Duckhunt! Benefits and risks of load disaggregation and end use metering for determining end use loadshapes

Terese Decker, Justin Spencer, Hemang Nerlekar, Mark Bielecki, Navigant, Boulder, CO
Vergil Weatherford, Navigant, Tallahassee, FL
Justin Elszasz, Navigant, Washington, D.C.
Kimberly Crossman, National Grid, Waltham, MA
Riley Hastings, Eversource, Westwood, MA
Bob Wirtshafter, EEAC, Chatham, MA

ABSTRACT

Many utilities face program design and resource-planning challenges that require them to understand their customers' hourly consumption behavior, particularly: quantifying how much energy demand side management saves, the potential for savings, and *when* the savings occur. With these needs in mind, three research teams working with a Southwest utility, a statewide partner in the Northeast, and a utility in the Southeast are currently engaged in projects to measure loadshapes to aid utilities with their unique challenges. New evaluation techniques allow for a major reduction in the cost of both direct measurement and load disaggregation approaches. These three research teams use different combinations of direct measurement and load disaggregation to derive loadshapes.

Building upon lessons learned from each of these data collection approaches, this paper explores the cost versus accuracy trade-offs between using a whole-building load disaggregation approach versus a traditional end use specific metering approach, and recommendations for when to use each approach. Evaluators can apply this to all sectors of evaluation and load research. These trade-offs may include costs, skills required to install, data accuracy, schedule, and end user inconvenience. The paper also describes how using a combination of techniques may improve accuracy and reduce costs.

Background

Some utilities and regulators have recently shifted their energy efficiency programs' focus from meeting energy savings targets toward optimizing both energy and capacity savings and cost effectiveness. Additionally, many utilities are experiencing changes in when their peak loads occur, which has caught some by surprise (PJM, 2015). This means that utilities have a greater need than ever to understand their customers' time of use consumption by end use to understand future load impacts of disruptive technologies and identify opportunities for load shifting and reduction. This fundamental shift has required utility program and resource planners to investigate which end use specific consumption behaviors are driving the system load now and which will drive peak loads in the future. For example, some end uses tend to consume energy with a constant load or off-peak (dehumidifiers), while some end uses are typically running at the same time as the utility peak (central air conditioners). Thus, different energy efficiency measures have disparate impacts on capacity. Further, emerging technologies such as heat pump water heaters and electric vehicle chargers may add a lot of uncertainty to future system-wide loads. When combined with advanced controls these loads may be managed to reduce their peak impacts.

Net metered solar photovoltaic systems are one such technology that can pose challenges to stable grid operation. If the utility's peak aligns with the solar system's highest production time (say, 1:00pm), then solar is helping the grid. However, in most cases, the utility's residential peak is not coincident with solar's maximum generation. When this is the case, utilities are also faced with the "duck curve problem", in which the curve depicting the overall system load less solar generation over a typical day has an increasingly steep slope just after the sun goes down for the day (NREL, 2015). This large slope indicates that utilities are facing requirements to provide an increasingly rapid ramp which constrains their systems. This phenomenon is illustrated in Figure 1 (SolarGain, 2016).

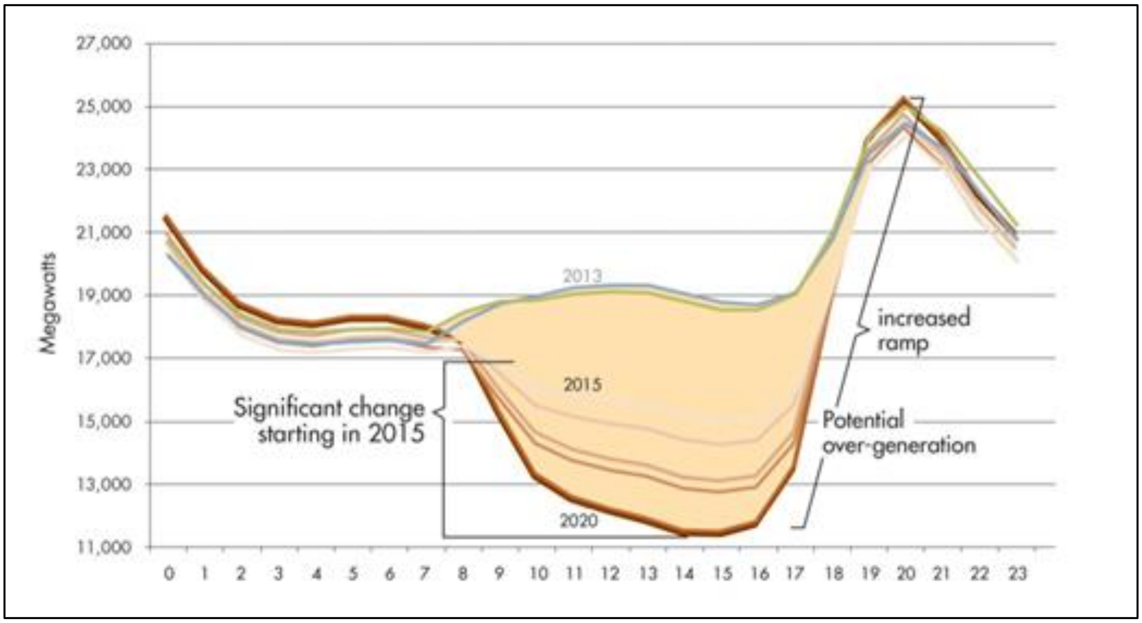


Figure 1. Duck Curve Illustration. *Source:* SolarGain, 2016

For these reasons and others, end use specific loadshapes – which would clearly identify when a load is present for each specific technology – are becoming increasingly valuable to utilities as energy efficiency and demand response programs aim to target more specific segments and shift specific loads. Further, having end use specific loadshape data can help inform a host of other research needs, including energy efficiency potential studies, transmission and distribution planning, load forecasting and optimization, rate design, and market effects research.

Introduction

Three separate research teams at Navigant have been working alongside their utility clients to provide the information needed to address the unique challenges they face. Each of the utilities involved has their own challenges and different specific objectives, and therefore each team has employed slightly different approaches. Each team aims to optimize the value provided to the utility while minimizing data accuracy concerns and risk associated with primary data collection.

This paper presents each of these research studies in case study format, emphasizing the loadshape determination method used. For each case study, the authors elaborate on the specific challenges faced in the region in which the utility is located, how the research team approached the challenges, the outcome of the research, and a matrix detailing the specific risks and benefits of the data collection method used. To distill these case studies into a decision making framework, the project evaluators have determined a few key variables that will influence the decision, namely:

- **What level of rigor is desired at the end use level?** In this case, rigor is associated with confidence and precision at the end use level, as well as data granularity.
- **How many end uses per site are of interest to the project initiator?** A project should be approached differently in terms of equipment and methods depending on how many end uses need to be researched.
- **Is advanced metering infrastructure (AMI) data available in the territory of interest?** Availability of AMI data will add a few options for project initiators looking for low or medium rigor answers to their questions.
- **Are measures of interest suitable for disaggregation?** Recent research shows that not all end uses are suitable for disaggregation based on their unique load profile in the micro and macro senses.

The Discussion and Conclusions portion of this paper will outline a method for determining when to use each approach.

Case Study: Southeastern state DR program evaluation

This section outlines work underway for a utility in a Southeast state that is interested in understanding savings and savings potential of high impact demand response (DR) measures.

In the Southeast, the goal is to estimate impacts and develop annual loadshapes for pool pumps, water heaters, and HVAC to understand the impacts and potential impacts of a utility's DR program. The relevant decision variables for the evaluation method are shown in Table 1.

Table 1. Southeast State Decision Variables

Decision/Context Variable	Case Study Specific Variable
Objective	Understand savings and savings potential of high-impact DR program measures
What level of rigor is desired at the end use level?	High
How many end uses per site are of interest?	Three: pool pumps, domestic hot water, HVAC
Is AMI data available?	No
Are measures of interest suitable for AMI disaggregation?	N/A
Are measures of interest suitable for measured data disaggregation?	No

Traditionally, deploying passive CT amperage loggers has been the most cost-effective way to gather residential energy consumption data evaluating demand response. Paired with spot measurements of voltage and power factor, interval amperage readings can provide a reasonable estimate of true power, but come with some limitations. The estimate of real and reactive power is extrapolated, not measured directly. The interval meters don't collect information on ramp rates or short duration impacts. These passive loggers are typically battery-powered, and data cannot be read unless a technician visits the site with a laptop and data transfer cable to retrieve it, which makes it difficult to do periodic quality control checks on the data.

More recently, in a year-long residential direct load control evaluation in the southeast, Navigant deployed 5-minute interval passive loggers for approximately 80% of the sampled homes, and deployed 1-minute internet-connected interval loggers for the remainder of sites. Logged loads include central

HVAC equipment, water heaters, and pool pumps, primarily logged at the main service panel of the home. Customers were given an added incentive for allowing use of their broadband internet connection so the evaluation team could access the data remotely. This study of 190 sites carries limited risk and involves a relatively narrow scope, but provides great value and high quality data around the most important end uses for the utility's demand response programs.

With relatively simple goals and a relatively simple metering approach, the team answered the basic utility needs with high confidence. The modest increase in cost and installation time for the networked loggers resulted in a valuable QC tool. During early stages of the study, the evaluation team worked with the utility program team to ensure study participants were programmed in the utility's direct load control system correctly, and are truly called separately from the larger population, as intended. Furthermore, since the technicians configured the meters to log true power as well as apparent power and voltage, the evaluation team can use the data to confirm the accuracy of spot measurements combined with amperage to estimate true power. Final data from this study will be available in early 2018.

Case Study: New England statewide pilot for baseline study

In Massachusetts, a consortium of eight program administrators (PAs), including seven utilities, are sponsoring research to derive baseline loadshapes for all major residential electric end uses. The results of this study will be used to inform their energy and demand program planning and other initiatives. They have the following stated research questions:

- What kind of energy-consuming equipment is present in homes in the state, and what is the efficiency of that equipment? What are baseline saturations of various inefficient devices?
- How and when are people using the electric equipment described above?
- Why do people use the equipment described above the way that they do?
- What is the distribution of inefficient equipment across the varying program administrators, building types, building ownerships, or other key parameters in the state?
- How do energy efficiency opportunities align with current program delivery channels?

Some basic information about the PAs' objectives and data needs are outlined in Table 2 below.

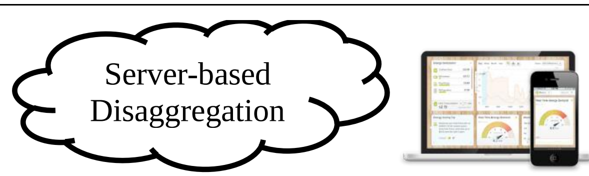
Table 2. Massachusetts Decision Variables

Decision/Context Variable	Case Study Specific Variable
Objective	Fundamentally understand current and future statewide energy usage and peak demand by end use
What level of rigor is desired at the end use level?	High
How many end uses per site are of interest?	All significant residential end uses (see next paragraph)
Is AMI data available?	No
Are measures of interest suitable for AMI disaggregation?	N/A
Are measures of interest suitable for measured data disaggregation?	No

The PAs decided upon specific end uses to meter and a target relative precision for each end use associated with their stated priority of understanding the end use's load. Many end uses on the list were determined to be high priority and targeted a relative precision of at least 15% at a one-tailed 90% confidence level, such as central air conditioners and heat pumps, space heaters, boiler circulator pumps, furnace fans, electric water heaters, and dehumidifiers. Several end uses were determined to be medium and low priority – associated with 20% and 30% relative precision, respectively – such as room air conditioners, electric baseboard heat, major kitchen appliances, heat pump water heaters, laundry equipment, sump pumps, and pool pumps. Other end uses were given no priority, meaning that the team would meter them as they encountered them but there were no targeted precisions associated with the sample design. These end uses included ground source heat pumps, tankless water heaters, aquaria, large battery chargers, well pumps, electric vehicle chargers, and television peripherals that are on a power strip with the television. The team also agreed to meter the whole home load and any other substantial loads that were encountered onsite to help understand the total size of other loads in Massachusetts homes.

To achieve an appropriate confidence and precision for each measured end use, the research team determined that the study would include a large sample. The research team identified that this would be a good opportunity to deploy a nested sample of end use metered sites within a larger sample of non-intrusive load monitored (NILM) sites in order to reduce costs while achieving a high rigor result. The approach would include leveraging disaggregated NILM data in conjunction with metered data to reduce the sample size of fully metered sites. However, the team and PAs determined that it was only worthwhile to employ NILM if it provided cost savings as compared with metering all the end uses of interest at a larger sample of sites.

Because AMI data is not widely available in this state, the research team needed to use a whole home monitor and visit each site in the sample individually, thus adding to the cost and risk of the study. There are four primary methods by which NILM is implemented in the home; each of which is characterized by different advantages and disadvantages in terms of cost, installation, accuracy, and granularity of disaggregation. The team considered each of these options for measuring whole home usage. Figure 2 illustrates the different NILM technology types and identifies key providers of said technologies that the research team considered.

Technology A: Software-Only Solutions *Require third-party hardware for data collection. Data transmitted to software vendor server where load is disaggregated.* Providers include Bidgely, Plotwatt, and EEme.	
Technology B: Utility Smart Meters *A utility smart meter accompanied by a Wifi-connected gateway. Disaggregation occurs after data are transmitted to a software vendor server.* Providers include Sensus, Landis + Gyr, Rainforest Automation, Schneider Electric, Itron.	
Technology C: Utility Meter-Reading Devices *A device installed at the utility meter used to digitize and transmit meter data. Includes both optical sensors (cheaper) and meter base sensors. Disaggregation occurs after data are transmitted to a software vendor server.* Providers include Blue Line Innovations, Wattvision, Enetics.	

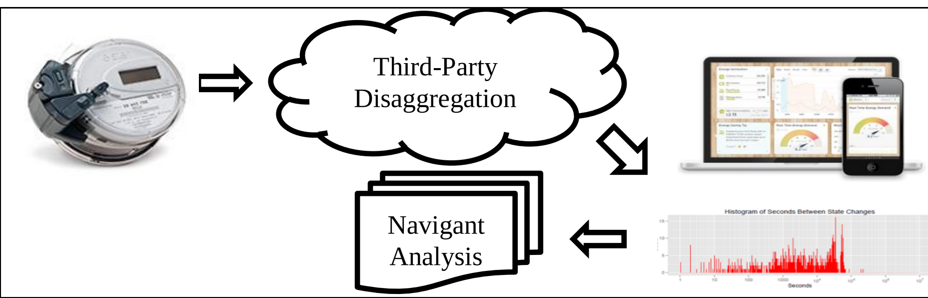
Technology D: Current Transformer (CT) Based Devices *Hardware for monitoring voltage and current installed in the home electrical panels. Disaggregation occurs after data are transmitted to a software vendor server.* Providers include Navetas, Energy Inc. (TED), Neurio, Current Cost, eGauge, Smapee.	

Figure 2. NILM Hardware and Software Technology Options. *Source:* Navigant analysis, images courtesy www.bidgely.com, www.bluelineinnovations.com, www.theenergydetective.com, www.enernetics.com, and www.itron.com

Among the differentiating characteristics of the hardware options described above, the frequency at which data are collected is particularly important because it directly impacts the level of disaggregation possible (Armel, 2012). Alternatively, electrical panel current transformer devices and optical meter-reading sensors designed for home energy monitoring are capable of much higher frequency in data reporting and storage. This allows for much more granular disaggregation of electrical end uses. However, their installation, maintenance, data collection processes, and failure rates present challenges for large-scale studies and drive up overall costs.

Given the limitations associated with Technologies B and D, the research team chose to leverage a combination of Technologies A and C. The most prominent commercial configurations currently used to implement this NILM solution are delineated in Figure 3.

Scheme 1
Optical sensor (Technology C) attached to utility meter and third-party disaggregation software (Technology A) (BlueLine PowerCost)



Scheme 2
Utility meter-based device (Technology C) with accompanying disaggregation software (Enetics)

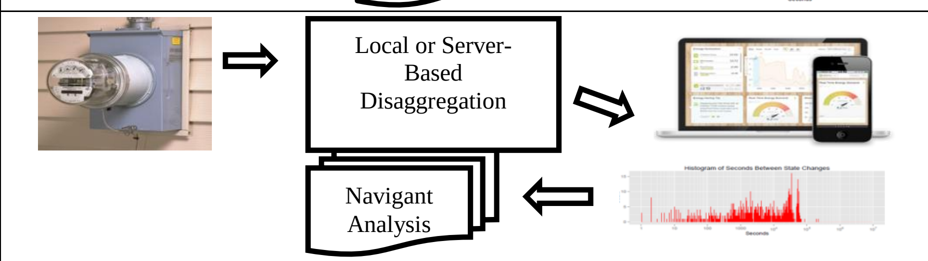


Figure 3. NILM Schemes

Both of these commercially viable data schemes require third-party disaggregation. For the Phase 1 pilot, Navigant contacted three third-party disaggregation firms who were reputable and proven providers. Of these three, one of the providers agreed to participate in the Phase 1 project.

This third-party software needs to be paired with a hardware device collecting data directly from the home as shown in Figure 3. The two options above differ in exactly how the NILM hardware device integrates with the meter. In Scheme 1, the device is physically separate from the electrical wiring and uses an optical reader to “watch” the meter. This means it can be installed without an electrician and without interfering with the operation of the meter. However, optical readers can experience issues with weather impacting the readings, introducing another source of error. In Scheme 2, a device is physically integrated with the meter and the mains electricity coming to residence. This means it requires a utility electrician to install it. It may also require additional regulatory approval, which introduces another layer of risk.[1]

Given the uncertainty around regulatory approval and the cost of installing a physically-integrated device, Navigant pursued the lower cost optical reader device, while planning for and allowing for the fact that hardware would fail and that there may be accuracy issues that increase the intensive end use metering requirements to "true up" the results.

The uncertainty in this technique created too much risk to proceed without first testing the methods on a smaller sample – 600 online surveys, 45 onsite validation and load disaggregation sites, and 23 full end use metered sites (Hastings, 2017). The nested sample allowed the team to use a triple ratio estimation method to extrapolate results to a wider sample while reducing risk (Spencer, 2013). The major advantage of this approach is that it did not require the chosen methods to provide perfect results given that the team could leverage higher accuracy data to true up the sites sampled for less accurate methods. In theory, this approach optimizes strengths by utilizing both big data analysis and innovative engineering to provide a cost-optimized, low risk, high quality solution. The team moved forward with a pilot project which they called "Phase I". The objectives of the Phase I metering effort were to:

- Develop and rigorously test onsite data collection tools and protocols
- Determine the predictive power of NILM for each end use being tested and refine the circumstances under which NILM will prove useful for the full-scale study
- Validate coefficient of variation (CV) assumptions to refine sampling plan
- Determine the feasibility and expense of measuring each desired end use individually

Because of the volume of data desired at each site and the complexity of metering all end uses in a home, the team used a sophisticated sampling scheme and remote data collection to further optimize costs. This effort involved the collection of both whole house electric consumption data via home energy monitors as well as end use metering data via the use of plug load meters and circuit current transducers (CTs).

The team took advantage of the pilot nature of the project by using various analytical and regression techniques. Both in-house Navigant staff and a third-party disaggregation provider disaggregated the whole house data into hourly load curves for specific end uses (Elszasz, 2017). The resulting load curves were then compared against the actual load curves. Coefficients of variation were determined for each end use for the disaggregation data and for the end use metering data. The CVs for each approach were then compared to ascertain the accuracy of the disaggregated load curves.

Throughout the Phase I data collection, Navigant found that the average per site cost of end use metering was less than anticipated (about $3000/site). In contrast, the average per site cost of whole house metering via home energy monitors was greater than anticipated (about $1000/site). This difference was primarily due to fewer hours required per site for the set-up and collection of end use metering data, and substantial additional time necessary to ensure that the whole house energy monitors stayed online and provided continuous, usable data[2]. These costs did not include the data disaggregation or any other analysis and data cleaning.

[1] In the study being conducted in NY, the study was delayed for many months while the Enetics device sought regulatory approval as a metering device.

[2] The whole house energy monitors chosen for this study were found to be systematically unreliable. They were sensitive to misalignment and regularly disconnected from the network. The monitors are intended for end user

The combination of these findings heavily influenced the trade-off between utilizing whole house data that is disaggregated using NILM in combination with end use metered data for end uses that cannot be readily disaggregated, as compared to end use metering for all end uses. The ten loads that were tested using NILM techniques and their disaggregation test status are shown in Table 3 below (Elszasz, 2017)[3].

Table 3. Tested End Uses and the Disaggregation Status of Each Using Two Methods

End Use Category	NILM Tested End Use	Disaggregation Method 1 Result	Disaggregation Method 2 Result
Heating and Cooling	Central air conditioners	PASS	PASS
	Room air conditioners	PASS	FAIL
	Ductless heat pumps	Inconclusive – too few sites	Inconclusive – too few sites
Domestic Hot Water	Electric water heater	FAIL	Not tested
Kitchen	Refrigerator	FAIL	Not tested
	Second Refrigerator	Inconclusive – too few sites	Inconclusive – too few sites
Laundry	Clothes dryer	FAIL	FAIL
Miscellaneous	Dehumidifier	Not tested	FAIL
	Sump pump	Not tested	FAIL
	Pool pump	FAIL	FAIL

This effort proved that the disaggregation of whole house energy usage data into individual end uses is only worthwhile to the study if this approach allows the researchers to gather data from a larger number of houses at a lower cost, while simultaneously producing comparable results for individual end use load curves. The lower than expected ratio of the cost for end use metering at a single household relative to whole house monitoring with disaggregation undercuts one of the main motivations of utilizing both metering techniques.

In addition, the predictive power of the disaggregated load curves estimated using NILM techniques was found to be inconsistent. Limitations in the reliability of the whole house meters lowered the quality of the data available for end use disaggregation and increased the costs of whole house data collection, due to the need to frequently monitor the status of meters.[4]

Further, the most difficult loads to meter would provide the most leverage by eliminating some of the cost and burden of metering. However, the most difficult loads to meter are 120V hardwired loads because they require an electrician to meter at the panel and sometimes meticulous circuit tracing if they are not on a dedicated circuit (which cannot be determined definitively until the electrical panel is accessed). Examples of 120V hardwired loads are boiler circulator pumps, furnace fans, some dishwashers, and most pumps. On the other hand, 240V hardwired loads and 120V plug loads are more easily metered because they do not require extensive circuit tracing. The only load types that passed the NILM test – cooling loads – were easier to meter, thus gaining less leverage than originally hoped.

After significant evaluation of the approaches and based on these results and findings regarding the cost of data collection for whole house and end use metering approaches, Navigant recommended that the full scale Massachusetts Residential Baseline study not depend on disaggregation of whole house

engagement, and thus, were not ideal for research purposes without proactive troubleshooting.

[3] The data used in this analysis was collected starting in July 2016 and ending in October 2016. No heating season data was available to test heating equipment for NILM methods. Therefore, heating end uses are excluded from Table 3.

[4] Alternate hardware could fix this problem.

data. The reasons for this recommendation are 1) the relative cost of whole house monitoring and whole house with end metering data collection, 2) the low quality of the data collected via low cost whole house monitors, and 3) the inconsistent accuracy of the estimated end use load curves based on disaggregation techniques such as NILM.

Case study: Southwest state study to overcome the Duck curve

This section is a case study for a utility in the Southwest that is interested in targeting the Duck Curve problem and making informed program planning decisions.

In the Southwest, utilities are facing several challenges that can be helped by leveraging more granular data collection including: a steep Duck Curve, stringent cost effectiveness testing for their existing energy efficiency measures, a need to expand their efficiency program portfolio to meet their targets, and to understand potential distribution benefits from conservation voltage reduction (CVR).

Table 4. Southwest State Decision Variables

Decision/Context Variable	Case Study Specific Variable
Objective	Assessing hourly impacts of DSM to identify high-value measures and programs to alleviate Duck Curve issues
What level of rigor is desired at the end use level?	High
How many end uses per site are of interest?	Four+: pool pumps, water heaters, central air conditioning, other common appliances
Is AMI data available?	Yes
Are measures of interest suitable for AMI disaggregation?	Potentially
Are measures of interest suitable for measured data disaggregation?	No

In the Southwest, the research team is exploring load disaggregation solutions in combination with direct measurement at approximately 50 residences – including about 25% with net metered solar photovoltaic systems – to provide end use loadshapes for program and resource planning models. For each site, an independent software provider will disaggregate the hourly AMI load data associated with each of the 50 sites in the onsite sample as well as a random selection of 50,000 residences that will be used as a control group. The evaluation team will use the end use metered data to validate and true-up the disaggregated data and extrapolate the true-up to the larger sample in order to leverage load disaggregation for an annual loadshape update using high quality AMI data. This approach optimizes resources to allow for annual updates of program planning and evaluation while reducing risk and recurring costs associated with primary data collection. The presence of AMI data differentiates this study from the Massachusetts study referenced above. This further optimizes resources by allowing for a larger sample of whole home interval data without incurring onsite data collection costs. If leveraging AMI data proves to be unsuccessful due to the one-hour granularity, researchers could consider gathering additional whole home data at a finer granularity, which may increase chances of success.

Discussion and Conclusions

Each study detailed in this paper has achieved the goals desired by the utility clients. Based on the scoping considerations determined from each study and the data gathered during the study, Navigant has simplified the decision-making process into a basic decision tree. The decision tree below in Figure *4* outlines how utilities can determine which methods might be ideal for their projects.

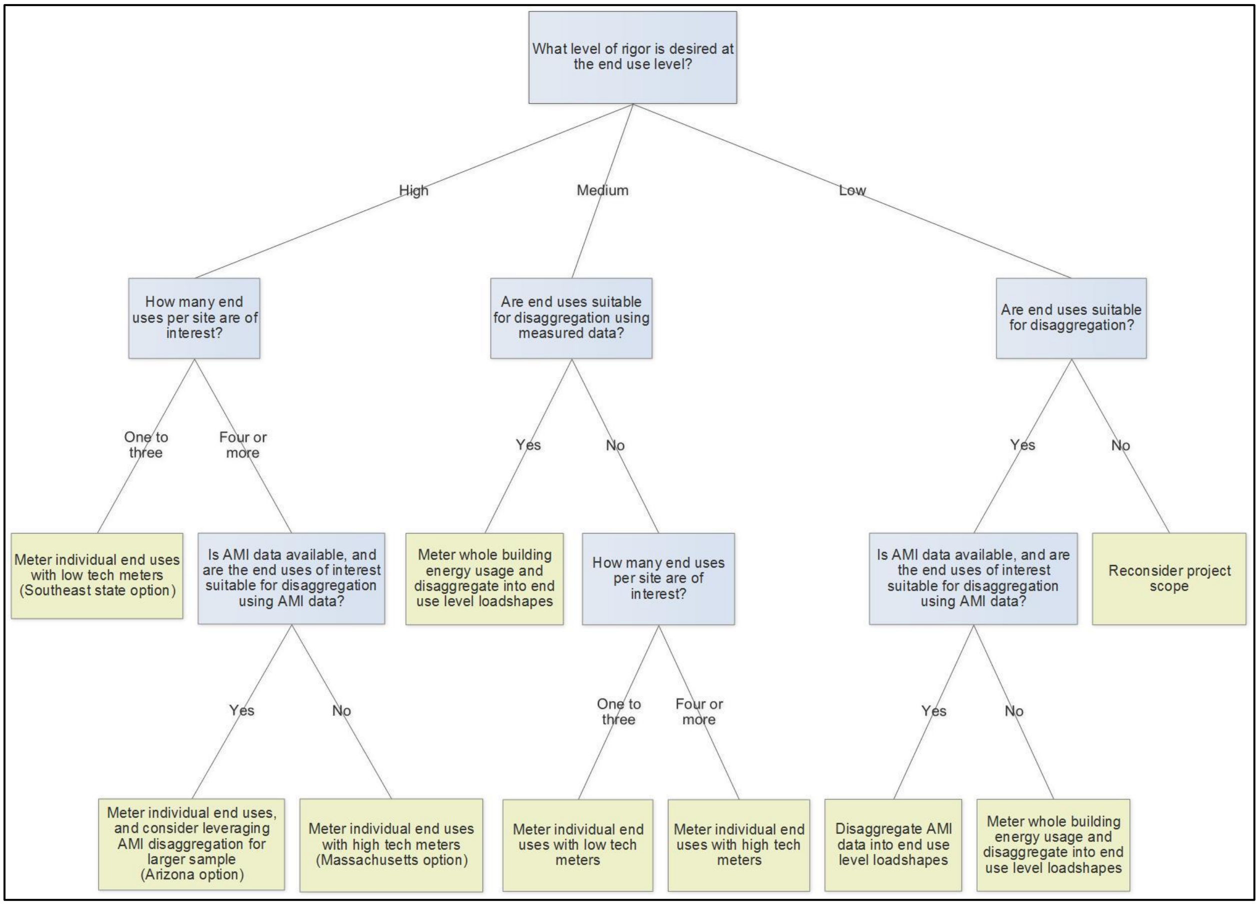


Figure 4. Decision Tree for Loadshape Measurement Approach Options

The first question of the decision tree is related to the level of rigor desired by the project initiators. In this case, rigor is correlated with the desired quality of the data in terms of data granularity and data accuracy. For example, a high rigor project is defined as needing a high level of confidence and precision at the end use level. These projects will be ones related to understanding timing of specific loads down at the hourly level. All high rigor tracks of the decision tree conclude with metering. Navigant's research and tests completed to date have shown that while load disaggregation can work in some scenarios, if a study's goal is to achieve high data granularity, high data accuracy, or high confidence and precision, then it is not a good fit for the study.

All three of the case studies outlined in this paper followed the high rigor track in the decision tree because the clients requested either hourly or better data at the end use level or high confidence and precision at the end use level. From there, the projects in the three case studies differ in the number of end uses desired and the availability of AMI data.

The decision to use low tech meters or high tech meters is dependent on how many end uses the project is targeting. If the study's objective is to understand nearly every residential end use loadshape, such as in the case of the Massachusetts study, then high tech meters that collect true power consumption and are remotely connected via a powerline carrier connected to the home's wireless internet or a cellular router should be employed. If, on the other hand, the project is only geared toward understanding a few key end uses, such as in the Southeast state DR program evaluation, low tech passive CT amperage loggers that require manual data retrieval are still the most cost-effective option.

In general, Navigant found evidence that there are opportunities to leverage disaggregation (NILM) for the estimation of several end uses, but it will not provide a high rigor result (Baker, 2016). In particular, disaggregation methods estimated loads most accurately for cooling loads. However, there were limitations on the ability to disaggregate many of the hardwired loads, which are more challenging to meter directly. This is where the combined disaggregation and end use metering approach would have the greatest value, but it does not appear as though this approach is sufficiently developed as yet to achieve this aim as all non-cooling loads tested in these studies were either inclusive or failed to provide a benefit. Other disaggregation algorithms may work better in the future. Most difficult artificial intelligence problems are currently solved using deep learning algorithms, which are relatively brutish, but almost guaranteed to work eventually if given enough data for training.

If the project initiator wants medium rigor loadshapes– indicated by needing less granular and/or accurate data at the end use level – then the research team can consider leveraging a smaller sample of metered sites in conjunction with AMI data to get a good answer. However, if AMI data is not available to leverage then the research team should meter the desired end uses with high tech meters for four or more end uses and lower tech meters for three or fewer end uses.

It is assumed that low rigor projects will not use end use specific metered data, but instead either leverage AMI data or use inexpensive meters to measure whole-building usage for disaggregation. If the loads desired in a low rigor project are not suitable for disaggregation, then the research team should consider using existing data or reconsider the project scope.

At a time when real time evaluation is more appealing than ever, utilities and researchers should remember that the cost of metering at the end use level has fallen appreciably in recent years. In some cases, the best choice may be to meter everything needed to answer the research questions. The availability of internet-connected metering devices has great promise to improve the capability to perform quality control checks on collected data and analyze preliminary results on an ongoing basis or expedited timeframe. In other cases, the evaluation team may be able to effectively leverage AMI data disaggregation.

References

Armel, K. Carrie et al., 2012. "Is Disaggregation The Holy Grail of Energy Efficiency? The Case of Electricity," Energy Policy, Vol. 52 p. 213-234.

Baker et al., "Can We Find the End Use in Smart Metering Data?", ACEEE 2016.

Elszasz et al., "A Snapshot of NILM: Techniques and Test of Non-Intrusive Load Monitoring for Load Shape Development". IEPEC 2017.

Hastings et al., "Measuring up – How do our baselines compare?". IEPEC 2017.

NREL, 2015. "Overgeneration from Solar Energy in California: A Field Guide to the Duck Chart" http://www.nrel.gov/docs/fy16osti/65023.pdf. Retrieved 4/16/2017.

PJM, 2015. "The Polar Vortex, One Year Later." http://pluggedin.pjm.com/2015/01/the-polar-vortex-one-year-later/ Retrieved 4/14/17

Spencer et al., Revisiting Double Ratio Estimation for Mitigating Risk in High Rigor Evaluation. IEPEC 2013.

SolarGain, 2016. The Duck Curve -- And What It Means For You. https://www.solargain.com.au/blog/duck-curve-and-what-it-means-you. Retrieved 4/16/2017.