

Evaluation Planning, Methods, and Practices Across the United States: Emerging Issues and Opportunities for Cooperation

Michael Messenger, Itron, Oakland, CA.,

Charles Goldman, Lawrence Berkeley National Laboratory, Berkeley, CA.

Steven R Schiller, Schiller Consulting Inc. Oakland, CA.

Abstract

Understanding the strengths and weaknesses of the current methods and practices used to evaluate the effectiveness and energy savings from energy efficiency programs is likely to become increasingly important for electricity policy makers as funding levels and projected savings from these programs continues to grow. This paper presents the results of a survey of efficiency policy experts, evaluation project managers, and evaluation practitioners in fourteen states and the Pacific Northwest on the following areas:

1. key purposes, users, and uses of evaluations of energy efficiency programs;
2. the pros and cons of methods used to evaluate these programs and report the results;
3. the processes used to determine what types of evaluation to pursue and how much to spend; and
4. emerging issues in the energy efficiency program evaluation field.

The survey results suggest that the relative importance of resolving these issues will vary depending on the outcome of various efforts to develop energy efficiency resource and or carbon savings goals at the national and or regional levels. We conclude with a summary of recommended evaluation issues that merit further analysis and potentially engagement or resolution by evaluation organizations with an interest in increasing the effectiveness of evaluation, measurement and verification efforts.

Introduction and Methodology

This project was conducted on behalf of the Leadership Group of the National Action Plan for Energy Efficiency to support its efforts to improve the effectiveness of evaluations conducted across the country and promote the use of common terms to describe the impacts and effectiveness of energy efficiency programs. The results presented here are the first step in a planning process designed to identify the most important opportunities for the organization to address emerging evaluation issues and then develop an action plan to catalyze change or improvements in selected issue areas

Three survey questionnaires were developed in this study to elicit both information and opinions about EM&V practices. One questionnaire was targeted to policymakers and regulators in each of the 14 case-study states. The second questionnaire was targeted to program administrators (e.g. utilities, Energy Trust of Oregon, NEEA) and practitioners (i.e. evaluators and consultants). The third questionnaire was targeted at national-level EM&V experts who have worked for several years in multiple jurisdictions and in different capacities in the U.S.

The survey was distributed to a total of 90 state/regional experts and 17 national EM&V experts. The response rate was 59% for the state experts and 65% for the national experts. Response rates were higher than the expected response rate of 50% given the summer holiday season that made it difficult to reach respondents. The fourteen states selected in this study span range of experience with ratepayer-funded EE programs from states just starting out to states with over 25 years of experience. These states include California, Connecticut, Florida, Iowa, Idaho, Illinois, Massachusetts, Maine, Minnesota, New York, Oregon, Pennsylvania, Texas, and Wisconsin. In addition, we interviewed experts at the Northwest Energy Efficiency Alliance that covers several states in the Northwest.

Findings

Key Uses of EM&V studies

Figure 1 shows how respondents ranked the relative importance of key uses of EM&V studies for energy efficiency programs. The most important current uses reported by respondents can be grouped into two categories, those receiving scores above or below the average ranking of important. The uses related to direct estimation of net and gross load impacts estimates , assessing cost effectiveness of programs and assessing the cost effectiveness of program designs received the top 3 rankings while uses that focused on indirect uses of these savings estimates such as assessing progress toward energy and environmental goals , use in resource planning or use in quantifying participation payments in forward capacity markets received rankings lower than 4.0 or a score of very important.

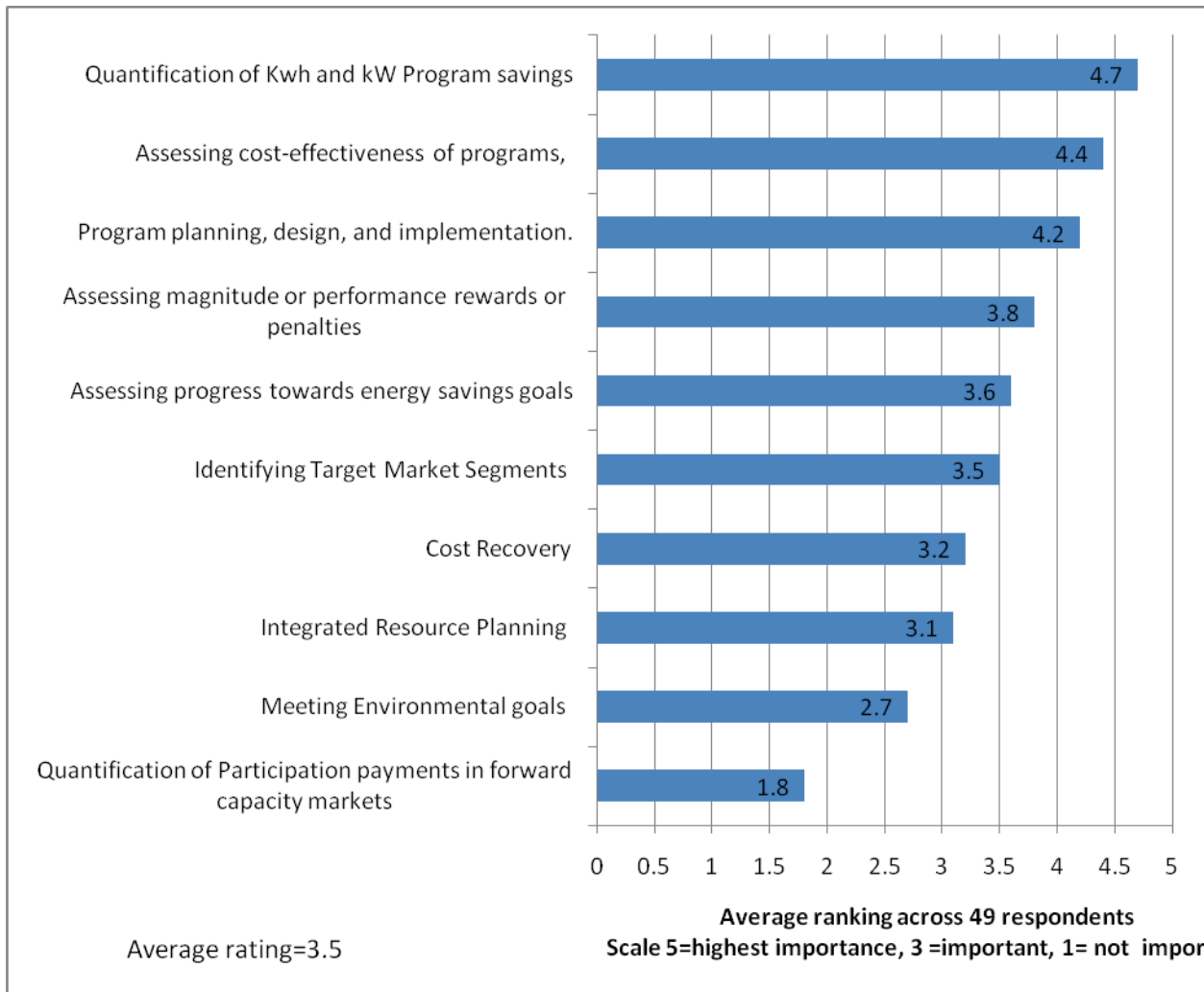


Figure 1- Average Ranking of the Importance of Evaluation Objectives

Key Users or Audiences for EM&V studies

Respondents indicated the most important users of EM&V studies include regulators, and utility program managers – see Figure 2. Third-party implementers and state energy offices are also an important audience for the EM&V studies. Consumer groups, environmental agencies, reliability organizations, Independent System Operators, and Energy Service Companies were seen as somewhat less important while market monitors were ranked as the least important users of EM&V studies.

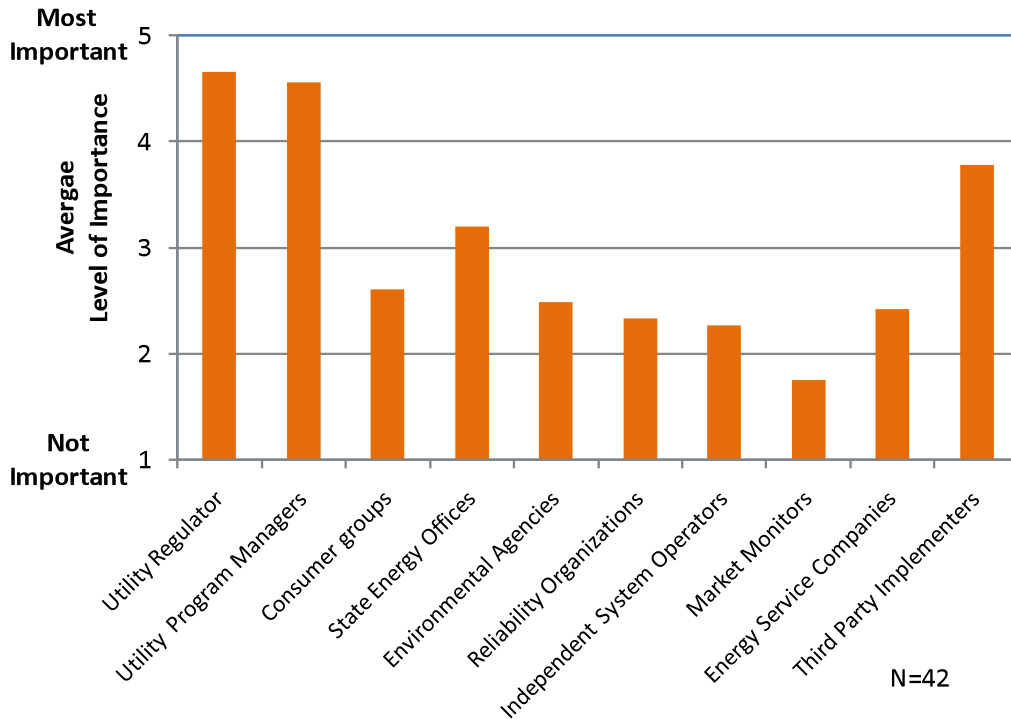


Figure 2. Ranking of Most Important Users of Energy Efficiency Evaluations

Planning of EM&V Studies

Survey respondents were asked to select among three alternative processes that are currently used to define the approach taken on which entities have responsibility for initiating and leading the overall planning of evaluation studies and activities in their states (e.g., evaluator, program administrator, and/or regulatory or oversight body;). Table 1 presents these results.

Table 1. Comparison across all case-study states and NEEA with respect to process used for selecting types of evaluation research

State	Evaluation planning is initiated and completed by:		
	Evaluator or firm hired to conduct evaluation studies based on objectives or goals from the program administrator	Program administrator in a public setting within budget constraints set before	Regulatory or oversight body who will ultimately approve or adopt a research plan & budget
CA		X	X
CT			Evaluation Planning performed in a collaborative setting
FL	X	X	
IA	X	X (but not public setting)	
ID	X	X	
IL	X	X	X
MA			
ME			X
MN	X	X	
NEEA		X	
NY	X	X	X
OR	X	X	
PA			X
TX		X	
WI	X		

Readers note: X’s in multiple columns indicate there was disagreement among respondents about the most important process used to conduct evaluation planning in this state.

Scope and Level of Methods to Estimate Energy Savings

Evaluators have chosen to evaluate programs at different levels of detail depending on administrator needs and regulatory policy objectives in each state. Our survey found evaluation foci at four different levels:

1. Average Savings over a population for specific energy efficiency measures;
2. Average Savings at the end use level where more than one measure may have been installed;
3. Average Savings at the Program level; and
4. Average Savings at the Portfolio level;

The granularity of savings estimates chosen is a function of both regulatory policy and the extent to which program administrators rely on deemed or ex-ante savings values per efficiency measure or system installed. Veteran states tend to focus on estimating savings at the measure level or end use level because these results are needed to assess the cost-effectiveness of the program and understand

differences between planned and achieved program savings. However many states currently only provide public estimates of energy and peak savings at the program or portfolio level.

Table 2 shows the majority of the case-study states either have or are moving toward development of standard measure data or deemed savings data bases. Practices differ in terms of whether use of deemed savings is mandatory for program administrators or encouraged and whether deemed values are verified by an independent party. Three states (CA, CT, and MA) require that the EM&V methodology incorporate uncertainty of various estimates while IL is considering this requirement. The majority of states either have or are considering an audit requirement to verify a sample of installations. In three states (CA, NY, and WI), EM&V protocols exist or they require the use of specific methods to evaluate programs. States such as FL and IA rely, primarily, on IPMVP protocols while IL and PA are considering the development of protocols.

Table 2. EM&V Guidance and Reporting Requirements Across States

State	Is there a Technical Resource Manual or Deemed Savings database?	Requirement to identify range of uncertainty in program savings estimates?	Audit requirement to verify sample of installations?	EM&V protocols exist to guide or require use of specific methods to evaluate program or market savings?	Requirement to report Gross or Net Savings or both?
CA	Yes	Yes	Yes	Yes	Both
CT	Yes	80/20	Part of program	No	Gross
FL	No	No	No	No; Rely on IPMVP	Both
IA	No	No	No	No; Rely on IPMVP	Sometimes both
ID	No	No	No	No	No
IL	Not yet	No, but likely	Yes	Not yet, in process	Both
MA	Yes; in next two months	Yes	Yes for most programs	No	Net
ME	Yes	No	No	No	Gross
MN	Yes	No	Yes	Yes for custom projects	Gross
NEEA	No	No	No	No	Sometimes both
NY	Yes	Not yet	Yes	Yes	Net
OR	No	No	No	No	Sometimes both
PA	Planned	No	Under review	Not yet	Under review
TX	Yes	No	In practice yes	No	Gross
WI	No	No	Yes	Yes	Both

Reporting Conventions and Requirements by Governing or Oversight Organizations

There is tremendous variation among all case-study states and NEEA with respect to which type of program savings metrics (i.e. Net or Gross or both) are required to be reported. Four states require both net and gross savings to be reported in all completed load impact evaluations (California, Florida,

Illinois, and Wisconsin) while three states require both to be reported on an optional basis(Iowa, NEEA and Oregon).

In addition to this reporting terms variation, the effects included when calculating gross and net program savings also varied significantly across the sample states. Table 3 shows which specific effects related to net savings are analyzed in respective states. In the majority of states, free-ridership and spillover/market effects are analyzed and ex-ante savings estimates are trued-up based on ex-post evaluations. In California, free ridership is always estimated for all EE programs but market effects are only periodically evaluated on an ad hoc basis. Leakage effects when incented products move across state lines are also estimated in California.

Table 3. Types of effects analyzed in load impact evaluations

State	Free-ridership	Leakage to Other States?	Spillover/ Market Effects	Truing-up of ex-ante estimates based on ex-post evaluations
CA	Yes	Yes and No	Yes in few cases*	Yes
CT	Yes	NA	Yes in some cases	Yes
FL	Yes	No	Yes	Yes
IA	No	NA	No	No
ID	No	No	No	Yes
IL	Yes	NA	Yes	Yes
MA	Yes	No	Yes	Yes
ME	Yes	No	No	No
MN	Yes	No	No	No
NEEA	No	No	No	Yes
NY	Yes	No	Yes	Yes
OR	Yes	No	Yes	Yes
PA	No	No	NA	NA
TX	No	No	No	No
WI	Yes	No	Yes	Yes
Total Yes	10	1	8	9

Roles and Responsibilities for Executing Evaluation Plans

In all case-study states, except ME and PA, the actual evaluation of most programs is carried out by third-party contractors (or evaluators/practitioners) and managed by program administrators or regulatory staff. In many cases, the process of designing the evaluation plan for process or impact evaluations was collaborative with substantial inputs and oversight by both regulators and program administrators. In two states (CA and ME), regulators directly manage load impact evaluations. Interviews conducted by project team suggest that the trend is toward more involvement by regulatory staff in the actual evaluation process. In some states, the utilities’ and/or program administrators’ internal staff will conduct evaluations of specific programs, however, these are, typically, smaller efforts.

Evaluation Project Funding

Data on EM&V budgets for energy efficiency programs is compiled by the Consortium for Energy Efficiency and their recent report (2008) provides information for at least some of the case-study states. However, information about how the EM&V budget is allocated is not easily available in the public domain. In the survey, the respondents were asked to provide data from the most recent year available about three elements of EM&V budgets – size, comparison with total program budget, and allocation among three main types of EM&V studies (i.e. impact, process, and market research).

Table 4 indicates that size of the annual EM&V budget varies from \$45 million in CA to \$500,000 in ME. When normalized with respect to the total EE program budget, however, we find that in two states, the EM&V budget exceeds 5% while in two other states it is ~1% or less. For the majority of the case-study states and NEEA, the fraction of EE budget allocated for EM&V purposes varies from 2 to 5%.

Table 4. EM&V funding levels across States in Sample

State	EM&V Funding (million \$)	EM&V Funding as % of Total EE Funding	Allocation of EM&V Budget (% Impact / % Process / % Market Research)
CA	25 to 45*	> 5%*	75 / 15 / 10
CT	2.0	3%	NA
FL	NA	NA	NA
IA	8	> 5%	50 / 30 / 20
ID	NA	NA	NA
IL	2.6	2-3%	75 / 15 / 10
MA	6.3	3-5%*	75 / 15 / 10
ME	0.5	2-3%	30 / 50 / 20
MN	0.7	3-5%	NA
NEEA	NA	3-5%	60 / 30 / 10
NY	8	3-5%	80 / 10 / 10
OR	2*	2-3%*	50 / 30 / 20
PA	NA	<1%	50 / 30 / 20
TX	0.8	<1%	75 / 15 / 10
WI	3 to 4	3-5%	100 / 0 / 0

* = includes funding for evaluating codes and standards also.

The allocation of EM&V budget among different types of studies also varies substantially across states. The portion of EM&V budget allocated for impact studies was the highest (ranging from 50% to 100%) in all states except ME where process studies account for the highest (i.e. 50%) allocation. WI does not allocate any budget to either process or market research studies. Respondents from four states (CT, FL, ID, and MN) were unable to provide estimates of the budget allocations by study type.

Emerging Evaluation Issues

Six key EM&V issues were identified as likely to be important to resolve based on the observed differences in evaluation methods and reporting practices in the 14 states interviewed. These issues were identified through interviews conducted with a wide range of national and state EM&V experts as well as the expert opinion of the report authors.¹ The issues are grouped into six major categories:

1. Consistency in Reported Savings Issues- Differences between jurisdictions in how program savings are estimated and how they are defined (e.g., net versus gross) makes comparison of reported savings and their comparison against energy sales or growth benchmarks very difficult if not impossible. (See Reference 2 for a comprehensive glossary of EM&V terms)
2. Net Savings Method Issues– Although standard project and program EM&V methods are well documented for determining energy and demand savings, there is no standard agreement on how to address such basic measurement issues as: (a) how, if at all, to address program attribution; (b) how to define and set standards for rigor and accuracy given different policy objectives, and (c) how to assess broader market effects/impacts of energy efficiency programs on the future demand for energy services.
3. Quality Control and Accuracy Issues- The quality of program evaluations produced in different states and regions is perceived to vary widely because of differences in the level of independent review of program saving estimates required in each state and the failure to require that the level of uncertainty associated the program savings estimates be reported.
4. Evaluation Resource Allocation Issues - The wide range of program planning processes and state by state (versus comprehensive national) approaches used to allocate public funds to evaluation projects may be biased toward simple program savings reporting and underfunding analysis on the effectiveness of alternative program designs and market changes stimulated by the programs.
5. Independence of Evaluator compared to working cooperatively with Program Administrators – EM&V has two primary objectives, (1) to assess the savings resulting from an EE program or portfolio and (2) to provide feedback for program/savings improvement. Particularly in a regulatory environment where program savings levels affect administrator compensation, there is a natural friction between the need for independence between program administrators and evaluators when estimating program savings and the need for a close working relationship between evaluator and program administrators/implementers to provide useful feedback on programs. This tension combined with the need for more timely evaluation feedback has often resulted in less funding being available for efforts to assess the effectiveness of, and hopefully improve, program designs.
6. Integration of EE evaluation load impact results in utility planning and forecasting –Failure to address the analytic challenges associated with aggregating estimates of net program savings into load forecasting frameworks may result in under- or over-counting of the impacts of EE programs on current and future economy wide energy use and the resulting GHG emissions.

¹ The opinions expressed in this paper represent the analysis and opinions of the LBNL-Itron team and do not necessarily represent the views of the Technical Working Group or its EPA sponsors.

Importance of Resolving Evaluation Issues is Likely to be Dependent on Evolution of Energy Efficiency Policy at the Federal and Regional Levels

Respondents to this survey reported that the relative importance of solving critical evaluation issues was dependent on how legislative efforts to develop national savings or carbon reduction goals evolve over the next five years. Given this observation, we developed three alternative energy efficiency futures designed to elicit survey respondent views on the overall importance of each EM&V issue group. Evaluation priorities and practices are driven by state policy in Scenario 1 (Business as usual), by national policy in Scenario 2 (National Drivers), and by regional initiatives in Scenario 3 (Regional Drivers). Table 5 describes each scenario and presents the results of the LBNL team assessment of the likely relative importance of each EM&V issue group in the three energy efficiency scenarios.

Table 5. Future Scenarios for Energy Efficiency: Key EM&V Issues

Description of Scenario	Relative Importance of Solving EM&V Issues (Source of rankings: LBNL team)
<i>Scenario #1 – Energy Efficiency Markets under “Business as Usual” Conditions</i>	
<p>Interest and funding for EE programs continues to grow at the local and state level in response to legislative and regulatory requirements and other drivers (e.g. high, volatile energy prices, high cost and regulatory uncertainty surrounding supply-side resources, concerns about climate change). Program funding levels grow from \$3.1 billion in 2008 to \$7.6 billion by 2020 (see Ref #1, Medium scenario in Barbose et al 2009).</p>	<p>Consistency Issues- Low Measurement Issues- High Quality Control and Accuracy- Medium Evaluation Resource Planning- Medium. Timing/Feedback issues- Low to Medium Integration of Program Results into Forecasting- Low to High</p>
<i>Scenario #2 – Energy Efficiency Markets under a National EERS or Climate Change legislation</i>	
<p>In this scenario, Congress enacts some form of nationwide energy efficiency savings goals, which may or may not be part of national carbon legislation. A national Energy Efficiency Resource Standard contains requirement for nationwide reporting of program savings and a federal agency is given the responsibility of developing consistent EM&V protocols that are used by entities to report on their progress towards achieving the EERS.</p>	<p>Consistency of Program Savings terms- Very High Measurement Issues- Medium Quality Control and Accuracy- High Evaluation Resource Planning- Medium. Timing/Feedback issues- Low Integration of Program Results into Forecasting- Low to High depending on adopted savings goals metrics</p>
<i>Scenario #3 – Energy Efficiency Markets effected by Regional Policies</i>	
<p>EM&V practices and reporting requirements are driven by the efforts of regional planning and/or reliability organizations (e.g. ISO, RTO) to standardize evaluation methods for use in regional transmission planning processes and wholesale markets (e.g. forward capacity markets). Program Administrators begin to recognize and achieve economies of scale in EM&V by working with regional EE organizations (e.g. NEEP or NEEA) to develop common evaluation methods and data.</p>	<p>Consistency of Program Savings terms- Medium Measurement Issues- High Quality Control and Accuracy- Low to High Evaluation Resource Planning- Low Timing/Feedback issues- Low Integration of Program Results into Forecasting- Very High (depending on type of adopted savings goals metrics)</p>

Informal polling of experts in the Evaluation Technical Working groups revealed that Measurement Issues related to Program Attribution (EM&V Issue Group #2) and Efforts to increase Quality Control and Accuracy of Load impact estimates (EM&V Issue Group #3) had the highest priority across all three energy futures. Interestingly, resolution of the problem of how to best integrate estimates of net program impacts into forecasts of future energy and peak demand was scored very high in the regional policy and transmission organization driver future, but low in the other energy futures. Conversely, efforts to increase the use of common Program Savings terms (EM&V Issue Group #1) was ranked as a top 3 priority in the business as usual and national policy driver future but relatively low priority in the regional policy driver future.

Conclusions

Most respondents agreed that it will be important to improve the consistent use of program savings terms and more importantly to be clear whether program administrators are reporting gross or net program savings in their summary reports. Given the wide range of methods and different definitions of net program savings found in our survey, addressing and reconciling these differences should be a high priority for national organizations seeking to compare savings reports from different states. Both of these issue areas represent a significant opportunity for state or regional organizations to collaborate or cooperate in efforts to increase consistency in both methods and savings terminology

Looking to the future, it will be important to ensure that the assumptions used to estimate baseline energy usage in program impact evaluations are consistent with the assumptions used in load and peak demand forecasts of utilities and Independent System Operators. This issue is important in order to increase confidence among resource and transmission system planners and policymakers in energy efficiency as a long-term resource and to ensure that estimates of future electric system resource needs are consistent and do not over or under-estimate the impacts of energy efficiency programs on future load growth. .

References

Barbose, Galen, Goldman Charles, and Schlegel, Jeff. 2009, The Shifting Landscape of Ratepayer-Funded Energy Efficiency in the U.S. October:(LBNL- Environmental Energy Technologies Division)

Schiller, Steven R. 2009, Model Energy Efficiency Evaluation Guide (Prepared for the National Action Plan on Energy Efficiency)