Towards Demand Response Measurement and Verification Standards

David Holmberg, NIST
Gaithersburg, MD
david.holmberg@nist.gov

Dave Hardin, EnerNOC
Boston, MA
dhardin@enernoc.com

Ed Koch, Akuacom/Honeywell
San Rafael, CA
ed@akuacom.com

Keywords: demand response, M&V

Abstract

The value and impact of demand response (DR) at grid-scale depends upon the widespread and active participation of a diverse customer base. The level of participation by customers is impacted by financial compensation and return-on-investment. The role of demand response measurement and verification (M&V) is to determine the quantity of energy or power that is "delivered" by a DR resource under the conditions imposed by a DR program. The value and impact of demand response depends upon M&V standards that are fair, simple and accurate.

This paper reviews the state of demand response M&V standards as they relate to DR programs, summarizes technical issues, and offers recommendations for addressing these issues in an effort to strengthen the DR market.

1. DEMAND RESPONSE AND BASELINES

The timely reduction of customer energy demand has proven to be a valuable complementary resource to traditional generation as a means of balancing the supply and demand of electrical energy on the grid. This is evidenced by the growing number and variety of demand response markets and programs, FERC support for demand response [1], and the growth of automated DR (AutoDR). However measuring it effectively and efficiently remains a challenge.

Demand response M&V refers to the capability to quantify and validate that a residential, commercial or industrial customer has reduced or increased load in response to a DR signal while reducing the opportunities for gaming the system.

Well-defined M&V standards that are fair, simple and accurate will enhance interoperability at the organizational and business levels, as defined in the GridWise Architecture Council Interoperability Framework[2] [3], and facilitate market growth of demand response through clearly defined benefits and the consistent application of M&V techniques matched to the characteristics of customers, markets and programs.

DR markets and programs span Independent System Operators, Regional Transmission Organizations (ISO/RTOs) and balancing authorities at the wholesale level, and utilities and energy service providers at the retail level. Through DR programs, a wide variety of customers (including residential, commercial, institutional and industrial) provide energy and power resources to the grid.

Demand response can be either pay-for-performance programs or pay-for-participation. Compensation for these approaches is tied to the measurement and verification that the customer satisfied the conditions as defined by the program.

In the pay-for-participation case, verification consists of determining that the curtailment signal was sent, received and acted upon by the customer. This is often the case for direct load control (DLC) programs which toggle power to high-load devices such as air conditioners.

Pay-for-performance programs in energy, capacity, reserve and regulation markets each have specific performance evaluation criteria and associated methodologies [4]. These programs range in time from commitments made years ahead down to real-time response within 4 seconds.

In cases where pay-for-performance is measured by comparison to an absolute value, accurate measurement is essential and verification is straightforward. In cases where performance is measured relative to a baseline, both the definition of the baseline and energy measurement are critical. Baselines are estimates of the energy that would have been consumed during a DR event if the event had not occurred. The challenge is to obtain a simple but accurate estimate of a customer's energy usage reductions relative to a baseline during a specific time interval (i.e., the DR deployment period) that is fair to all parties.

As estimates, baselines are inherently imperfect. However, good baselines balance three main attributes: 1) accuracy – giving customers credit for the curtailment achieved; 2) integrity – harmonizing incentives for all parties and making it difficult to "game" the system; 3) simplicity – performance calculations should be easily understandable by all stakeholders.

2. DR M&V BACKGROUND

NAESB [4] identifies four wholesale demand response programs or services; 1) Energy, 2) Capacity, 3) Reserve and 4) Regulation. Energy programs require resources to "deliver a quantity of electricity measured in MWh"; Capacity programs require that resources help balance the grid by controlling electrical demand "over a defined period of time measured in MW"; Reserve programs require that resources be available for deployment by the balancing authority based on reserve capacity requirements; and Regulation programs require that resources "increase and decrease load in response to real-time signals" from the balancing authority.

Within each of these program types, several methods exist for evaluating the performance of customer resources: 1) Maximum Base Load, 2) Meter Before/Meter After, 3) Baseline Type I, 4) Baseline Type II and 5) Metering Generator Output. (Note: The exception is the regulation program where only Meter Before/ Meter After or Metering Generator Output applies.)

Maximum Base Load refers to the ability of a resource to operate at an electrical load level at or below a specified level. Meter Before/Meter After refers to performance measured against a baseline defined by meter readings prior to deployment and similar readings during the sustained response period [4]. Baseline Type I refers to a baseline created by using historical interval meter, weather and/or calendar data and Baseline Type II uses statistical sampling to generate a baseline.(See reference [5] for detailed definitions.)

Of these methods, Maximum Base Load, Meter Before/Meter After, Baseline Type I, and Metering Generator Output rely primarily upon obtaining accurate time-stamped historical energy data with potentially straightforward calculations based on the meter data, if the baselines are well-designed.

Maximum Base Load (MBL) is a static technique that utilizes data, often from the previous year, to draw a line at a certain power level below which the customer must maintain demand when called upon. This demand level is often non-representative of current load conditions due to changes within the customer's facility. This technique often bases the MBL on previous year peaks either coincident or non-coincident with system peaks) [7].

Meter Before/ Meter After is usually used only for fast-response programs and reflects actual load changes in real-time, reading the meter before and after response to measure the change in demand.

Metering Generator Output is applicable to behind-the-meter onsite generation and determines the demand reduction based on generator output data, assuming that all load taken served by the generator would otherwise have been on the system.

Type I baselines are widely used and include techniques such as rolling averages, comparable day values, and period averages. Characteristics of Type I baselines include:

1) Baseline shape based on average load profile,
2) Use of meter data from each individual site,
3) Use of historical meter data from days immediately preceding the event,
4) Use of other data to adjust the baseline such as weather and calendar data.

Rolling averages usually use historical meter data weighted towards more recent data and depend upon having sufficient data to reflect representative conditions.

Comparable day methods identify a representative day in the past but these methods suffer from: 1) a lack of objective criteria for selecting a specific day and 2) they rely upon after-the-fact identification.

Period averaging methods create baselines by averaging historical energy data to estimate load for specific time intervals that are "representative" of the load. These are also called High/Mid X of Y baselines where Y is the number of most recent days with X of those days having the highest load for High X of Y baselines or middle load for Mid X of Y baselines. As examples, PJM has used High 4 of 5 baselines, Ontario has used High 15 of 20 and California has used High 10 of 10.

High X of Y methods are examined and recommended for NAESB Type I baselines with adjustments in the EnerNOC "Demand Response Baseline" White Paper [6] and the KEMA "PJM Empirical Analysis of Demand Response Baseline Methods" [7] as the best baseline method in most cases. The selection of the number of days in X of Y baselines depends upon many factors[1] including, but not limited to, the following: [6]

1) Look-back Window – The range of days prior to the event that are considered (i.e. the value Y).
2) Exclusion rules – Some days are excluded from consideration such as holidays, previous DR event days, weekends, thresholds and scheduled shutdowns (as these are not representative of "normal" operation).
3) Ratio of X to Y – The selected subset of X days in the range of Y days relates to the characteristics of the DR program and the customer's general energy usage patterns.
4) Time intervals – More frequent data capture provides greater detail about load behaviors.
5) Baseline adjustments – Adjustments are based on day-of-event load conditions to improve baseline accuracy. Adjustments may also be made based upon weather, calendar days, etc.

[1] Many of these factors are also used in other baseline calculation methods.

6) Adjustment Duration - If the time period associated with the adjustment is either too short or too long, it may not be representative.
7) Multiplicative vs. Additive Adjustments – Multiplicative reflect percentage demand comparisons and additive reflect actual differences. Additive and multiplicative adjustments both use the difference between the baseline and observed load but the additive adjustment is constant across the entire event period while the multiplicative adjustment adjusts as a percentage of loads during the event period. This can produce an adjustment more appropriate for a load shape that changes during the event period.
8) Capped vs. Uncapped Adjustments – Limits to adjustments.
9) Symmetric vs. Asymmetric Adjustments – Symmetric adjustments can increase or decrease the baseline while asymmetric adjustments only allow adjustment in one direction.
10) Aggregation level – Calculations can be done at the facility level vs. at a portfolio level.

Type I baseline methods are widely used and can provide simple, accurate and fair M&V for DR programs [6]. At the same time, these methods can lead to inaccurate results if the Type I method variables are not carefully matched to DR program needs and customer energy usage characteristics. An example of this is the look-back window. If this window is too long, then out-of-date, non-representative data might be used in calculating the baseline. Over-estimating performance penalizes the system operator while underestimating performance penalizes the customer.

Type II baseline methods are often used in scenarios where aggregated meters are available but individual site meters are not. Aggregated historical meter data is used to create a baseline that is appropriately allocated to individual sites or loads that are not metered. This method is typically more appropriate for residential DR as commercial and industrial facilities can cost-effectively meter energy usage. Type II methods are often more complex, require more meter data and may not produce timely results leading to a lack of real-time visibility.

3. CHALLENGES AND RECOMMENDATIONS FOR DR M&V

Based on the foundation of work that has already been done examining DR M&V methods in [6] and [7], each of which have limitations and point to further work, and based on the analysis of ongoing and developing DR market needs, the authors identify some key challenges to advancing DR and then make recommendations for moving forward.

3.1. Flexible DR Framework

At the wholesale level, the NAESB WEQ-015 Business Practices for Wholesale Electricity Demand Response Programs [5] provides a valuable framework within which ISO/RTOs develop M&V requirements and specifications and implement specific programs. As a framework, it specifies the data attributes that apply to a method, but not the specifics of how the methodologies address energy provider or customer characteristics and how the different M&V approaches impact DR resource participation.

The NAESB framework is flexible and can be implemented in different ways in different wholesale markets. This results in localized market-specific M&V specifications which inhibit domestic and international market growth due to the regional customization and increased complexity required for implementing demand response solutions. In addition, the NAESB standard applies only to ISO/RTOs and does not apply to other balancing authorities.

At the retail level, M&V standards are still evolving as retail markets move toward pay-for-performance driven by smart meters.

Recommendation: These factors together indicate the need for best practice application guidelines that can lead to clear M&V practices in wholesale and retail markets for different DR program types.

3.2. Bi-Directional Energy Flows

Smart grid will require bi-directional energy flows with demand side resources (DSR). Resources can act as both loads and supply but historically wholesale markets have not permitted resources to both “take and make” power. Wholesale electrical markets and contract settlement are currently oriented toward generation assets. This is an inhibiting factor that prevents demand side resources from fully participating in demand response.

Recommendation: M&V methods should explicitly permit metered power flow to go negative when on-site generation is operational.

3.3. Increased Use of Energy Storage

It is generally accepted that facility energy storage (in all its forms) can and will play a key role as a demand side resource that can help in the integration of renewables. Storage resources are somewhat unique in that they can be used to affect load profiles in two different ways:

- The charging of the energy storage is itself a load that can be started and stopped just like any other demand side load.
- The energy storage can be discharged. In some cases the discharge of the energy storage is used to affect other loads such as HVAC, but in other cases

the discharge of the storage can inject electrical energy directly on to the grid.

One of the main attributes of energy storage is the concept of a "charge state" of the resource. The balancing authorities will need to take this attribute into account to most effectively dispatch these resources. Energy storage thus has a unique combination of both energy and power capacities and there will need to be new models for how their performance is measured and verified so they can be appropriately compensated.

ISOs, such as the CAISO, are in the early stages of deploying new resource models that try to accommodate storage, but more work needs to be done to come up with a consistent model that can be applied everywhere.

In addition, energy storage M&V is currently treated the same as generation in many open markets and therefore a stakeholder consensus does not exist that storage should be handled in the same manner as demand response.

Recommendation: Analyze the role of storage in demand response, develop a clear statement of the requirements for effective participation of storage in DR programs, and then, if appropriate, refine storage M&V methods that can be deployed more broadly.

3.4. Potential for Customer Participation in Multiple DR Programs

Increasing cardinality in the relationships between demand response programs and customer resources is resulting in customers participating in several concurrent or overlapping programs. Customers can participate in different programs based upon the energy characteristics of their resources. Successful participation requires decoupling the overlapping programs so that the resulting M&V methods accurately reflect individual program participation. A hypothetical example of this scenario might be a military base that wants to sell DR capacity to an ISO/RTO but is also on a local utility tariff that is dependent upon limiting peak demand under specified conditions.

Recommendation: Develop best practice guidelines that define the dependencies and interactions between programs.

3.5. M&V Quality Measurements

M&V techniques do not typically incorporate quality of response measurements such as monitoring how well a resource follows the specifications unless such qualities are explicitly included in the relevant product definition. Having more refined M&V techniques that incorporate the quality of response may allow generation and demand side resources to be better differentiated and compensated as opposed to treating everything like a generator as is typically done today. This, however, needs to be balanced with the additional overhead and cost required to accurately measure and analyze the quality data which may result in constrained participation in DR programs.

Recommendation: Evaluate M&V measures and guidelines for the quality of response that are comparable (though not identical) to generation.

3.6. Commissioning Costs

Many financial barriers-to-entry exist for demand response resources. Some of these costs are related to M&V requirements and include upfront capital costs for meters, telemetry and security. In some markets and programs there is a tight coupling between the metering required for M&V and the systems that are dispatching the resources in the DR program. This can require that the DR dispatching entity be in direct communication with the customer loads in exactly the same manner for all customers. Given the range of loads that could potentially participate in a DR program (from generators to HVAC systems) it may not be cost effective to utilize the same communication infrastructure for all loads. As an example, the requirement for real-time telemetry adds significant costs when standard open communication protocols (e.g. OpenADR 2.0 Profile B) are capable of meeting performance requirements (i.e. communication throughput and latency).

Recommendation: Develop best practice guidelines for the communication infrastructure for integration of DR resources.

3.7. Difficulty of Calculating Customer Performance

Benefits are difficult to clarify due to complex tariffs and DR program M&V techniques which make customer performance calculations difficult in real-time. M&V methods should enable the straight-forward calculation of customer benefits in real-time.

Recommendation: Eliminate M&V methods for which the customer cannot know in real-time whether performance is in compliance. Analyze remaining methods in terms of cost to customer (up front and operational).

3.8. Customer Dynamic Loads

Customer load curves are often dynamic and change unpredictability.

Recommendation: M&V techniques should be refined to accommodate a wide range of customer load variability. Best practice guidelines would be beneficial.

3.9. M&V for Sub-Loads

Some M&V techniques require meter readings on an asset basis but must be able to take into account both aggregations of loads in general as well as sub-loads within a facility. It is problematic to use a whole-facility meter to perform M&V on a specific sub-load. Directly metering sub-loads increases accuracy but also increases costs. In some cases it may be possible to forgo sub-metering a load and use the existing meter on the whole facility, but only if changes in the sub-load's demand can be clearly seen within the baseline of the entire facility. Using the whole facility's meter has obvious cost benefits and adds flexibility to the mix of loads that may be used within the facility, but there currently are no standard guidelines for when this more cost effective methodology can and should be used.

Another related issue that needs to be addressed is the potential for reducing load on a sub-metered portion of the facility while subsequently increasing the load in another section of the facility resulting in zero net change to the load presented the grid.

Recommendation: Develop best practice guidelines for the M&V of DR sub-loads.

3.10. Encouraging Load Shifting

Customers are often not incentivized to dynamically shift load consumption due to M&V baseline calculations, or load shifting (e.g. for pre-cooling strategies) is discouraged because it is viewed as gaming the system.

Note that methods such as TOU (time-of-use) can aid in incentivizing this behavior as can programs such as ISO New England's dispatchable asset-related demand (DARD) which dispatches on price to either consume more or consume less.

Recommendation: Analyze M&V techniques and make recommendations for how to encourage load shifting within customer facilities while at the same time preventing customers from "gaming" baselines for advantage.

4. GENERAL RECOMMENDATIONS

The referenced reports [5][6][7] provide a good foundation but do not address several outstanding issues and do not provide clear guidelines for what M&V methods are best applied to specific types of DR programs along with information detailing how those methods should be tuned.

4.1 Recommendations for Research

1) Assimilate, catalog and identify gaps in existing public and private research (i.e. LBNL) related to M&V. One potential area for further research includes baseline estimation algorithms that could provide simple, consistent, fair and accurate baselines in real-time with reduced data requirements.

2) Pursue further research as needed to develop best practice guidelines for DR M&V, specifically to address issues highlighted in section 3 above: customer differences (including load variability), infrastructure consistency among methods, customer implementation costs, ability of customer to judge compliance in real-time, and how to encourage load shifting.

3) Further research topics identified above: effective integration of storage, methods for decoupling DR programs, measures of DR response quality, techniques for implementing DR for customers with variable load profiles, and methods for the accurate accounting of sub-load response.

4.2 Recommendations for Standards

1) Using the NAESB framework as a starting point, launch an effort within an appropriate standards coordination body such as the Smart Grid Interoperability Panel (SGIP) to define the roadmap and requirements for extending and refining M&V standards to address the issues and concerns outlined herein through a collaboration of both supply-side and demand-side stakeholders including FERC. This would enhance market growth by reducing M&V diversity and increasing simplicity, consistency and clarity in the application of M&V. This effort would make specific recommendations for further domestic and international standards work in coordination with the retail NAESB M&V standard. These recommendations would include residential, commercial and industry customers.

2) Develop application guidelines for DR M&V that detail best practices for the range of methodologies. This includes both the baseline techniques and the interface techniques used to collect the necessary M&V data. Emerging smart grid standards such as OpenADR could play a role in making these methodologies more cost effective and reach a much broader range of loads than is typically used today.

3) Demand response M&V standards are related to energy efficiency (EE) validation, estimation and editing (VEE) of energy data. It is recommended that standards activities related to DR M&V be coordinated with activities related to EE VEE so as to achieve consistency in methodologies and techniques.

4.3 Recommendations for Policy

1) In order for the demand response market to grow, DR should be measured and treated differently from generation based on inherent bi-directional and dynamic characteristics. It is recommended that additional analysis be undertaken to evaluate existing policy and recommendations be made for developing consistent policy nationwide.

Biographies

David Holmberg, NIST

David is a mechanical engineer in the Mechanical Systems and Controls Group, Energy and Environment Division, Engineering Laboratory at National Institute of Standards and Technology. He is co-chair of the OASIS Energy Interoperation Technical Committee, co-convener of the IEC PC-118 Smart Grid User Interface WG2 Power Demand Response, chair of the SGIP Building-to-Grid (B2G) Domain Expert Working Group and member of the NIST Smart Grid team supporting SGIP and standards coordination.

Dave Hardin, EnerNOC

Dave is Sr. Director of Smart Grid Standards at EnerNOC. He is active in a number of Smart Grid initiatives including the Smart Grid Interoperability Panel as vice-chair of the Architecture Committee and chair of the Industrial-to-Grid Domain Expert Working Group, the OpenADR Alliance Board of Directors, the OPC Foundation Technical Advisory Council and is a member emeritus of the U.S Department of Energy's GridWise Architecture Council.

Ed Koch, Honeywell

Ed is currently a Senior Fellow at Honeywell and CTO/Co-Founder of Akuacom, a wholly owned subsidiary of Honeywell. Ed was the leader of the workgroup at LBNL that drafted the OpenADR specification and currently sits on the Board of Directors of the OpenADR Alliance. Ed also is the co-chair of the OpenADR Task Force within UCAIug and is actively involved in a number of Smart Grid standardization efforts including the NIST Building to Grid Domain Expert Working Group, the OASIS Energy Interoperation Technical Committee, and the NAESB Smart Grid Standards Taskforce.

References

[1] *National Assessment & Action Plan on Demand Response*, U.S. Federal Energy Regulatory Commision, http://www.ferc.gov/industries/electric/indus-act/demand-response/dr-potential.asp

[2] *GridWise Interoperability Context-Setting Framework*, March 8, 2008, GridWise Architecture Council, http://www.gridwiseac.org/pdfs/interopframework_v1_1.pdf

[3] *NIST Framework and Roadmap for Smart Grid Interoperability Standards, Release 2.0*, Feb. 2012, http://www.nist.gov/smartgrid/upload/NIST_Framework_Release_2-0_corr.pdf

[4] NAESB MC11010, *Minor Correction Request, Retail Electric Quadrant, Correction to Diagram for Meter Before/Meter After Definition*, April 2011, NAESB, http://www.naesb.org/pdf4/req_mc11010_rec_040811.doc

[5] *NAESB report: Business Practices for Measurement and Verification of Wholesale Electricity Demand Response*, March 16, 2009

[6] *The Demand Response Baseline*, 2011, EnerNOC, Inc, http://www.enernoc.com/images/whitepapers/pdfs/demandresponsebaseline.pdf

[7] *PJM Empirical Analysis of Demand Response Baseline Methods*, April, 2011, KEMA, http://pjm.com/markets-and-operations/demand-response/~/media/markets-ops/dsr/pjm-analysis-of-dr-baseline-methods-full-report.ashx