Bring Me the Right Data!
An In-Depth Look at What Data Is Useful for the Utility of the Future

Isaac Wainstein, Ryan Bossis – Energy and Resource Solutions (ERS)

ABSTRACT

It is often recognized that gathering large and robust data sets is a critical component for the utility of the future to successfully integrate energy efficiency as a grid resource. However, collecting massive amounts of data without specific goals and associated methodologies will result in large, high-cost studies that are not necessarily useful for saving energy in buildings. The right data will enable utilities to implement effective Integrated Demand-Side Management (IDSM) and energy efficiency (EE) programs, forecast constrained networks with improved accuracy, and provide building performance measurement and verification (M&V) on a real-time basis. This paper answers the questions of what data is needed to accomplish these tasks, as well as how to cost-effectively gather and analyze that data. New York's Con Edison has committed to meeting more than 50 MW of capacity requirements using targeted energy efficiency for permanent demand reductions in a network experiencing rapid growth and change. The authors of this paper and Con Edison have embarked on a pioneering metering and market characterization effort to gather a comprehensive data set including characteristics for all energy-consuming equipment and the corresponding metered load shapes at the building, end-use, and equipment levels. This data collection effort spans 150-plus small businesses and 40 multifamily buildings covering shared spaces and more than 100 dwelling units. This paper will outline a data-driven approach designing data collection efforts that gather the right data for successful implementation and defense of IDSM and energy efficiency as a resource.

Introduction

Planning and delivering reliable customer-side demand reductions as a reliable grid resource requires a high resolution understanding of how customers use energy. Con Edison launched their nationally recognized Brooklyn Queens Demand Management (BQDM) program to reduce forecasted overloads of a substation serving parts of the New York City boroughs of Brooklyn and Queens. Current projections, as presented in Figure 1, illustrate that demand will exceed capacity in the 12-hour span from noon to midnight, with the greatest shortfall occurring at 9:00 p.m. in the targeted area. With a 12-hour peak, typical demand response programs will not be sufficient.

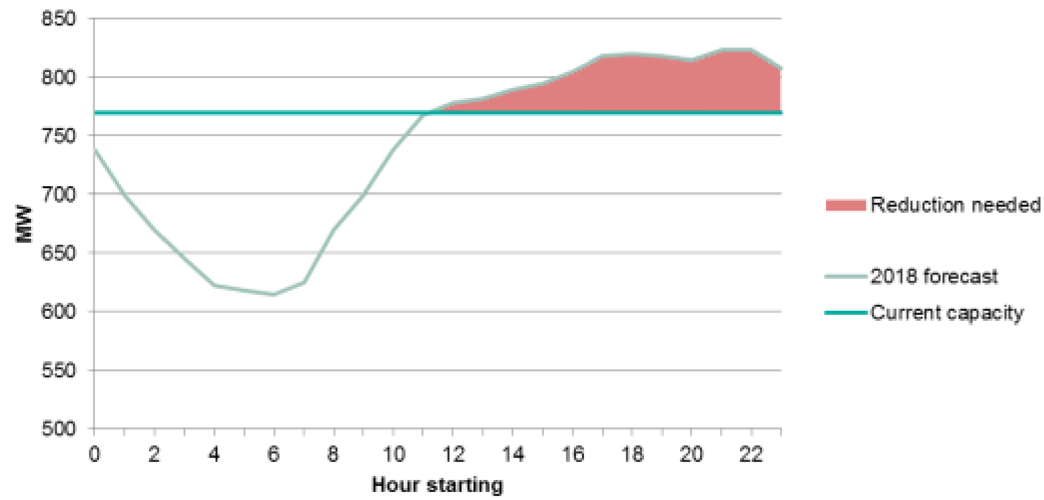

Figure 1. The BQDM challenge.

Gathering data about how customers use energy is critical for successful implementation of demand reduction solutions. What measures would work in this constrained network? What is the savings potential at each hour on a peak demand day? How can we measure reductions to a level of accuracy that will provide system planners with confidence to defer or cancel planned infrastructure upgrades? No doubt, answering these questions requires data. Yet simply gathering more data without a thoughtful approach can lead to costly efforts that do not significantly improve program delivery.

When planning a data collection effort, it is difficult to determine what data is useful. What equipment should be metered? How many meters should be installed? What should the meter sampling rate be? Trying to answer these questions without direction quickly leads to the common debate about the merits of owning your very own "Big Data."

Through this data collection effort we were forced to ask the question: How will this data be used? We were then able to develop a process of how to apply a data-driven approach to implement successful targeted demand side management (DSM). With this process in place, it is much easier to define the scope and purpose of a data collection effort. Upon reading this paper, we hope that system planners, program managers, policy makers, and other stakeholders can use this process to initiate and design their own data collection efforts that support successful implementation of demand side reductions as a reliable grid resource.

Data Collection Summary

The data-driven approach to DSM programs presented in this paper stems from a study commissioned by Con Edison to gather a comprehensive and representative data set including (1) characteristics of all electric energy-consuming equipment and (2) the corresponding hourly load shapes down to the equipment level. This effort focused on three important customer segments: multifamily in-unit (MFIU), multifamily common area (MFCA), and small business (SB). These three segments were of interest since they account for almost two-thirds of the billed energy use in the territory and require a mass-market approach to achieve the magnitude of demand reduction needed.

ERS installed over 2,700 metering devices at a statistically representative sample of 127 small businesses, 42 multifamily buildings, and 108 in-unit apartments across the 42 multifamily buildings. This data collection effort had direct access to each site, which provided the ability to collect additional data of critical value. Specifically, at each site field engineers inventoried all electric energy-consuming equipment and collected equipment counts, nameplate wattages, and efficiency metrics. The team also surveyed 156 customers to gather business and demographic data that can be used to translate this study's results to future constrained areas or populations of interest. The 8-month data collection period encompassed the warmest periods (most grid-constrained days), as well as the winter and swing seasons.

The resulting load shapes, in conjunction with the equipment inventories, were analyzed to develop an understanding of how BQDM customers consume electricity in order to improve forecasting, to better estimate program-sponsored demand reductions, and to identify cost-effective and reliable methods to achieve further reduction during the hours of greatest need. A summary of ERS's comprehensive data collection efforts is presented in Table 1.

Table 1. BQDM Metering and Market Characterization Data Collection Summary

Segment	Sites Visited	Loggers Deployed				Pieces of Equipment Inventoried	Customer Surveys Completed
		Lighting	Plug Load	Amperage	Total		
MFIU	108	537	416	86	1,039	3,100	62
MFCA	42	376	16	117	509	3,400	0
SB	127	530	285	397	1,212	4,200	94
Total	**277**	**1,443**	**717**	**600**	**2,760**	**10,700**	**156**

This data collection effort obtained building-level and equipment-level metered data, and the direct interface with customers provided the ability to collect an extensive set of equipment and customer characteristics. This data collection effort was designed based on the needs and existing data sources of Con Edison. We intend for other utilities or stakeholders to use this paper to review what data is needed for their DSM program, and then review what data they already have.

Applying Data to Effective Program Delivery

The data collection efforts – metering, equipment inventories, and customer surveys – can be sources of insight at all stages of program delivery, from the initial planning through the verification of program results. This section provides a summary of the data-driven approach as applied to the BQDM program. Each step features an overall question to be answered, a description of what data is needed to provide actionable insights, and a demonstration of the value in applying a data-driven approach. The proposed process is summarized in Figure 2 and broken out in the sections that follow.

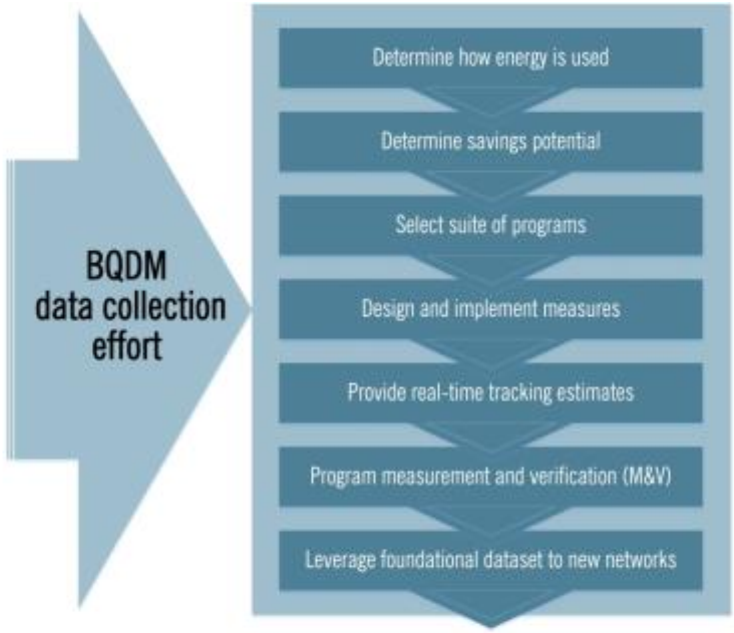

Figure 2. Program delivery process

Determine How Energy Is Used

Understanding how energy is used is critical for successful implementation of energy efficiency as a resource. Is it sufficient to meter the full building load or is equipment-specific metering needed as well? Should all energy-consuming equipment be metered or only specific equipment of interest? What defines an end use "of interest"? How granular should the data be? For example, what is the value in metering window air-conditioner (A/C) energy use for each customer segment, multiple space types, or by various demographic populations?

These questions illustrate how data collection efforts can quickly become overwhelming. Hence, the logic of where energy is being used must be incorporated when designing data collection efforts. This allows for an iterative data collection approach that only targets the collection of valuable data.

One BQDM example highlights the value in understanding the breakdown of energy use of the hundreds of pieces of equipment prevalent in multifamily apartments. Early discussions with Con Edison indicated that before the company understood where energy was being used, it was difficult to assign value to gathering data on specific pieces of equipment. For example, microwaves and toasters are highly prevalent with large rated wattage, but what is the resulting energy use after their intermittent operation is accounted for? Conversely, cable boxes have constant operation due to their need to always be connected to their respective server, yet does their lower prevalence and rated wattage make them more or less desirable for data collection than the previous appliances? How do these devices compare to equipment such as a window A/C that has significant energy use and also associated energy efficiency measures?

Figure 3 shows the peak-hour demand for these key end-uses within a typical apartment in the BQDM territory. It quickly highlights how the conversation about toasters, microwaves, and cable boxes is irrelevant when compared to the energy use of window A/Cs. And with the 52 MW reduction target, scalability is critical.

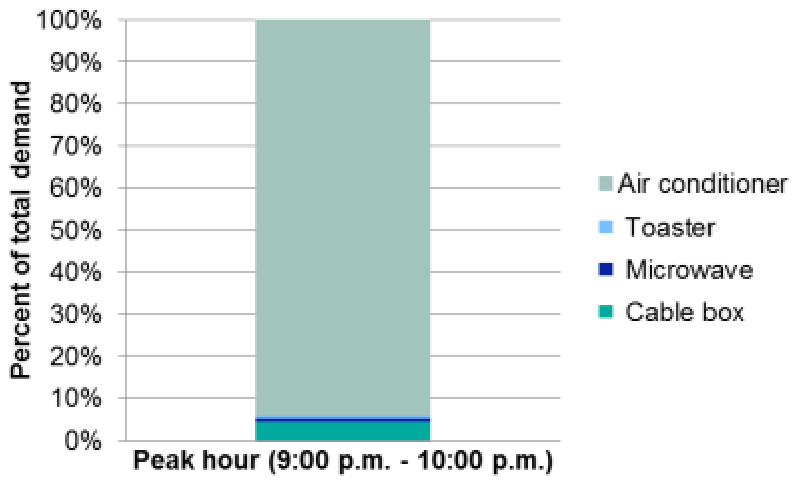

Figure 3. Multifamily in-unit peak-hour demand for key end uses.

Once the variables are well understood, the question becomes: What equipment should be metered? The focus of equipment metering should be for equipment that has viable demand reduction potential. The energy use of all other equipment is only of value for battery installations to offset facility use, or to anecdotally understand where energy is being used in the territory of interest. The potential for battery installations to offset a facility load can be determined though the whole-building load profile.

Determining what programs can be realistically implemented affects the targets of a data collection effort. For example, nearly every small business was found to have a lottery and ATM and nearly every residence had a cable box, all of which operate continuously. There may be some upstream potential for these measures, but if that potential cannot be reached in time for the network need, that equipment should not be targeted. However, other entities within the utility, such as forecasters or planners, may find such load shapes informative.

Data collection efforts would produce additional value per dollar by recording key characteristics through a survey for all other equipment that is not expected to have a viable demand reduction measure. From this survey data, such as operating schedule, rated wattage, and equipment quantity, a high-level estimate of the demand reduction potential can be developed. Stakeholders can use this estimation to value the benefits of future data collections that target specific equipment for potential new measures. Simply metering all pieces of equipment in hope of finding unexpected low-hanging fruit is not worth the added expense and complications.

Once a decision is made on the equipment of interest, one must determine the volume of data to be collected. Specifics such as the number of sites, data logging duration, and granularity (logger interval) can lead to "big" data, but without "big" value. There are three concepts to account for when determining how much data to collect:

1. Will the extra data allow me to implement my program more effectively?
2. Will the additional data allow me to better measure demand reduction?
3. Will the additional data allow me to better leverage this data to new networks?

Answering these questions requires diving into the specific stages of DSM delivery. These are explored in the following sections.

Determine Savings Potential

After determining the equipment of interest, the next step in the data-driven process is to understand the demand reduction potential. This requires metrics that will translate energy use into an estimate of demand reduction. The key data needed is the existing equipment type saturation within the targeted area, which was gathered by ERS as part of comprehensive equipment inventories throughout the BQDM territory.

For example, in order to estimate the total potential demand reduction from the replacement of inefficient A/Cs, the energy-use profile (Figure 2) must be combined with the difference between the average observed energy efficiency ratio (EER) from the end-use inventories and a high efficiency EER replacement.

Table 2 provides results of the equipment inventories and illustrates the distribution of EERs for all window A/Cs in the territory. With this information, Con Edison now has the ability to answer questions such as: What is the potential savings if all air conditioners are replaced with new ENERGY STAR units? What if the program only replaces units that have an EER of 10 or less? What is the difference in cost per MW of the two approaches?

Table 2. BQDM Multifamily In-Unit Window AC EER Values

Window A/C EER Range	Percentage of Total
<10	63%
10–11	35%
>11	1%
Unverifiable	1%
Total	**100%**

Note: Additional granularity was gathered as part of the data collection effort.

A second example involves calculating the lighting savings for various customer populations. The end-use inventory provides the type of fixtures that can be replaced, while metering indicates the potential demand reduction at each hour. As part of the data collection effort described above, the potential for lighting savings was provided across many different customer segments. First, lighting potential was provided for the multifamily in-unit, multifamily common area, and small business segments. Then the small business lighting potential was broken down by business type and even space type, such as ”sales area” and “stockroom.” These breakdowns feed into planning the optimal suite of programs as discussed in the next section.

An example of an actionable finding is that 59% of multifamily in-unit lighting was still incandescent. Furthermore, a survey of customers indicated that customers regularly went to their local hardware or grocery store where incandescent bulbs were still sold. Multiplying the percentage of prevalence by the energy-use profile, and combining with a LED replacement savings yields the hourly potential demand reduction for an incandescent to LED program.

Select Suite of Programs

Once the measure-specific savings potential is determined, program selection can be optimized to implement the specific programs that will achieve cost-effective reduction at the constrained hours. With the magnitude of savings known at each hour, planners can estimate the cost-effectiveness of implementing various programs. The additional data needed for this step is the cost of implementing the energy efficiency measures. With the lighting example, this would include the cost per fixture, installation cost, M&V, and administration of the program.

With this cost data, planners can explore the complementary nature of programs to achieve the demand reduction needed at all hours. The BQDM effort has a 12-hour span from noon to midnight, with the highest load occurring at 9:00 p.m. Small business lighting has a peak during the day, while multifamily in-unit lighting peaks in the late evening hours. When combined, these two programs provide consistent savings over multiple hours of need. Figure 4 illustrates the complementary shape from a small business and multifamily in-unit lighting programs.

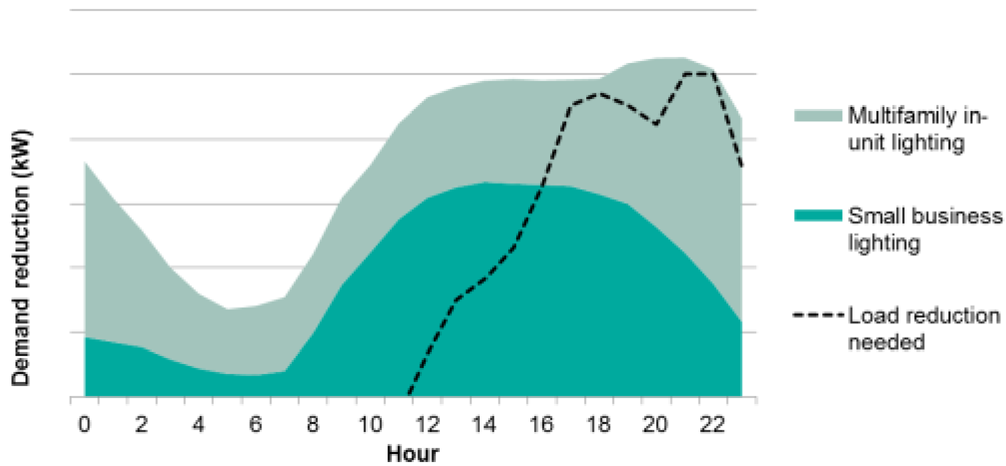

Figure 4. Complementary nature of Small Business and Multifamily In-Unit lighting programs

Design and Implement Measures

Program design should be taken into account when determining the different breakdowns of equipment to meter. This detailed information makes it easier to achieve cost-effective programs through pay-for-performance incentives or targeting the highest-saving customers. Program designers can target measures that reduce customer usage during a network's constrained hours.

For example, this study found considerable differences in lighting coincidence factors (CFs) by small business type. Specifically, the potential savings can differ by a factor of 10 simply based on the type of business that was selected. Restaurants were among the best performers (highest CF) in the evening hours, while offices were among the worst (lowest CF). The difference in CF between offices and restaurants was almost a factor of 10. This study also found a strong correlation with lighting operation and business hours. Gathering metering data for these types of breakdowns can aid in program design and inform program managers how to target customers strategically.

Provide Real-Time Tracking Estimates

Load shapes by business type, space type, end use, or other indicative variables can be used directly to compute measure demand reduction in program tracking systems. Using these factors to calculate demand reduction in real time will lead to more accurate tracking estimates, giving program administrators an effective indicator of actual program savings in real time.

The previous section discussed using data to target customers that would provide cost-effective savings at the hours of need. Data can also be used to help the program managers track how effective the implementation team is in utilizing this information. For example, a real-time dashboard would indicate whether savings were being achieved at the hours of interest. If specific business types were of interest, the dashboard could tell how many businesses of each type had efficient equipment installed and their respective demand reductions. Ultimately, data enables planning, implementation, and evaluation to operate in a feedback loop rather than operating as separate aspects of the delivery process.

Program Measurement and Verification

Even with so much data in hand, ongoing verification will be required to ensure that the program is installing the correct equipment, and ongoing metering to ensure that the operating profile has not changed. Implementing load shapes into tracking, as discussed in the previous step, will lead to fewer deviations when M&V results are finalized and a reduction in the amount of subsequent on-site verification metering to achieve the desired confidence and precision of results.

Leverage Foundational Data Set to New Networks

In order to make data collection efforts cost effective, efforts should be designed with the intent to translate results to multiple populations. To translate data to new networks, detailed customer interviews were conducted to gather demographic, physical, and operational variables. For example, in-unit lighting load shapes are expected to be dependent on tenant work schedules and the number of occupants. Results can therefore be translated to other networks once parallel survey data is gathered from a sample of the new networks' customers.

Conclusion

When designing a new IDSM programs and using energy efficiency as a resource, stakeholders would benefit from reviewing this data-driven process before any data collection effort is designed. Through this practice stakeholders can determine what questions they need the data to answer, what data is needed to answer these questions, and what data is currently available. With this data-driven approach, Con Edison is improving its ability to implement successful IDSM from to initial planning through M&V. For the BQDM effort, Con Edison is learning to use this data set to vet newly proposed measures, optimize the implementation of current programs, and provide more accurate and cost-effective M&V. Furthermore, Con Edison plans to use this data set to assess the potential for utilizing energy efficiency as a viable resource in future constrained networks.

Armed with the correct data set, program administrators will find themselves in a position to develop a data-driven approach to customer-level demand management. End-use load shapes, in conjunction with equipment inventories, enable stakeholders to understand how customers consume electricity in order to improve forecasting, better estimate program-sponsored demand reductions, and identify cost-effective and reliable methods to achieve further reduction during constrained hours. This approach provides the necessary tools to successfully defend and implement demand side management as an effective grid resource.