it is not obliged to evaluate its programs in the same manner as are regulated utilities. As indicated in its framework, however, the Green Bank is "committed to evaluating its programs in order to ensure that the Clean Energy Fund, cap-and-trade allowance proceeds, and other investments are yielding value to the Green Bank's objectives and that the Green Bank's programs effectively and efficiently operate and deliver their services to customers." The Green Bank hired a consultant team to complete certain EM&V projects including assisting the bank in developing an evaluation framework to assess, monitor, and report program impacts and processes. The consultant team, according to the framework document, received feedback on the framework from the Board of Directors of the Green Bank, the Joint Committee of the Energy Efficiency Board, and the Green Bank. As a framework for a specific financing program, this framework is unique in that it starts with describing a logic model for the financing program to set the stage for evaluations strategies that assess performance in the context of the logic model and metrics identified by that model.

5.3. Single Utility EM&V Framework

5.3.1. Indianapolis Power and Light Company (2015)

The Indianapolis Power and Light Company (IPL) is an investor-owned utility that, in collaboration with its Oversight Board, developed an evaluation framework for its efficiency portfolio based on a 2012 statewide framework that was prepared for what are now discontinued statewide core programs. Consisting of IPL, Office of Utility Consumer Counselor, and the Citizens Action Coalition, the Oversight Board works collaboratively to guide the implementation of IPL's DSM program. As with other frameworks, the purpose of this document is to provide (as indicated in the document itself) a "consistent platform from which evaluations can be designed and implemented so that evaluation results are both reliable and comparable across programs, evaluators, and program implementers." The IPL framework has sections on evaluation-related policies that cover objectives; budget planning; tracking; evaluation report contents; cybersecurity; and evaluation standards, expertise, and ethics. Other sections cover how evaluations are to be conducted with guidance on risk mitigation and reliability of evaluation results, M&V field protocols, survey research approaches, baseline approaches, savings persistence, net energy impact attribution approaches, and data needs.

5.3.2. Puget Sound Energy (2015)

Puget Sound Energy (PSE) is an investor-owned utility in Washington State. Working with its Conservation and Resource Advisory Group, it prepared its EM&V framework. The document is intended to meet the interests and intentions of the relevant Washington Utilities and Transportation Commission (WUTC) dockets. It thus describes PSE's approach to evaluating, measuring, and verifying the results of the efficiency measures, programs, and portfolio funded by its customers. As of the writing of this guide, the PSE framework is being revised with minor

changes as part of a regular updating process. The framework is included with PSE's filings with the WUTC for approval of its Biennial Conservation Plans and is similar to one filed by at least one other Washington utility, Avista.[52]

Table 5.1. Topics Substantially Covered in Example Framework Documents

Topics	Subtopics	AK	CA	DE	HI	NY	MA	MD	PA	TX	CT GB	IPL	PSE
Fundamental Topics/Issues	Definitions	X	X	X	X	X	X		X	X			X
	Portfolio Description				X		X		X		X		
	Evaluation Objectives	X	X		X	X	X	X	X	X	X	X	X
	Evaluation Principles	X	X	X	X		X			X		X	X
	Application of Updates			X				X		X			X
EM&V Scope Topics	Metrics	X	X	X	X		X	X	X	X	X	X	X
	Evaluation Documents	X		X	X	X		X	X	X		X	X
	Evaluation Budget	X	X	X	X		X					X	X
	Evaluation Timing	X	X		X	X	X	X	X				X
	Roles/ Responsibilities			X	X		X	X	X	X			X
Impact Evaluation Approach Topics	EM&V Methods	X	X		X	X		X	X	X		X	X
	Baselines	X	X	X	X			X		X		X	X
	Sampling Design	X	X	X								X	X
	Interactive Effects		X									X	X
	Persistence	X	X									X	X
	Certainty	X	X	X		X						X	X
	Program Specific Guidance	X	X				X						
Approaches for Other Evaluations Topics	Program Specific Guidance	X	X				X						
	Process Evaluations	X	X	X		X	X		X				X
	Market Evaluations		X	X	X		X						X
	Cost Effectiveness		X	X	X		X	X	X		X	X	X
Logistics Topics	Data Management	X	X			X			X	X		X	X
	Reporting/Public Access		X			X		X	X				
	Dispute Resolution						X						

[52] Avista Utilities 2016.

6. References

6.1. Framework Example References

Arkansas Public Service Commission. 2017. *Arkansas Technical Reference Manual*, Volume 6.1, Volume 1. Approved in Arkansas Public Service Commission Docket 10-100-R. http://www.apscservices.info/EEInfo/TRM6-1.pdf.

Avista Utilities. 2016 (April). Evaluation Measurement and Verification Framework. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=0ahUKEwiqte6rg6bXAhVQ0mMKHSMKAb4QFggxMAE&url=https%3A%2F%2Fwww.myavista.com%2F-%2Fmedia%2Fmyavista%2Fcontent-documents%2Fenergy-savings%2Fava_emv_framework.pdf%3Fla%3Den&usg=AOvVaw0Ng5LFfjpuBj9DC5_cUeCO.

California Public Utilities Commission. 2014 (June). California Evaluation Framework. http://www.calmac.org/publications/California_Evaluation_Framework_June_2004.pdf.

California Public Utilities Commission. 2006 (April). California Energy Efficiency Evaluation Protocols: Technical, Methodological and Reporting Requirements for Evaluation Professionals [a.k.a. Evaluators' Protocols]. www.cpuc.ca.gov/WorkArea/DownloadAsset.aspx?id=5212.

Connecticut Green Bank. 2016 (July). Evaluation Framework. http://www.ctgreenbank.com/wp-content/uploads/2017/02/CTGreenBank-Evaluation-Framework-July-2016.pdf.

Cooney, Kevin. 2015 (August 27). Why, When, and How to Utilize the "Right" Methods for Evaluation of Energy Efficiency Programs. Presentation at EEDAL 2015—Lucerne.

Delaware Department of Natural Resources and Environmental Control Division of Energy and Climate Evaluation, Measurement and Verification Procedures and Standards, Secretary's Order No.: 2016-EC-0048. Date of Issuance: December 15, 2016. Effective Date of the Amendment: January 11, 2017. http://www.dnrec.delaware.gov/Info/Documents/Secretarys-Order-No-2016-EC-0048.pdf.

Franconi, Ellen, Matt Gee, Miriam Goldberg, Jessica Granderson, Tim Guiterman, Michael Li, and Brian A. Smith. 2017. "The Status and Promise of Advanced M&V: An Overview of "M&V 2.0" Methods, Tools, and Applications." Rocky Mountain Institute and Lawrence Berkeley National Laboratory. LBNL report number LBNL-1007125. https://eta.lbl.gov/sites/all/files/publications/lbnl-1007125.pdf.

Granderson, Jessica, Philip N. Price, David Jump, Nathan Addy, and Michael Sohn. 2015. "Automated Measurement and Verification: Performance of Public Domain Whole-Building Electric Baseline Models." *Applied Energy* 144, 106–113. http://eis.lbl.gov/pubs/lbnl-187596.pdf.

Hawaii, Public Utilities Commission of. 2012 (January). Exhibit A. "Framework for Energy Efficiency Portfolio Standards". Section VIII. EEPS Evaluation, Measurement & Verification (EM&V). Decision and Order NO.30089, Docket No. 2010-0037. State of Hawaii Public Utilities Commission. http://www.hawaiicleanenergyinitiative.org/wp-content/uploads/2017/04/EE-Charrette_PUC_Decision-Order30089.pdf.

Hoffman, Ian M., Megan A. Billingsley, Steven R. Schiller, Charles A. Goldman, and Elizabeth Stuart. 2013. Energy Efficiency Program Typology and Data Metrics: Enabling Multi-State Analyses Through the Use of Common Terminology. LBNL-6370E. August. https://emp.lbl.gov/sites/default/files/lbnl-6370e.pdf.

Hoffman, Ian M., Steven. R. Schiller, Annika Todd, Megan A. Billingsley, Charles A. Goldman, and Lisa C. Schwartz. 2015. Energy Savings Lifetimes and Persistence: Practices, Issues and Data. Lawrence Berkeley National Laboratory. https://emp.lbl.gov/publications/energy-savings-lifetimes-and.

Indianapolis Power & Light Company, Evaluation Framework, 2015 (April). https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=6&cad=rja&uact=8&ved=0ahUKEwiksdW345vXAhWhrVQKHeRHArkQFgg-

MAU&url=https%3A%2F%2Fwww.iplpower.com%2FOur_Company%2FRegulatory%2FFilings%2FDSM%2FIPL_Miller_Direct_Testimony_and_Attachments_051717_compressed%2F&usg=AOvVaw35taXq19AWwhrhzDiozWDS.

Khawaja, M. Sami, Josh Rushton, and Josh Keeling. 2013. Chapter 11: Sample Design Cross-Cutting Protocols, NREL/SR-7A30-53827 April https://energy.gov/sites/prod/files/2013/11/f5/53827-11.pdf.

Lazar, Jim and Xavier Baldwin. 2011 (August). Valuing the Contribution of Energy Efficiency to Avoided Marginal Line Losses and Reserve Requirements. Regulatory Assistance Project. http://www.raponline.org/wp-content/uploads/2016/05/rap-lazar-eeandlinelosses-2011-08-17.pdf.

Lemoine, Peter, Kapil Kulkarni, and Tyler Huebner, and Val Jensen, 2008 (August). Integrating Risk Assessment into Energy Efficiency Program Portfolio Design. ACEEE Summer Study on Energy Efficiency in Buildings.

Malinick, Todd E., Jane S. Peters, Kevin Cooney, Greg Wikler, and Jay Luboff. 2017 (October 11). Evaluation Measurement & Verification Framework Refresh Needs Assessment CALMAC Study ID: SCE0414.01. https://pda.energydataweb.com/api/view/1940/EMV%20Refresh%20Final%20Report.2017.10.11.pdf.

Maryland 2016 (April 4). EmPOWER Energy Efficiency Programs Strategic Evaluation Guidance.

Maryland 2017 (October 5). EmPOWER Energy Efficiency Programs Strategic Evaluation Guidance.

Massachusetts Department of Public Utilities 2015 (October). Exhibit 1 2016-2018 Massachusetts Joint Statewide Three-Year Electric and Gas Energy Efficiency Plan. D.P.U. 15-160 to D.P.U. 15-169 Three-Year Plan 2016–2018. https://www.columbiagasma.com/docs/default-source/default-document-library/three-year-plan.pdf.

Northeast Energy Efficiency Partnerships (NEEP). 2016.

New York. 2016 (November). Evaluation, Measurement & Verification Guidance. Office of Clean Energy Clean Energy Guidance. http://www3.dps.ny.gov/W/PSCWeb.nsf/96f0fec0b45a3c6485257688006a701a/255ea3546df802b585257e38005460f9/$FILE/CE-05-EMV%20Guidance%20Final%20%2011-1-2016.pdf.

Pennsylvania Public Utility Commission. 2016 (August). Evaluation Framework for Pennsylvania Act 129 Phase II Energy Efficiency and Conservation Programs. http://www.puc.state.pa.us/Electric/pdf/Act129/SWE_PhaseIII-Evaluation_Framework082516.pdf.

Puget Sound Energy. 2015 (August). Exhibit 8: Evaluation, Measurement & Verification (EM&V) Framework. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwi6rqbD76nVAhVKh1QKHeOlDnwQFggpMAA&url=https%3A%2F%2Fwww.utc.wa.gov%2F_layouts%2F15%2FCasesPublicWebsite%2FGetDocument.ashx%3FdocID%3D15%26year%3D2015%26docketNumber%3D152075&usg=AFQjCNFc8fbN01LahFle57XhVV9lbz-cNA.

Schiller, S. R. 2012 (December). *Energy Efficiency Program Impact Evaluation Guide*. State and Local Energy Efficiency Action Network (SEE Action).

Schiller, Steven R., Greg Leventis, Tom Eckman, and Sean Murphy. 2017 (June). Guidance on Establishing and Maintaining Technical Reference Manuals for Energy Efficiency Measures. Prepared by Lawrence Berkeley National Laboratory for the State and Local Energy Efficiency Action Network.

Texas, Public Utilities Commission of (2017). Section 25. Substantive Rules Applicable to Electric Service Providers. Subchapter H. Electrical Planning. Division 2. Energy Efficiency and Customer-Owned Resources. §25.181. Energy Efficiency Goal. (q) Evaluation, measurement, and verification (EM&V). Effective date 3/30/17 https://www.puc.texas.gov/agency/rulesnlaws/subrules/electric/25.181/25.181.pdf.

Violette, Daniel M., and Pamela Rathbun. 2014 (September). Chapter 23: Estimating Net Savings: Common Practices. The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures. https://www.energy.gov/sites/prod/files/2015/02/f19/UMPChapter23-estimating-net-savings_0.pdf.

6.2. General References

Bonneville Power Administration. 2017 (April). Residential Lighting Market Characterization Study. https://www.bpa.gov/EE/Utility/research-archive/Documents/Momentum-Savings-Resources/2017_Residential_Lighting_Final_Report.pdf.

California Public Utilities Commission (CPUC). 2001. "California Standard Practice Manual: Economic Analysis of Demand-Side Programs and Projects." www.cpuc.ca.gov/workarea/downloadasset.aspx?id=7741.

National Efficiency Screening Project. 2017. "The National Standard Practice Manual for Assessing Cost-Effectiveness." Prepared by Tim Woolf, Synapse Energy Economics; Chris Neme, Energy Futures Group; Marty Kushler, American Council for and Energy-Efficient Economy; Steven R. Schiller; and Tom Eckman. https://nationalefficiencyscreening.org/national-standard-practice-manual/.

NEEP EM&V Forum. 2011. Glossary of Terms and Acronyms. http://www.neep.org/emv-forum-glossary-terms-and-acronyms.

Northeast Energy Efficiency Partnerships (NEEP). 2015 (August). Scoping the Certification of Energy Program Impact Evaluators. http://neep.org/scoping-certification-energy-program-impact-evaluators.

Schare, Stuart, Kevin Cooney, Jordan Mann, Andrea Lewis, Vijeta Jangra, Shannon Dorato, and Josh Arnold. 2015 (August). Why, When, and How to Utilize the "Right" Methods for Evaluation of Energy Efficiency Programs. In Proceedings of the EEDAL'15, Lucerne, Switzerland.

Schiller, Steven R., and Charles A. Goldman. 2011 (August). Developing State and National Evaluation Infrastructures—Guidance for the Challenges and Opportunities of EM&V. Steven R. Schiller, Schiller Consulting, Inc., Charles A. Goldman, Lawrence Berkeley National Laboratory. International Energy Program Evaluation Conference.

Schwartz, L. C., M. Wei, W. Morrow, J. Deason, S. R. Schiller, G. Leventis, et al. 2017 (January). Electricity End Uses, Energy Efficiency, and Distributed Energy Resources Baseline. Lawrence Berkeley National Laboratory.

Slattery, Bob. 2015. Reported Energy and Cost Savings from the DOE ESPC Program: FY 2014. Oak Ridge National Laboratory. https://energy.gov/sites/prod/files/2015/04/f21/2014_savings_espcs.pdf.

State and Local Energy Efficiency Action Network 2015 (September). Energy Efficiency Collaboratives—Driving Ratepayer-Funded Efficiency Through Regulatory Policies Working Group. Michael Li and Joe Bryson. https://www4.eere.energy.gov/seeaction/working-group/driving-ratepayer-funded-efficiency-through-regulatory-policies.

Glossary[53]

Evaluation, Measurement, and Verification (EM&V): The conduct of any or all of a wide range of assessment studies and other activities aimed at determining or assessing efficiency portfolio, program, project, or measure impacts or cost-effectiveness; assessing processes associated with program implementation; documenting the effects of an efficiency program, project, or measure; and understanding or documenting program, project, or measure performance, program or program-related markets and market operations as well as program-induced changes in energy efficiency markets, demand or energy savings, or program cost-effectiveness.. Sometimes the terms EM&V and evaluation are used interchangeably.

EM&V Framework: A primary, guiding document that describes a jurisdiction's EM&V infrastructure. It should cover fundamental topics and issues (such as evaluation definitions, objectives, and principles), EM&V scope topics (such as metrics, expected plans and reports, budget, timing, and roles and responsibilities), and impact evaluation approaches topics (such as baseline definitions and allowable, prescribed or proscribed EM&V methods). A framework also can cover other evaluation topics including those associated with cost-effectiveness analyses, process, and market evaluations and logistics.

EM&V Infrastructure: The components (e.g., policies, principles, metrics, processes, methods, products, organizational structure) deployed to conduct evaluation, measurement, and verification in a jurisdiction.

Impact Evaluations: Assessments of program-specific, directly or indirectly induced changes (e.g., changes in energy or demand use) associated with an energy efficiency program.

Market Evaluations: Studies that assess the energy efficiency marketplace. These include market effects studies, potential studies, and baseline studies. Market effects studies asses the change in the structure or functioning of a market or the behavior of participants in a market that result from one or more efficiency program efforts. Potential studies investigate how much efficiency savings might be available through various measures. Baseline studies determine indicators of market development before program intervention.

Measurement and Verification (M&V): Methods used to determine energy or demand savings at a single site or project by a combination of implementation verification, direct metering, agreed to or deemed calculations and analytical methods, and measurements and stipulations of key independent variables and factors. Commonly defined by International Performance Measurement and Verification Protocol (IPMVP) Options A, B, C, and D. Does not include the use of fully deemed savings values.

Process Evaluations: Systematic assessments of energy efficiency programs for the purposes of documenting program operations at the time of the examination, and identifying and recommending improvements to increase the program's efficacy or effectiveness for acquiring energy resources while maintaining high levels of participant satisfaction.

Protocol: Procedural methods and details for conducting one or more specific EM&V activities.

Program Administrator: An entity selected by a regulatory or other government organization to manage an energy efficiency portfolio within a specific geographic region or market. Typical administrators are publicly owned utilities, investor-owned utilities, nonprofit organizations, and state government agency, as may be determined by legislation or regulatory order.

[53] This glossary defines some key terms used in this guide and in the evaluation, measurement, and verification of efficiency programs. Most of the definitions contained in this glossary are derived from the glossary contained in Schiller 2012. Some of the definitions in this glossary reflect usage in the specific context of efficiency programs that have an oversight (regulatory) body; thus, they might not be applicable to non-regulatory contexts such as in commercial agreements between energy services companies and their clients. Additionally, definitions of terms do vary across different jurisdictions and thus readers are encouraged to seek clarity on the definitions used or to be used in specific EM&V frameworks.

Program Evaluator: A person or entity that conducts evaluations or EM&V associated with an efficiency program. Often the adjectives "third-party" and "independent" are used to further describe program evaluators. The designation of independent or third-party is determined by those entities involved in the use of the evaluations and can include evaluators retained, for example, by the program administrator or a regulator.

Program Implementer: An entity selected and contracted with or qualified by a program administrator to provide products and services to consumers either directly or indirectly.

Stakeholders: Entities with an interest or concern associated with the planning, implementation, or evaluation of efficiency portfolios. These can include state and local governments, utility regulators, administrators of energy efficiency programs, efficiency program implementers, evaluation consultants, environmental groups, industry representatives, trade allies, consumer advocates, and individual consumers.

Technical Reference Manual: A resource that contains energy efficiency measure information used in program planning, implementation, tracking, and reporting, and for evaluation of impacts associated with the subject measures.

Appendix A. Energy Efficiency EM&V Background—Impact Evaluations

Impact evaluation includes a range of retrospective assessments and activities aimed at determining the effects of efficiency policies, portfolios, programs, projects, or measures. Impact evaluation is one of four broad categories of efficiency evaluations: impact evaluations, process evaluations, market evaluations, and cost-effectiveness evaluation. Impact evaluations document direct and indirect performance metrics, such as energy and demand savings and avoided air emissions.

A.1. Impact Evaluations Fundamentals

The impacts of end-use efficiency (and demand response and conservation) activities, such as energy and demand savings, cannot be directly measured. Instead, impacts are based on a comparison between what happened and a set of assumptions about what would have happened under the same set of operating conditions (i.e., the baseline or counterfactual scenario). In effect, efficiency impacts are always "estimates." The need for counterfactual assumptions (*see* text box below) results in inherent uncertainty and adds complexity to the impact evaluation process.

KEY DEFINITIONS FOR DETERMINING ENERGY

Baseline: Conditions, such as energy consumption and demand, which would have occurred without implementation of the subject energy efficiency measure. Baseline conditions are sometimes referred to as the counterfactual. There are several options for establishing baselines and a range of definitions for these options used in the efficiency industry.

Gross Savings: The difference between energy consumption of the affected equipment or facility with versus without the EE project or EE measure in place, without consideration of program influence or attribution. Gross savings is calculated relative to a specified baseline determined without regard to program influence.

Net Savings: The difference between energy consumption with the program or intervention in place and that which would have occurred absent the program or intervention, accounting for program influence and attribution.

Operating Conditions: The conditions in which the EE project or EE measure or affected structure or equipment is used or operated.

Further, for a given program or project, the specific impact evaluation (or EM&V) method that is applied depends on a number of factors, including the type of efficiency activity, overall policy objectives, access to data, and available budgets. Thus, these factors and the counterfactual result in the need to balance the accuracy of savings estimate against the cost and effort to determine that estimate. EM&V practitioners select one of the three EM&V methods described below, or one or more of the numerous variations of these methods, that they believe creates the right balance of cost, accuracy, and timeliness for the subject measure(s).

DEFINING THE COUNTERFACTUAL SCENARIO

Energy savings and associated impacts of efficiency actions are estimated to varying degrees of accuracy by comparing the situation (e.g., energy consumption) after an efficiency measure is implemented (the reporting period) to what is assumed to have been the situation in the absence of the measure (the "counterfactual" scenario, also known as the baseline). For energy impacts, the baseline and reporting period energy use are compared, making adjustments for factors unrelated to efficiency actions (such as weather or building occupancy). These adjustments are a major part of the evaluation process; how they are determined can vary from one measure type to another and between EM&V methods.

There is also variation in the definitions and assumptions used for establishing the (counterfactual) baseline. For example, common practice, preexisting condition, and codes and standards are examples of different options for baselines are selected throughout the industry. Thus, a framework document can be a vehicle for establishing clear baseline definitions. See Section 3.3.2 for some example baseline categories, and Appendix B for standard industry resources that address and define baseline issues.

EM&V has been used primarily for, and is most advanced for, utility customer–funded energy efficiency and demand response programs as well as performance-based projects implemented directly by energy service companies (ESCOs) for their clients. Thus, efficiency EM&V strategies in wide use today—including budget levels, oversight procedures, and preferred methods—are mostly derived from utility regulatory agency requirements, together with industry standard energy efficiency guides and protocols developed to support regulatory requirements and ESCO projects. For those interested in more information on EM&V practices and resources, refer to Appendix B: Energy Efficiency EM&V Resources.

A.2. Impact Evaluation Methods

"Evaluation" is the typical term associated with assessing programs (and program portfolios and policies). "Measurement and verification (M&V)" is a method associated with assessing project and individual measure impacts; it also is one way that programs are evaluated. For example, M&V can be applied to a sample of projects, and the results extrapolated to the entire program population of projects.

Besides M&V methods, there are two other methods commonly used for efficiency program impact evaluation: (1) *deemed* (also called *unit* energy *savings* or *stipulated*) savings methods and (2) comparison group methods. Solely using fully deemed savings values is not considered M&V. Measurement and verification, as defined by the efficiency industry, always requires some level of site measurements. These three methods are described briefly below.

Deemed savings method. These are processes by which fully deemed savings values are determined and applied. The deemed savings method can also overlap with the M&V method, as both can involve developing deemed variables, factors, and calculations, which are used for determining fully deemed savings values (deemed savings method) as well as partially deemed savings values (M&V method).

The focus of the deemed savings method, however, is the fully deemed savings values that are stipulated estimates of energy or demand savings (or potentially other impacts) for a single unit of an installed efficiency measure that:

- 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
- Are applicable to the condition (e.g., office building lighting system retrofit, residential refrigerator upgrade) under which the measure is being implemented.

EVALUATION, MEASUREMENT, AND VERIFICATION METHODS

"Evaluation, measurement, and verification (EM&V) is a process of assessing an energy efficiency program, including applying M&V and other methods to estimate program savings. EM&V can include:

- The M&V methods applied at the building level, with results expanded to the program level.
- The use of deemed savings values, with installations and key parameters verified by the evaluator, but without direct measurement of site performance (thus deemed savings is not considered a true M&V approach). [deemed savings method]

Analysis of consumption data for program participants and a comparison group to determine savings for the program as a whole, and not necessarily for any individual facility or measure. [comparison group method]."[54]

As part of the deemed savings method and to fully quantify impacts, a separate verification process is usually needed to confirm the quantity of measures installed (and for some programs, whether they are operating correctly) and that the measure implementation conforms to the conditions and applications (e.g., installation specifications) defined for use of the deemed savings value. More information on the deemed savings method can be found within Schiller et al. 2017.

[54] Franconi et al. (2017), p. 7.

The deemed savings method is used to stipulate values (i.e., unit energy, demand savings) for projects with well-known and documented savings values and for which there is a strong central tendency in the distribution of savings across sites or installations—that is, not much variation in savings across most installations. Examples include energy-efficient appliances such as washing machines, computer equipment, and refrigerators as well as lighting retrofit projects with well-understood operating hours. Many performance contracting projects document their savings with deemed savings values.

FULLY DEEMED VERSUS PARTIALLY DEEMED VALUES

M&V is distinguished from deemed savings methods by M&V's requirement for some field-based or project-specific measurements—which is not a requirement for deemed savings. Without any such measurements, a value is a *fully* deemed savings value, period. However, because M&V can involve the use of predetermined (deemed) calculations and values or factors (such as those for long-term weather data) there is some possible overlap in terminology when savings values assigned to a measure are essentially *partially* deemed. In this guide, the separation between M&V and deemed savings methods is defined such that deemed calculations—which result in partially deemed savings values—are an element of M&V.

A.3. Measurement and Verification Methods

The industry standard M&V document—the International Performance Measurement and Verification Protocol (IPMVP)[55]—defines four M&V options, two retrofit isolation options and two whole-facility options.

- *Retrofit isolation options:* Assessing savings from each efficiency measure individually (IPMVP Options A and B). Verification is an integral part of Options A and B because the measurement process involves direct observation of all or a sample of the affected equipment.
    - Option A: Savings are determined by field measurement(s) of the *key* performance parameter(s) that define the energy use of (and thus savings of) a measure or project and stipulation of other factors. Option A thus could be considered a partially deemed savings value approach, resulting in partially deemed savings values.
    - Option B: Savings are determined with field measurement(s) of *all significant* performance parameter(s) that define the energy use of (and thus savings of) a measure or project; unlike Option A, Option B does not allow stipulations of any major factors.[56]
- *Whole facility options:* Collectively assessing energy (and demand) savings from all energy efficiency measures in a facility, taking into consideration the interactions between measures and systems within the facility.
    - Option C: Facility energy meter(s) data are used to compare energy use before and after implementation of the efficiency measures.
    - Option D: Calibrated simulation models[57] are used to estimate energy use before and after measure implementation. Option D often is used with new construction efficiency actions, as the baseline does not exist but can be simulated.

[55] Efficiency Valuation Organization (EVO). International Performance Measurement and Verification Protocol (IPMVP) (multiple dates). http://evo-world.org/en/products-services-mainmenu-en/protocols/ipmvp.

[56] For example, in a lighting retrofit, the parameters may be a change in wattage and operating hours. With Option A, only operating hours might be measured and a change in wattage is stipulated. With Option B, they would both be measured over some specified period.

[57] Whole-building and building component energy simulation programs are physics-based tools that engineers, architects, and researchers use to model both energy consumption (for heating, cooling, ventilation, lighting, and plug and process loads) and energy savings opportunities in buildings. A wide variety of building energy simulation programs have been developed and enhanced, and are in use throughout the building industry.

One study of the U.S. Department of Energy's (DOE's) Energy Savings Performance Contract (ESPC) program further indicated that for those ESCO projects, the most common M&V approaches were IPMVP Options A and B.[58]

Options A and B have historical limitations associated primarily with cost of metering (equipment and labor), which project participants might not be interested in paying for, particularly over the life of projects. This could change with the development of what is known as "M&V 2.0."[59] Option D, calibrated computer simulations, is used when the savings for individual measures are desired but only whole-premise metered data are available.

A.4. Comparison Group EM&V

This method involves determining program savings based on the differences in energy consumption between a comparison group and program participants. Comparison group methods include randomized control trials and quasi-experimental methods.[60]

Because the effects of implemented measures are reflected in the observed participant-comparison differences, separate verification is not typically required. Control groups have been used for decades for residential efficiency programs with large numbers of relatively homogenous participants. There has been renewed interest in this method for a wide range of program types, as a potential gold standard of savings determination. Some M&V 2.0 applications also employ this method.

At least in theory, comparison group analyses assess the savings associated with just the efficiency activity and not changes in energy consumption or demand associated with outside factors, such as changes in the economy and energy prices or savings from those consumers who would have completed the projects outside of program influences (i.e., "free riders").[61] This is done by comparing data between a treatment group (participants) and a control group of consumers that are determined to be statistically similar. The challenges for comparison group approaches include reasonably applying them to populations of non-homogenous, customized projects (such as efficiency in commercial, institutional, and industrial facilities) and structuring a control group. Particularly if done randomly (at least in part to avoid self-selection biases), that might mean that some eligible consumers do not get to participate in the efficiency activity.[62]

A.5. Impact Evaluation Activities

With respect to the activities that take place as part of an impact evaluation of efficiency programs, Figure A.1 indicates another means for organizing the components, or toolbox, of efficiency EM&V. This figure shows that beyond measurements, verifications, and analyses, there often are reviews that are done of the tracking systems used to record efficiency actions by the implementers or administrators of the programs, reviews of project documentation (referred to as "engineering reviews" in the figure, but also known as "desk reviews"), and interviews with participants (utility customers) as well as others involved in program implementation (e.g., contractors and equipment suppliers). Figure A.2 shows that the major activities can be conducted for

[58] Slattery 2015.
[59] For more information on M&V 2.0, see Franconi et al. (2017).
[60] Randomized control trials (sometimes referred to as *full experimental designs*) are evaluations that derive savings estimates by comparing the energy use of customers who are randomly assigned to receive an energy efficiency measure to a control group that does not. Randomization minimizes self-selection bias, and the different comparison groups enable the evaluator to determine the impact of the measure when compared with the no treatment (control) group, and other variables are kept constant. In practice, this often is problematic because consumers generally self-select to participate in programs, so quasi-experimental evaluation methods are more frequently used. A quasi-experimental estimate of the savings still compares the energy use of participants with a control group, but without random assignment of consumers to either a control group or participant group.
[61] In practice, how well the control group method determines true incremental, net impacts depends on the specific approach applied (randomized control trials are more reliable than quasi-experimental methods) and how well the approach is implemented.
[62] This is particularly problematic for where the anticipated energy savings from a measure or program is small (e.g., a few percentages) relative to the total energy consumption. In such cases, very large samples of both the control and treatment groups are required for determining whether there is a statistically significant difference in use between the two groups.

different subsets of the projects in a program—generally larger to smaller subsets as the activities move from less to more complex/expensive efforts.

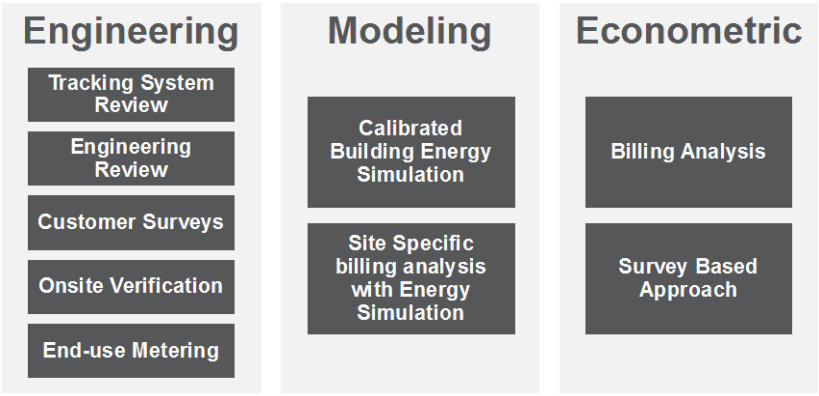


Figure A.1. EM&V activities (Schare 2015)

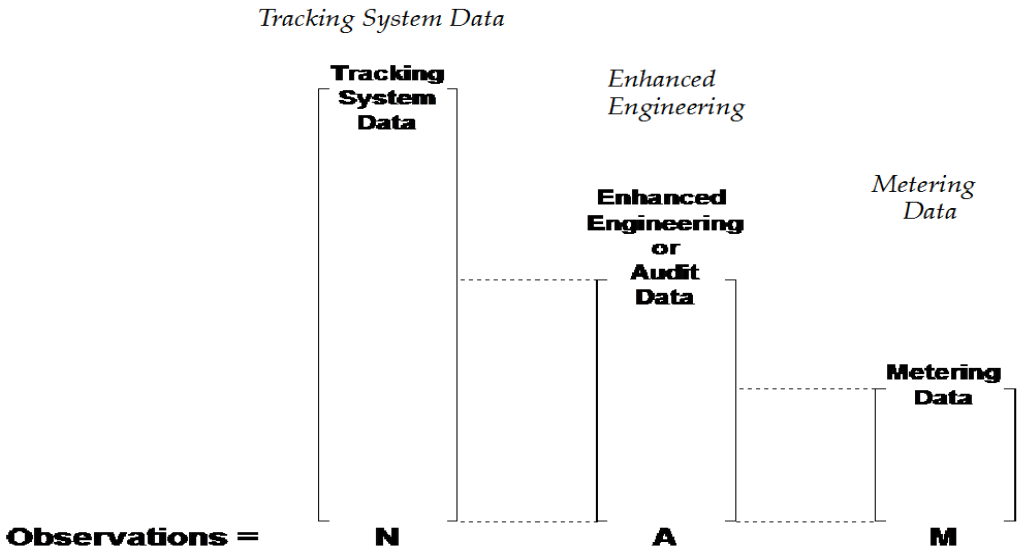


Figure A.2. Leveraging EM&V activities data with different sample sizes (Cooney 2015)

Appendix B. Energy Efficiency EM&V Resources That Are Focused on Impact Evaluation

B.1. SEE Action Energy Efficiency Program Impact Evaluation Guide[63]

This industry-standard guide to EM&V describes and provides guidance on concepts and methods for determining and documenting energy and non-energy impacts resulting from energy efficiency programs and portfolios of programs funded by utility customer funds. It specifically focuses on impact evaluations for programs designed to reduce facility energy consumption, demand, or both—as well as related air emissions.

B.2. SEE Action EM&V Portal[64]

The State and Local Energy Efficiency Action Network (SEE Action) offers resources, discussion forums, and technical assistance to state and local decision makers as they provide low-cost, reliable energy to their communities through energy efficiency.

B.3.International Performance Measurement and Verification Protocol[65]

The International Performance Measurement and Verification Protocol (IPMVP) provides an overview of current best practices for determining and verifying results of energy efficiency Internationally; it is the most recognized M&V protocol for demand-side energy activities. The IPMVP provides a framework, and definitions that can help practitioners develop M&V plans for their projects.

B.4. Uniform Methods Project[66]

Published by DOE, Uniform Methods Project protocols provide standardized, common-practice M&V methods for determining gross energy savings for many of the most common residential and commercial measures and programs offered by administrators of energy efficiency programs in North America for utility customers. The UMP also includes cross-cutting protocols for topics such as net savings determination, metering, and persistence of savings determination.

B.5. FEMP M&V Guidelines: Measurement and Verification for Performance-Based Contracts, Version 4.0[67]

Prepared for DOE's Federal Energy Management Program, the purpose of this document is to provide guidelines and methods for documenting and verifying the savings associated with federal agency performance contracts. It contains procedures and guidelines for quantifying the savings resulting from energy efficiency equipment, water conservation, improved operations and maintenance, renewable energy, and cogeneration projects.

B.6. ASHRAE Guideline 14-2014: Measurement of Energy and Demand Savings. American Society of Heating, Refrigerating and Air-Conditioning Engineers[68]

Guideline 14 provides a standardized set of energy, demand, and water-savings calculation procedures. This publication provides guidance on minimum acceptable levels of performance for determining energy and demand savings, using measurements, in commercial transactions.

---

[63] https://www4.eere.energy.gov/seeaction/publication/energy-efficiency-program-impact-evaluation-guide.
[64] https://www4.eere.energy.gov/seeaction/topic-category/evaluation-measurement-and-verification.
[65] Efficiency Valuation Organization (EVO). International Performance Measurement and Verification Protocol (IPMVP). http://evo-world.org/en/products-services-mainmenu-en/protocols/ipmvp.
[66] http://energy.gov/eere/about-us/ump-protocols.
[67] http://energy.gov/eere/femp/downloads/mv-guidelines-measurement-and-verification-federal-energy-projects-version-40.
[68] American Society of Heating, Refrigerating and Air-Conditioning Engineers (www.ashrae.org).

B.7. M&V Guidance from Regional Transmission Organizations

These are protocols for quantifying and verifying the demand reduction value of EE programs, projects, and measures for the forward capacity markets in these regional transmission organizations (RTOs).

- ISO-NE Measurement and Verification of Demand Reduction Value from Demand Resources—Manual M-MVDR (2014). https://www.iso-ne.com/static-assets/ documents/2017/02/mmvdr_measurement-and-verification-demand-reduction_rev6_20140601.pdf
- PJM Manual 18B: Energy Efficiency Measurement & Verification (2016). https://www.pjm.com/~/media/documents/manuals/m18b.ashx