M&V 2.0: hype vs. reality

Jason Kupser
DTE Energy
Detroit, Michigan
USA
www.dteenergy.com

Sophia Francois
DTE Energy
Detroit, Michigan
USA

Joshua Rego
Navigant Consulting
Ann Arbor, Michigan
USA
www.navigant.com

Peter Steele-Mosey
Navigant Consulting
Toronto, Ontario
Canada

Toben Galvin
Navigant Consulting
Burlington, Vermont
USA

Craig MacDonald
Navigant Consulting
Boulder, Colorado
USA

Kathleen Gaffney
Navigant Consulting
London
United Kingdom

Keywords

measurement and verification, econometric evaluation, non-compliance, data analytics

Abstract

The rollout of advanced metering infrastructure has created an opportunity to derive additional value from energy efficiency programs by obtaining more timely and more granular estimated impacts than those made available through traditional measurement & verification (M&V) approaches. This paper presents two different, but related, approaches to leveraging residential high frequency energy consumption data for program management and evaluation: custom econometric analysis and third-party software tools (commonly referred to as M&V 2.0). These approaches are still relatively new, and the question remains for many utilities: how much of M&V 2.0 is hype, and how much is real?

In the fall of 2014, DTE Energy set out to better understand the true potential of emerging M&V approaches and to determine the relative benefit of these approaches in a deemed savings environment. To accomplish this, the company chose to pilot test both the econometric and software-based M&V 2.0 approaches to determine which represents the most effective method for performing M&V.

This paper presents the findings of this evaluation to-date, beginning with the market research performed to determine the potential options for packaged "M&V 2.0" type software tools capable of evaluating residential energy efficiency programs. It goes on to discuss the structure of this evaluation, including the process of establishing program goals, identifying a testbed energy efficiency program, and selecting a vendor. Last, the paper highlights lessons learned along the way, in the hopes that other utilities interested in testing M&V 2.0 approaches will find this information beneficial. Results from comparing traditional and M&V 2.0 methods are forthcoming, and may be shared at the time this paper is presented.

Introduction

The ability to quantify the impacts of energy efficiency efforts without the need for survey-based estimation or appliance-specific metering is considered by many to be the future of evaluation, measurement, and verification (EM&V). Recent expansions in the nationwide deployment of advanced metering infrastructure (AMI) have placed this goal possibly within reach for the first time by automatically capturing and storing interval electric consumption data at a level far more granular (hourly or sub-hourly) than ever before. In the last decade, there has been a significant increase in the number of hardware and software tools capable of leveraging high frequency customer electricity demand and consumption data. These tools offer a range of abilities, from customer intelligence, to end-use load disaggregation.[1]

One area of increasing discussion in the world of interval data analytics is advanced measurement and verification, often referred to as "M&V 2.0". New tools and approaches offer the ability to leverage existing data from AMI infrastructure to estimate the impacts of residential and commercial energy efficiency programs. As these tools have become more preva-

1. Todd, Anniuka and Hult, Erin. 'Big Data: New Opportunities for EM&V of Energy Efficiency Programs'. 2014. Lawrence Berkeley National Laboratory.

lent, an increasing number of utilities are asking whether or not these data could be leveraged to enhance the effectiveness of energy efficiency programs by increasing the timeliness and accuracy of impact estimates while reducing the cost of impact evaluation. Currently, two different, but related, approaches exist to analyze AMI data for the purposes of M&V:

- **Software Tools**: Off-the-shelf or packaged software tools capable of remote analysis of high frequency energy consumption data
- **Econometric Analysis**: Custom, user-performed statistical analysis of high-frequency consumption data

Even in a deemed savings regulatory environment, results from these advanced analysis tools and methodologies still represent significant value to utilities in three ways:

- **Real-Time Performance**: Insight into the potentially more precise performance of energy efficiency programs in near real-time can provide positive value to the company's system planning, program operation, and evaluation efforts.
- **Customer Engagement**: The increased granularity of data allows the utility to better tune its programs. To date, most applications of these data analytics approaches have been to better target customers for programs and to improve program management (e.g. more detailed and real-time tracking of trade ally involvement and savings per job site).
- **Resource Planning**: Understanding net impacts, as seen by the grid, can help the utility understand the real need for energy, distribution, and transmission capacity.

At the same time discussion of M&V 2.0 possibilities is on the rise, there is a growing body of evidence suggesting that these tools can, within a reasonable degree of accuracy, determine energy savings across a portfolio of structures. New research suggests that when the date of program participation is known and the sample size is large enough (40+), automated M&V approaches using whole-building modeling can be used to predict the energy use in a group of commercial buildings within a 5–8 % margin of error. [2,3] Use of automated M&V tools to evaluate program-level savings in residential structures has not yet been tested, but commercial results seem to indicate that this is possible.

Despite this, there is little evidence of these approaches actually being applied at a portfolio level to determine residential program impact in the real world. Therefore, the question still remains as to whether or not the capabilities offered by these emerging tools live up to the ever-increasing hype.

BACKGROUND

In the fall of 2014, DTE Energy set out to answer this question by determining whether or not M&V 2.0 tools are sufficiently flexible, scalable, and robust enough for use in residential energy efficiency program impact evaluation. The project was broken down into two phases, outlined in Table 1.

PHASE 1: MARKET SCAN

The goal of Phase 1 was to investigate whether or not commercialized software capable of analyzing residential interval consumption data exists, and if so, how the results of impact evaluation compare to those using a deemed savings approach. Working with Navigant Consulting, the company reviewed existing software tools to determine:

1. **Suitability**: What are the high-frequency energy consumption software tools that currently exist, what are their capabilities, and which are of greatest interest for residential energy efficiency program evaluation?
2. **Current Use**: What jurisdictions are currently using these software packages, what are they using them for, and have any used them for program evaluation?
3. **Results**: Of the jurisdictions using these software packages, what have their experiences been like in terms of calculated savings (compared to previously used estimates), cost, accuracy, scalability, and ease of use?
4. **Value Add**: How else beyond improved evaluation outcomes, could these tools deliver value to utilities and their customers?

To answer these questions, the team reviewed web resources, press releases, journal articles, industry publications, and research papers. After conducting interviews with industry representatives, utility personnel, and academic experts, over 20 companies were evaluated on five criteria to determine which tools most closely align with DTE Energy's M&V 2.0 goals:

- **Platform**: software tool is remote, and does not require hardware (sub-metering) or installed software to perform the desired analysis
- **Customers**: software tool is currently being used by, or marketed to investor-owned utilities
- **Target**: software tool is capable of evaluating energy use in residential structures
- **M&V Capabilities**: Software tool is capable of performing remote measurement & verification
- **Frequency**: Software tool is able to utilize interval (hourly or sub-hourly) consumption data

The Phase 1 market scan revealed that few existing software packages satisfy all five criteria simultaneously. At the time this research was conducted (December 2014), the majority of the identified tools (18 of 21) were capable of analyzing AMI data. The high number of tools able to analyze these data is not surprising, as this was the first criteria used to determine potential software tools.

Of the 18 software tools able to analyze AMI data, 12 were identified as having existing utility clients, or as being targeted to utilities. The market scan also found that the majority of these software packages focus on enhancing the energy savings from efficiency programs, rather than evaluating energy efficiency programs. This is done in three ways: enhanced engage-

2. Granderson et al., "Evaluation of M&V Accuracy and Savings Uncertainty", 2014, US Department of Energy.

3. Granderson, Jessica et. Al. 'Assessment of Automated Measurement and Verification (M&V) Methods', 2015, Ernest Orlando Lawrence Berkeley National Laboratory.

Table 1. DTE Energy M&V 2.0 Project Phases.

Phase 1: Market Scan	Phase 2: Pilot Test
• Literature review of sources describing use of M&V 2.0 software tools • Catalog jurisdictions where software are used, and describe any outcomes • Based on these findings, recommend software for testing	• Contingent on Phase 1 results, test both software and econometric M&V 2.0 approaches • Compare results, accuracy, methodology, and cost between M&V approaches

ment, advanced targeting of participants, and program intelligence. First, enhanced customer engagement allows utilities to provide customers with recommendations and tips specific to their consumption patterns. Second, advanced targeting allows program managers to identify high-value potential participants to engage. Last, program intelligence provides utility staff with deeper insight into top performing program characteristics like trade allies and energy saving measures. While undoubtedly valuable to program managers, these services do not satisfy DTE Energy's goal of providing timely and accurate M&V.

The team also found that current AMI data analysis software is focused on the commercial sector. In fact, six of the 12 tools capable of using AMI data and marketed to utilities were only capable of analyzing energy use in commercial structures. Three were capable of analyzing energy use in both residential and commercial buildings, and three of the identified software tools were capable of analyzing only residential consumption data.

Finally, the team identified few tools capable of performing remote M&V without the need for third-party hardware (sub-metering) or locally installed software. Of the 21 tools reviewed, only 8 were capable of remote M&V, and only one was capable of performing remote M&V in residential structures. Table 2[4] contains a full list of the reviewed software tools:

Based on the results of the Phase 1 market scan, the team identified four key takeaways about the market for software tools capable of analyzing energy consumption data:

1. **Focus on Commercial:** Existing software is focused on the commercial sector. This is likely due to the larger potential energy savings from per-customer recruitment in this sector.
2. **Focus on Energy Savings**: Tools are focused on achieving, rather than evaluating, savings from energy efficiency programs. This is likely due to the quantifiable value associated with creating energy savings, as opposed to the more timely and accurate measuring of said savings.
3. **Little Use by Utilities**: The team found no evidence of utilities using software to replace traditional impact evaluation. This is likely due to the fact that software capable of fully replacing traditional impact evaluation does not yet exist. If utilities are testing the M&V capabilities of these or similar software tools, it is likely that these projects are still in pilot phase, and results have not yet been released.

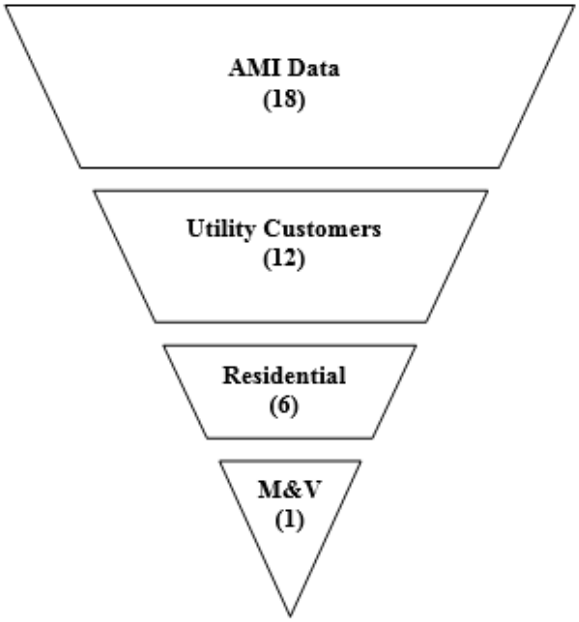

Figure 1. Narrowing of Software Capabilities.

4. **Rapidly Changing Market**: The market for these software tools is changing rapidly, both in terms of the capabilities, and the number of firms offering products in this space.

Research performed in Phase 1 also revealed that while there is a great deal of discussion surrounding M&V 2.0, there does not appear to be a robust demand for this service yet. Utility interest undoubtedly exists, but many operate in a deemed savings environment where program savings are determined by per-measure saving values set in a regulator-sanctioned technical reference manual (TRM), or by one-off evaluation of custom projects. Switching to a different method of impact evaluation has the potential to introduce an additional amount of uncertainty into the evaluation process, which may seem like a risk not worth taking to many utilities. In addition, it is unclear how these tools will be accepted by regulators. Across the country, regulators require a third-party evaluator to serve as an independent check on reported program savings. Many jurisdictions have not yet determined whether or not third party software tools, designed and managed by a vendor but used by the utility themselves, could fill this role. Until this is determined, utilities will still be required to retain an independent evaluator, and M&V 2.0 results, while potentially quicker

4. Kramer, Hannah et. Al. 'Inventory of Commercial Energy Management and Information Systems (EMIS) for M&V Applications'. 2013. Portland Energy Conservation, Inc. http://www.eeperformance.org/uploads/8/6/5/0/8650231/inventory_of_mv_applications.pdf.

Table 2. Software Tools Reviewed as Part of Phase 1 Market Scan.

Software	Platform	Utility Customers	Target	M&V	Frequency
EnergySavvy	Remote	Utilities (20+)	Both	Y	Sub-hourly
Opower	Remote	Utilities (95+)	Residential	N	Sub-hourly
FirstFuel	Remote	Utilities (2+)	Commercial	Y	Sub-hourly
Retroficiency	Remote	Utilities (15+)	Commercial	Y	Sub-hourly
PlotWatt	Remote	Utilities (2+ pilots)	Both	N	Sub-hourly
Bidgely	Remote	Utilities (2+ pilots)	Residential	N	Sub-hourly
Ecova	Remote	Utilities (10+)	Both	N	Sub-hourly
Apogee	Remote	Utilities (650+)	Both	N	Monthly
C3	Remote	Utilities (17+)	Commercial	N	Sub-hourly
BuildingIQ	Remote	Utilities (1+)	Commercial	Y	Sub-hourly
Eeme	Remote	Unknown	Residential	N	Sub-hourly
Gridium	Remote	Non-Utility	Commercial	Y	Sub-hourly
Energy Ai	Remote	Unknown	Commercial	N	Sub-hourly
EnerNOC	Remote	Utilities (25+)	Commercial	Y	Sub-hourly
UT3	Software	Utilities (1)	Commercial	Y	Sub-hourly
kWIQly	Remote	Unknown	Commercial	N	Sub-hourly
Wegowise	Remote	Unknown	Commercial	Y	Sub-hourly
Aclara	Remote	Unknown	Commercial	N	Sub-hourly
Lucid	Remote	Non-Utility	Commercial	N	Sub-hourly
Autodesk	Remote	Unknown	Commercial	N	N/A
Portfolio Manager	Remote	Unknown	Commercial	N	N/A

Source: Navigant market scan, December 2014.

and less costly, may seem redundant. In addition, there is still uncertainly around whether or not regulators will accept the methodology used by M&V 2.0 approaches.

Following the Phase 1 analysis, several questions still remained unanswered. First, the team was unable to determine the cost associated with various software tools. Most software providers appear to tailor pricing to each utility based on the level of customization required, projected energy savings, and the number of services (customer engagement, targeting, or verification) that are being requested. Smaller firms may offer fixed pricing models based on the number of data points (meters) analyzed by the software. Additionally, software accuracy was reported to vary between specific platforms, and little research exists into the accuracy of these tools when applied, at the program-level, to residential structures. Despite these outstanding questions, DTE Energy made the determination that enough opportunity exists in this market to move forward with Phase 2.

PHASE 2: PILOT STUDY

While Phase 1 focused on understanding the current uses and capabilities of M&V 2.0 software, Phase 2 focuses on comparing the methodology, results, accuracy, and cost between these approaches. It should be noted that work on Phase 2 is still ongoing. To determine this, DTE Energy set out to pilot test both the packaged software and econometric approaches to M&V 2.0, with the primary goal of evaluating the potential value of a software tool to perform real-time M&V. The plan is to use results from the parallel econometric analysis as a baseline by which to assess the accuracy of the software tool's calculated results. The team identified four specific goals to ensure the greatest chance for success, outlined in Table 3:

Next, the team determined the programs with which to test the two M&V 2.0 approaches. For M&V 2.0 analysis, data are required from all participants for a 12-month period prior to program participation to allow for accurate participant matching. This meant choosing a larger and more mature program to ensure a higher number of participants with adequate historical consumption data. Further, the number of desired participants is inversely correlated to the relative size of the anticipated per-measure impact, and selecting a program with a large number of participants would proactively address any potential concerns around the number of participants, or per participant savings. This led DTE Energy to select the Residential Heating, Ventilation, and Air Conditioning (HVAC) Program and Appliance Recycling (AR) Programs with which to pilot M&V 2.0 approaches. Both programs were selected due to their maturity, participation, and associated savings, as outlined in Table 4.

In fall 2015, DTE Energy's implementation contractor for the Appliance Recycling Program (ARP) stated it was unable

Table 3. M&V 2.0 Phase 2 Goals and Metrics.

Goal	Description	Metric
Validate Evaluation Work	The extent to which M&V 2.0 approaches provide more granular and more accurate savings estimates	• Program-level savings estimates • Measure-level savings estimates • Accuracy of calculations
Cost and Scalability	The cost, time, and scalability of M&V 2.0 approaches	• Ability to scale to portfolio level • Cost of using analysis method • Cost/time to add other programs
Use in Resource Planning	Ability and accuracy of M&V 2.0 approaches to calculate demand impacts	• Ability to determine peak impacts • Accuracy of peak impact estimates
Better Project Management	Other benefits of M&V 2.0 approaches, and general staff impression of approaches	• Staff feedback on approach/tool

Table 4. Selected Testbed Programs for M&V 2.0 Evaluation.

Residential HVAC	Appliance Recycling
• Heating/cooling measures represent significant time-sensitive savings, maximizing the value of M&V 2.0 analysis • Heating/cooling measures account for significant volume of savings, addressing potential data volume concerns • The large number of program participants' addresses potential data volume and scalability questions	• Large number of participants and volume of savings addresses potential data volume, impact granularity, and scalability questions • ARP represents a significant amount of peak demand savings due its scale • Given the lack of variable operating conditions for ARP measures, traditional and M&V 2.0 savings estimates would likely be similar, creating an ideal basis for comparison

Table 5. Vendor selection criteria.

Required Capabilities	Desired Capabilities
• Capable of determining energy savings impact • Capable of determining peak demand impact • Capable of analyzing AMI data • Web-accessible and easily-customizable dashboard • Generates easily-customizable and on-demand reports • Supports multiple program types • Scalable to include whole portfolio of energy efficiency programs • Capable of determining historical program savings	• Ability to analyze gas consumption • Ability to track performance of program elements including trade allies and geographical regions • Ability to track workflow stages • Capable of being customized by DTE staff, or with minimal input from vendor • Minimal required training for DTE staff

to continue implementing program, and as a result, the program was temporarily suspended until a new contractor could be identified. The implications of this are discussed in greater detail in subsequent sections of this report, however the immediate impact on Phase 2 work was to preclude the ARP from being selected as a testbed program.

Due to this lingering uncertainty on price and software capabilities, the team chose to solicit feedback from qualified vendors, rather than selecting a vendor outright. Releasing an RFP would address questions associated with cost and scalability upfront, and would help determine any unadvertised or new capabilities offered by the field of software vendors. Preliminary research in Phase 1 identified vendors who were contacted directly and invited to submit responses. The request for proposals and associated scope of work laid out a list of required and desired capabilities for a software tool, outlined in Table 5.

By the time potential suppliers were required to indicate their intent to submit a proposal, only one vendor had submitted their intent to bid for Phase 2 work. This further confirmed the results of the team's Phase 1 findings, that few vendors are offering software tools capable of performing remote M&V 2.0 for residential energy efficiency programs. The submitted proposal was reviewed and found to satisfy all required criteria laid out in the request. As of March 2016, this contract is in the final stages of negotiation, and the software is currently expected to launch as part of Phase 2 work in early Q2 2016.

The success of the Phase 2 work necessitated access to these data, and therefore beginning the process of gaining access early was critical should any complications arise. The team determined the criteria for the kinds of data that would be required to evaluate both testbed programs as including:

- Data for all residential customers with single meter, billed at standard residential electric rate
- Non-overlapping hourly usage data of Electric AMI meters at site level for identified customers

Data to be excluded from the extract were defined as:

- Data for customers with agreements effective and terminated on the same day
- Data for customers who have an account terminated
- Data for customers with landlord agreements
- Data for customers who do not have a responsible person last name on record
- Data for non-DTE Energy customers

Given the large volume and secure nature of the data, the team needed to create a plan to accept, store, and decrypt large volumes of data on an ongoing basis. Rather than start from scratch, the team decided to use the systems already put in place for DTE Energy's Behavioral Demand Response pilot program. The program's implementation contractor, OPower, was already receiving a similar AMI data feed, and the project team was able to utilize this interface. As it turns out, the need to develop a new interface or data access pathway would have been cost and time prohibitive, and would have prevented the project from moving forward.

In parallel to the evaluation and launch of an M&V 2.0 software tool, the team planned on conducting an econometric analysis of the selected testbed programs. This analysis began in early 2016, is still ongoing, and preliminary results may be shared when this paper is presented. The purpose of conducting an econometric analysis alongside the software analysis was two-fold. First, it would allow the approach to be tested relative to a deemed savings approach based on cost and calculated savings. Second, it would provide a baseline to assess accuracy of the selected software tool. Validation of the software tool has yet to begin, and will be performed based on a comparison of estimated impacts using quasi-experimental methods supported by AMI data. The estimated impacts that will be compared with packaged software outputs include:

- **Peak Demand Impacts**: The system-coincident ex-post on peak demand impact of the Residential HVAC program (average kW per participant)
- **Energy Use Impacts**: The ex-post energy use of the Residential HVAC program (average kWh per participant)
- **Accuracy**: Demand and energy use impacts compared between both M&V 2.0 methods
- **Timing**: Frequency and speed with which econometric and M&V 2.0 analyses can provide impact information
- **Cost**: Relative costs associated with econometric and M&V 2.0 software analyses

The econometric analysis conducted by the project team will seek to estimate both ex-post and ex-ante program impacts of the selected testbed programs. "Ex-post" impacts refer to the actual program impacts that have occurred in the historical period, or the energy conservation and peak demand impacts that accrued to the program in the previous program year. "Ex-ante" impacts refer to a forecast of program savings under some pre-determined scenario, such as the peak demand impact the program could expect to deliver in 2015 given "normal" weather and forecast enrolment. For ex-ante impacts, the team plans to rely on "normal" and "extreme" weather assumptions. Presently, the team is in the process of "matching" to estimate savings. "Matching" refers to the development of a non-participant comparison group with similar characteristics to the participant groups including consumption, home size, and other variables known to have a high correlation with post-program energy use. The team's approach will be to match on historic usage. In particular, the matched control customer for each participant will be selected to minimize the Euclidean distance between the participant's and potential control customer's electricity use during a pre-program period.[5] To estimate peak demand savings, the team will match on pre-program peak demand hours; to estimate energy savings, the team will match on pre-program average daily usage.

Next steps & conclusions

There is an increasing amount of discussion around M&V 2.0, and emerging research suggests that these tools can be used to accurately determine energy savings across a portfolio of buildings. This area of research is still relatively new, and analysis of M&V 2.0 tools capable of evaluating energy use in residential structures is yet to be performed. Despite this increase in attention, the market for M&V 2.0 is still in its relative infancy. Most of the tools capable of analyzing AMI data are geared towards commercial building analysis, and towards achieving, rather than evaluating, energy savings. Few off-the-shelf tools are capable of remote M&V, and the team found no evidence of utilities using these methods to determine program impact in place of traditional impact evaluation. This is likely due to the fact that the impact of these evaluation techniques on program savings, along with general regulatory acceptance of these methods, is still unknown.

Despite this, DTE Energy is determined to continue moving forward with answering questions around the impact, scalability, timeliness, and cost of these methods, starting with determining the savings arrived at through M&V 2.0 techniques. Econometric analysis is underway, and modeling estimated program savings and peak demand impacts for the 2015 program year will begin shortly. Additionally, the software tool still remains to be deployed in order to evaluate savings, accuracy, and cost of software-based M&V 2.0. Preliminary savings estimates from these two analysis methods will likely be available when this paper is presented in August 2016.

Once the software tool has been deployed and the first round of the econometric analysis is complete, the project team will need to compare results between both platforms, and against deemed savings estimates. Results from the econometric analysis and those delivered through the software tool will be compared to gauge the accuracy of the se-

5. Ho, Daniel E., et. Al. 'Matching As Nonparametric Preprocessing For Reducing Nodel Dependence In Parametric Causal Inference'. 2007. Political Analysis.

lected software tool. Savings and peak demand estimates from the econometric analysis and software tool will be compared to those calculated through methods currently used by DTE Energy to determine the potential impact on program savings. Finally, the costs associated with each method will need to be compared to determine which approaches are most cost-effective.

There remains a great deal of discussion around M&V 2.0, and for good reason. The increased granularity, frequency, and access provided by advanced metering data and analytic techniques are likely to have long term impact on how evaluation is performed. The question that still remains is not whether or not this will lead to a sea change in how evaluation works, but whether or not that time is now. In the event that the software is found to be accurate, cost-effective, and a useful tool for impact evaluation, DTE Energy plans to explore the expansion the use of the software and/or econometric analysis to multiple programs. Ultimately, DTE Energy hopes that the process explained here and the lessons learned will be of use to other utilities interested in pursuing M&V 2.0.

References

Granderson, Jessica et. Al. 'Assessment of Automated Measurement and Verification (M&V) Methods'. 2015. Ernest Orlando Lawrence Berkeley National Laboratory, http://eetd.lbl.gov/sites/all/files/lbnl-187225_0.pdf.

Granderson, Jessica et. Al. 'Evaluation of M&V Accuracy and Savings Uncertainty', 2014, United States Department of Energy.

Ho, Daniel E., et. Al. 'Matching As Nonparametric Preprocessing For Reducing Nodel Dependence In Parametric Causal Inference'. 2007. Political Analysis. http://gking.harvard.edu/files/matchp.pdf.

Kramer, Hannah et. Al. 'Inventory of Commercial Energy Management and Information Systems (EMIS) for M&V Applications'. 2013. Portland Energy Conservation, Inc. http://www.eeperformance.org/uploads/8/6/5/0/8650231/inventory_of_mv_applications.pdf.

Todd, Anniuka and Hult, Erin. 'Big Data: New Opportunities for EM&V of Energy Efficiency Programs'. 2014. Lawrence Berkeley National Laboratory.