San Diego Gas and Electric Company

Emerging Technologies Program

Project ID ET13SDG1031

Residential Disaggregation

Final Report

Prepared For

Kate Zeng, Nate Taylor, & Matt Smith

San Diego Gas and Electric Company

Emerging Technologies

8306 Century Park Court

San Diego, CA 92123

Prepared By

Bo White, Marc Esser

NegaWatt Consulting, Inc.

www.negawattconsult.com

08/22/2014

Acknowledgements

San Diego Gas and Electric Company (SDG&E) and the authors of this report would like to acknowledge the assistance and cooperation of our host site participants and families that went out of their way to allow for our onsite testing and continual coordination. We would also like to thank the technology vendors that we evaluated for their valuable time and effort.

Project Team

SDG&E – Kate Zeng, Emerging Technologies Program Manager; Nate Taylor, Matt Smith, Managers
Host Site Participants – Numerous homeowners
NegaWatt Consulting – Bo White, Project Engineer; Marc Esser, Principal

Disclaimer

Executive Summary

The focus of this case study of non-intrusive load monitoring (NILM) algorithms by four vendors at 11 non-randomly selected homes in the SDG&E region was to evaluate the rough accuracy of each algorithm in an effort to better understand the state of that industry. Whole building electrical power and energy use data was granted to the vendors at four approximate frequencies of 10-seconds, 1-minute, 15-minutes, and 1-hour. There were three data sources for this whole building data: two Rainforest Automation Eagle gateways obtaining high frequency data from the utility electricity smart meter, different only in firmware; and SDG&E Green Button Connect data. The average sampling and recording rate of the Eagle gateways was 10-seconds and the researchers also up-sampled this data to 1-minute and 15-minute intervals. The Green Button data was at 1-hour intervals for most homes and 15-minute intervals at two homes based on the homes' electricity rate schedule.

In addition, zip code location was granted to all vendors at the beginning of the study. Home appliance survey data was granted near the end. The vendors were asked to provide disaggregated predictions based on up to all four frequencies prior to and after receiving the appliance survey data. In respect of their valuable time, they were not required to provide all. Incidentally, most vendors used only the 10-second gateway data and the Green Button data and some only provided pre-appliance-survey predictions. The researchers suggested to the vendors to provide hourly and daily predictions but they were allowed to give predictions at higher or lower frequencies. The researchers reviewed all vendor predictions but focused the analysis on the daily predictions if provided by each vendor for comparison and simplicity's sake. In cases that only monthly predictions were provided, those were analyzed. Abbreviated hourly and minute interval results are provided in section Accuracy Calculations.

The researchers found that various complexities made it difficult to confidently calculate disaggregation accuracy across all vendor predictions and all measured major appliances at each home. First, measuring at only the home breaker-level meant that in numerous cases other plug loads or lighting were in the same data streams as major appliances. Second, some homes had multiple instances of appliances such as refrigerators and only some of those breakers were measured due to measurement and verification (M&V) equipment and installation cost and risk constraints. Third, there were occasional gaps in the whole building meter data and the breaker level data. For all of these reasons, the focus of the accuracy calculations are on those appliances that the researchers found to be on dedicated breakers.

The four NILM vendors were given anonymized names of Vendor A, Vendor B, Vendor C, and Vendor D. Two of these vendors (B & C) provided daily disaggregation predictions based on the 10-second whole building data, both before and after receiving the appliance survey data. Vendor B alone provided "improved" daily predictions after reviewing researcher provided preliminary plots[1]. Vendors C & D provided pre-appliance-survey daily predictions based on the Green Button whole building data. Vendor C also provided post-appliance-survey daily predictions. Vendor A provided only monthly pre-appliance-survey predictions based on all whole building frequencies while Vendor B provided only monthly post-

[1] After reviewing some individual daily plots with the timestamps hidden, the vendor tuned their algorithm. All vendors were given this option but only Vendor B participated.

appliance-survey predictions using the Green Button data. Table 1 shows calculated accuracy "*a*" for all of these daily predictions and Table 2 shows "*a*" for all of the monthly predictions. Please note that Vendor C provided microwave predictions and Vendor D provided air conditioning predictions but the researchers did not find enough true positive instances for analysis. Please also note that there is some overlap in the end use categories since that column is the union of all the categories self-chosen by the vendors. Grey boxes indicate that the vendor intentionally omitted the respective combination of end use category and input dataset. Using ASHRAE 14 (ASHRAE, 2002) as a rough guideline, green boxes indicate good accuracy (*a > 0.7*), yellow is fair (*0.6 < a < 0.7*), and red is poor (*a < 0.6*).

Input Data-Set →	**HAN 10s pre survey**		**HAN 10s post survey**		**GB pre survey**		**GB post survey**	**HAN 10s improved**
Vendor→ ↓End Use	**B**	**C**	**B**	**C**	**C**	**D**	**C**	**B**
EV	0.75		0.88	0.73		0.57	0.57	
Pool Pumps	0.81	0.74	0.74	0.74	0.67	0.66	0.81	0.74
Refrig.	0.72	0.54	0.78	0.54				0.78
HVAC	0.46		0.47					0.63
WH	0.59		0.59					0.67
HVAC & WH		0.26		0.25	0.36		0.36	
Dryer		0.69		0.69		0.69		
Oven		0.26		0.26	0.05		0.05	
Cook & W/D	0.54		0.35					0.49
Solar	0.34							0.83

Table 1: Disaggregation accuracy of daily predictions where provided by vendors

Input Data-Set →	**HAN 10s pre survey**	**HAN 1m pre survey**	**HAN 15m pre survey**	**GB pre survey**	**GB post survey**
Vendor→ ↓End Use	**A**	**A**	**A**	**A**	**B**
Pool Pumps					0.79
Refrig.	0.29	0.43	0.36	0.39	
HVAC	-1.09	-2.35	-2.27	-3.06	0.34
Microwave	0.25	0.3			
W/D	-0.29	-0.12	-0.15		
Dishwasher	-0.54	-0.45	-0.47		

Table 2: Disaggregation accuracy of monthly predictions where provided by vendors

The accuracy metric was calculated exactly as in previous research conducted by EPRI (EPRI, 2013). The accuracy (*a*) equation was $1 - RMSE/\bar{x}$ or equivalently after expanding

$$1 - \frac{1}{\bar{x}}\sqrt{\frac{1}{n_{obs}}\sum_{i=1}^{n_{obs}}(x_i - \hat{x}_i)^2}$$

where x_i are the ground truth measurements, $\bar{x}$ is the mean of x_i, $\hat{x}_i$ are the vendor predictions, and n_{obs} are the number of observations. An accuracy of 1 would be perfect and values below roughly 0.6 (and including negative numbers) are considered poor. Before applying these equations, the researchers first filtered (and in the sole case of refrigerators, processed) the raw 1-minute ground truth data on a home by home and breaker by breaker basis, the exact methods of which will be detailed later. Then, this filtered ground truth data set was grouped over all homes by appliance category and then compared to the respective vendor predictions. The appliance categories shown are the appliance category names chosen by the vendors so they are not always directly comparable to each other. Please note that values near zero were filtered out. False positives and negatives significantly decrease accuracy and the researchers chose not to focus on that issue. However, abbreviated results without the values near zero omitted and including the metric of F-score are given in section Accuracy Calculations.

As expected, the vendors performed their best disaggregation on end-uses with high power magnitude and/or regular runtimes (e.g. electric vehicles, pool pumps, and refrigerators). For daily electric vehicle predictions based on the HAN 10s data, accuracies ranged from 0.73 to 0.88. For pool pumps, accuracies ranged from 0.74 to 0.81 and for refrigeration, accuracies ranged from 0.54 to 0.78. This is valuable to utilities and the respective homeowners because they represent a high percentage of those homes' energy use. The next best categories were dryers and water heating. For dryers, Vendors C and D had accuracies of 0.69. For water heating, Vendor B had an accuracy of 0.67 in their "improved" dataset. While the vendors did poorly on most of the other categories according to the chosen accuracy metric, the reader should keep in mind the sample size, data quality, and other limitations of this study. Furthermore in the body of the report, individual home results will be presented where the vendors do significantly better.

As stated previously, one of the vendors (Vendor A) only provided monthly pre-appliance-survey predictions so they cannot be directly compared to any of the other vendors. Using the chosen accuracy metric, they fared poorly. A major reason could be that they had difficulties dealing with the occasional and sometimes large gaps in the whole building data. Since they only provided monthly predictions, no single days could be omitted. Concerning HVAC in particular, Vendor A's models rely on seasonal variations and were trained with data from areas where temperatures change significantly. So, when used in an area like San Diego, where HVAC isn't operational most of the time, their model over-estimated the portions of total power that is dictated by temperature. According to the vendor, this could easily be remedied by more training data collected from the specific location. In addition and due to time constraints, they only provided results for the first two months of the data collection period and did not give themselves extra time to train their algorithms.

The researchers conclude that in general the vendors have promising NILM algorithms. Their accuracy was found to be very good in some cases. However, accuracy did vary substantially across some homes. The researchers are unsure of the reasons and recommend further study, especially on less energy intensive and complex buildings. There are numerous other studies concurrently underway that would be worthwhile to stay updated upon.

Table of Contents

Document change tracking

Document Date	List of Changes
04/04/2014	Draft outline for limited SDG&E distribution only
05/29/2014	Draft for limited SDG&E distribution and review
06/27/2014	Complete draft for SDG&E, vendor, and peer review
08/22/2014	Incorporated SDG&E and vendor comments

List of Tables

List of Figures

Introduction

Residential NILM is of interest to the California IOUs because it could prove to be a cost effective way to make use of the large data set unleashed by the swift installation of residential smart meters in California. These smart meters already have a protocol in place (SEP) to securely and wirelessly send high frequency demand data to low cost gateways. The IOU authentication process for these gateways is relatively simple and there are a few gateways on the market. However, most of these gateways are currently used simply to display instantaneous whole building demand or simple plots. NILM is an opportunity to gather more useful information from the raw data. The purpose of this study is to simply test the accuracy of the some NILM vendors' algorithms.

Substantial previous research has been completed about residential NILM. The researchers focused their background research on the specific topic of measuring the accuracy of NILM algorithms. EPRI recently completed a large study (EPRI, 2013) that focused heavily on accuracy measurements. While the report could not be obtained for financial reasons, the researchers reviewed slides of their accuracy results and obtained their chosen metrics directly from one of the researchers. Research by Batra and Kelly et al. (Nipun Batra, 2014) focused even more precisely on the issue of metrics. The central focus was to provide a common framework for the NILM industry to measure the accuracy of their algorithms. Metrics from both of these sources were used by the researchers. In addition, research similar to the EPRI research is currently underway at PNNL (RS Butner, 2013). Their accuracy metrics also informed the researchers' work and the researchers collaborated by phone with the authors.

Project Objective

[Exact copy of a similar section in Appendix A: Project Plan]

As noted in Appendix A: Project Plan:

> The goal of this project is to judge the efficacy of several commercially available NILM residential disaggregation technologies. SDG&E would like to understand how accurately these vendors can detect the use of individual appliances. While they may in the future also evaluate the energy saving recommendations the vendors can provide, the focus for now is to only *verify the accuracy of the disaggregation*.

Project Methodology

[Full details can be found in Appendix A: Project Plan and Appendix B: Measurement & Verification Plan]

The Project Plan contains detailed information on the following:

- Technology description
- Description of incumbent technology
- Project Goals
- Technology Application
- Project Milestones

The M&V Plan contains detailed information on the following:

- Test Site Description
- Measurement & Verification Options
- Data Collection / Analysis Procedures
- Calculations of technology algorithm accuracy

Technology Overview

[Updated copy of a similar section in Appendix A: Project Plan]

Multiple vendors of non-intrusive load monitoring (NILM) technologies are studied here. They all disaggregate residential electrical utility smart meter data into constituent large appliance or appliance group energy use without any individually measured appliance data. High frequency (10 second interval on average) smart meter data is obtained using a Zigbee and Internet enabled third party gateway installed in the home. Hourly (or in two cases 15 minute) green button electrical data is also provided to the vendors as supplemental information. Gas usage was completely ignored. All vendors disaggregate to the best of their ability using their own techniques and all focus on appliances that use more than 100 Watts. All vendors' can optionally include energy saving recommendations but the focus of this study is the quality of the disaggregation itself. Most of the vendors prefer to use data at the highest resolution possible and all vendors have the ability to utilize well-structured raw data from any reliable sensing hardware. All have the technical ability to parse and analyze Green Button data but some (Vendors B and C) did not consider it very valuable on its own. While no vendors shared their algorithms with the researchers, Vendors B and C intimated that their algorithms yield much more accurate results with sub 1 minute raw whole building data versus 15 minute or 1 hour data.

Some vendors have the ability to display their results to the occupants using a website, smartphone, and/or tablet app (not tested by the researchers). Others primarily tailor their results directly to the utility, hope the utility will improve their infrastructure to store high frequency data without the use of on-site gateways, and hope to gain access to all residential smart meter data from utility servers.

All vendors consider home survey data to be valuable and can incorporate the information into their algorithms. However, it isn't mandatory for any of their algorithms. Some prefer to also have home addresses in order to obtain demographic and climate information. In this study, the researchers only shared zip codes with the vendors.

See Table 3 for an anonymized comparison of the basic features of the technologies. For more detailed information about NILM strategies currently in use, see (Zoha A, 2012).

Technology feature	A	B	C	D
NILM of residential electricity	X	X	X	X
Energy saving recommendations	X	X	X	
User-facing graphical user interface	X	X	Third party	
Analytics targeted at homeowner	X	X	X	
Analytics targeted at the utility	X	X	X	X

Table 3: Technology Comparison Chart

There is no incumbent disaggregation technology but there is an existing related standard practice: the federally supported Green Button Connect program. This program gives residential utility customers the ability to grant utility authorized third parties automatic access to their smart meter data. The standard practice is for those third parties to simply provide quick access and plotting within a smartphone or tablet app, sometimes augmented by rudimentary analysis or recommendations – the hourly resolution not permitting significant depth. With the purchase and installation of compatible third party hardware (supporting Home Area Network / SEP 1.x), the customer can obtain higher frequency electrical data (up to roughly every 10 seconds) from the smart meter.

Host Site Overview

[Updated copy of a similar section in Appendix B: Measurement & Verification Plan]

As stated in (NegaWatt Consulting, 2013a), "10 homes will be chosen from a group of about 27 homes owned by SDG&E employees". First, surveys were distributed to all 27 employees in order for them to self-report which major electrical appliances they owned. The appliance surveys were then reviewed and 10 homes were chosen with as diverse an appliance pool as possible, without regard to whether the selection is representative of the market (at n=10 this would not be realistic). One additional home with a grid-tied solar photovoltaic system was subsequently added for a total of 11 homes.

Here is a list of the major electrical appliance types that were collectively present at the homes:

- Electric Water Heater
- Pool Pump
- Air conditioners
- Fridge/Freezer Combo
- Freezer
- Beverage fridge
- Dishwasher
- Hot Tub
- Electric Vehicle
- Washer
- Electric Dryer
- Range
- Oven

- Stand-alone Resistance Heater
- Microwave

All of the homes are in the SDG&E region and all have smart meters.

Measurement & Verification Plan Overview

[Updated copy of a similar section in Appendix B: Measurement & Verification Plan]

Two SDG&E Home Area Network (HAN) approved gateways manufactured by Rainforest Automation with model name “Eagle” were installed at each home. One was customized for one of the vendors and the other was generic and connected to the researchers’ server. They were connected to the home’s Internet router and configured to communicate with the smart meter via Zigbee.

The smart meter data was shared as individually gzip compressed xml and csv files, each containing 1 hour of data, at 3 different resolutions: 10 second, 1 minute, and 15 minutes. The 10-second (on average) interval data was the raw instantaneous demand data and the other two resolutions were up-sampled by the authors.

Green Button electrical data (1 hour resolution in most cases) for each home was made available on the same server. The data followed Green Button data standards; the authors received this data from SDG&E and simply passed it on unprocessed.

The vendors were notified that SDG&E preferred results based on *all* available resolutions of 10s, 1m, 15m, and 1h. However, they were given the option to omit any at their discretion in interest of their time, professional judgment, and possible differences in their capabilities. Please note that 10s and 1h data is available for all smart meter customers as of today, with the 10s data requiring the purchase of a HAN gateway. The 15m resolution may be available for customers in the future.

Please note that occasionally there was missing data. The vendors addressed this issue as best they could. Where there were extensive gaps of missing data, vendor results were ignored.

Current transducers were installed on up to 12 circuits at each home. In the case of two homes where there were subpanels, the researchers installed 2 data loggers in each home. The researchers selected the circuits with the highest monthly energy usage. The priority at install time was to identify and include a) *all* circuits with *greater* than 20A capacity, and b) circuits with systems that are used regularly and that have a *non-flat profile* such as appliances (as opposed to flat profile systems such as lighting that were intentionally ignored). The current transducers were connected to a line-powered logger inside the electrical panel (see Appendix B: Measurement & Verification Plan for instrumentation details). The logger data was then made available for online readout with a username and password via power line communication. The logger saved 10 minutes of “volatile” 1 second data and 1 year of 1 minute data. The 1 second data was occasionally referred to as needed but not saved. The 1 minute data was downloaded remotely, stored securely, and used for analysis by the researchers.

Whenever there were multiple loads on one circuit, all of those loads were manually cycled and power data was recorded in 1-second intervals with either a Fluke Power logger, Extech clamp-on meter, or the circuit-level logger, whichever was most appropriate based on site conditions. In the case of refrigerators, this one-time data ("pattern samples") was then used to help the researchers remove anomalous data from those breakers. Data quality level was qualitatively noted for each circuit based on fieldwork conducted by the researchers. For those loads where the breaker was dedicated to the large load being measured, the logged data (per breaker) was used directly as "ground truth". Non-dedicated breakers aside from refrigeration were omitted from the analysis in order to increase the confidence in the results.

The vendors were asked to give two sets of final results, the first without the aid of home appliance surveys and the second with surveys such that the researchers could evaluate the usefulness of survey data. The presumption was that appliance surveys would be difficult to convince homeowners to complete and the researchers were interested in algorithm accuracy with and without it.

Applicable standards

While there is substantial past and concurrent NILM research and while some vendors have expressed interest in M&V standards, there are not any NILM specific codes or standards. However, one related standard is Smart Energy Profile (SEP) by the Zigbee Alliance (Zigbee Alliance, 2013).

The SEP standard is relevant to disaggregation in the California market since the smart meters currently exclusively use this standard to deliver demand data over Zigbee from the smart meter to HAN devices. There are multiple HAN devices readily available in the market that comply with SEP, are approved by the utilities, are relatively inexpensive ($50-$100), provide high-resolution demand data (8-11 second interval), and are user installable. All vendors were comfortable using SEP compliant and utility approved HAN gateways manufactured by a third party.

Project Results and Discussion

Detailed Host System Description

Overview

As stated in the Project Methodology section, 11 homes were outfitted with 2 HAN gateways and a circuit level power logger each. The gateway data and Green Button Connect 1h electricity data was shared with the vendors for their use in their disaggregation algorithms. The vendors were not given access to the homes, the home addresses, or the homeowner names. The researchers installed all of the instrumentation, obtained appliance surveys (with SDG&E assistance), and surveyed each home's appliances on-site. The researchers granted the vendors the appliance survey data near the conclusion of the project for their use in their second and final batch of results.

The researchers installed all of the hardware themselves and the homeowners were not requested to modify their behavior in any way. The researchers outlined the project to the homeowners and occasionally requested them to reboot some of the hardware to improve data collection and remote access.

Monitored Points

The researchers logged whole building real power via the gateways and breaker-level real power via a professional-quality power meter manufactured by eGauge. The power meter also measured or calculated and recorded apparent power at each breaker, voltage and frequency on both legs, and current on each pole of every chosen breaker.

NILM Algorithms

The vendors chose not to disclose details about their respective NILM algorithms for proprietary reasons.

Accuracy Calculations

As stated section Executive Summary, the chosen accuracy metric was $1 - RMSE/\bar{x}$. While that is a mathematically sound metric, there are numerous decisions that had to be made in choosing the data to enter into that equation. The equation requires two time synchronized vectors, one being "ground truth" measurements and the other being vendor predictions, at the same interval. So, one of the first important decisions is interval length. EPRI found in their study that longer prediction intervals led to better accuracy. The researchers choose to focus on daily intervals since it was the most common interval chosen by the vendors and there were an adequate number of observations.[3]

A second issue is that each source of raw data (two gateways and one breaker-level power meter) had its own data quality issue. Each gateway had different poor data quality days making it more difficult to directly compare vendors to each other. In an attempt to address this issue, the vendor used the same

[3] Abbreviated hourly and minute-interval results are also provided later in this section.

filter to omit days of gateway data. At 10-second intervals, the maximum possible points in a day are 8640. The researchers chose 7000 as a reasonable and consistent cut off point for both data sets.

The breaker-level data did not have any significant gaps during the measurement period but it did have a different and very significant data quality issue: the common presence of extraneous end-use loads on the measured breakers. In an effort to address this most simply, the vast majority of breakers with extraneous loads were completely omitted from the analysis. This choice to omit data significantly reduced the size of the data set but allows for more confidence in the results.

The only breakers where the researchers removed loads were the refrigeration breakers. The refrigeration could not be analyzed otherwise because all refrigerators were served by single pole breakers with additional outlets, often including regularly used kitchen outlets. Since refrigeration is a large end use, the researchers decided to clean that ground truth data rather than completely omit the category. To clean the data, the researchers first visually inspected all the refrigeration ground truth plots and omitted homes where there were substantial deviations from a typical refrigerator load curve. Then, the researches omitted very low power values (less than 25 Watts) under the assumption that there is no typical refrigeration energy use state in that range. Then, the data was stepped through in 7 hour windows, minimum energy use was calculated, and this minimum was subtracted from all data points in that time window. This removed any steady and extraneous plug loads. Finally, values above 225 Watts were replaced with interpolated values using the nearest points less than 225 Watts. This removed large magnitude extraneous plug loads like toaster ovens.[4]

Another issue was related to time synchronization. The vendors reported their predictions for different spans of time, making it more difficult to fairly compare them to each other. The researchers chose not to filter the vendor predictions to only include days where they all provided results since it would significantly reduce sample size. However, they did provide some plots of individual home/appliance combinations to portray the issue.

Finally, the researchers chose to remove days of low daily energy in the ground truth and the predictions in an effort to ignore the false positives and false negatives (and subsequently treat them separately by providing abbreviated results with those days included). The limit chosen was 400 Watt-hours. When those data points are not removed, accuracy in general decreases.

Thus far, only one accuracy metric aggregated over all homes has been provided (in the Executive Summary). Numerous additional metrics are shown in the remaining tables below. The first three columns of metrics were calculated exactly as in previous research conducted by EPRI (EPRI, 2013). The formula for the first metric (accuracy a) is provided in the Executive Summary. The second and third metrics $\bar{X}$ and σ are respectively the mean and standard deviation of the errors $x_i - \hat{x}_i$ and are in units of *Watt-hours*. r^2 is the coefficient of determination, *F* is the f-score, $\sum x_i$ is the sum of all the included ground truth measurements, $\sum \hat{x}_i$ is the sum of all the included vendor predictions, *ETE* or error in total

[4] Please note that this also unfortunately removes defrost cycles. However, that daily energy use is much less than daily cooling energy use.

energy assigned is $\sum \hat{x}_i - \sum x_i$, a_T or total accuracy is $|ETE|/\sum x_i$, and *NEP* or normalized error in assigned power is $\frac{1}{\sum x_i}\sum_{i=1}^{n_{obs}} |x_i - \hat{x}_i|$. *ETE, NEP,* and f-score were obtained directly from previous research by Batra and Kelly et al. (Nipun Batra, 2014). Values near one are desirable for the accuracy metrics, coefficient of determination, and f-score. Values near zero are desirable for the error metrics. The highlighting used in select columns of the tables is similar to that used in the Executive Summary where green is best, yellow is fair, and red is poor.

Table 4 shows all of the calculated metrics for the daily predictions provided by the vendors. The calculations were performed after filtering the data and aggregating the data from all included homes. The values near zero in both the ground truth and vendor predictions were ignored. The dataset is identical to that used to produce Table 1 in the Executive Summary, therefore the accuracy *a* column is identical. The dataset used to create Table 5 is identical to that used in Table 4 except that values near zero were not ignored. It is the only table or figure provided in this report where those values were not ignored. Table 6 shows all of the calculated metrics for the monthly predictions and is therefore an expanded version of Table 2 from the Executive Summary.

Table 7 and Table 8 show metrics for the daily predictions itemized by home. Similarly, Table 9 shows itemized results for the monthly predictions albeit a subset for brevity. They are organized similarly to the previous tables except that each row represents an appliance category instance at only one home. The "End Use" column now includes an identifying number for the home and the appliance category names have been abbreviated.

Table 10 shows aggregated hourly interval results and Table 11 shows aggregated minute interval results. These two tables are the only instances of results at those intervals in this report. For the hourly datasets, values less than 17 Watts were ignored and hours with less than 291 HAN gateway data points were ignored. For the minute-interval datasets, values less than 25 Watts were ignored and minutes with less than 4 HAN gateway data points were ignored. While these values are somewhat arbitrary, the researchers found them to be reasonable and more importantly kept them consistent across similar data sets.

Vendor B, Pre-Survey, HAN 10s, Daily, Included Homes Aggregated													
End Use	α	$\bar{X}$	σ	n_{obs}	$\bar{x}$	rmse	*NEP*	r^2	Σx_i	$\Sigma \hat{x}_i$	*ETE*	α_T	*F*
EV	0.75	1372	2607	76	11.58	2.95	0.12	0.77	880	871	-9.7	0.99	1
Pool Pumps	0.81	860	1383	106	8.73	1.63	0.1	0.85	925	845	-80.3	0.91	1
Refrig.	0.72	375	270	89	1.63	0.46	0.23	0.68	145	128	-17.3	0.88	1
HVAC	0.46	1055	754	67	2.4	1.3	0.44	0.54	161	99	-61.7	0.62	1
WH	0.59	3452	1606	26	9.32	3.81	0.37	0.91	242	153	-89.8	0.63	1
Cook & W/D	0.54	889	827	45	2.66	1.21	0.33	0.75	120	91	-28.7	0.76	1
Solar	0.34	4713	2953	33	8.47	5.56	0.56	0.33	280	415	135.5	0.52	1
Vendor C, Pre-Survey, HAN 10s, Daily, Included Homes Aggregated													
PoolPump	0.74	1186	1622	141	7.73	2.01	0.15	0.68	1090	952	-138.3	0.87	1
Refrig.	0.54	458	489	168	1.46	0.67	0.31	0.02	244	168	-76.2	0.69	1
HVAC & WH	0.26	3658	3334	81	6.7	4.95	0.55	0.61	543	267	-276.4	0.49	1
Dryer	0.69	2168	1553	23	8.54	2.67	0.25	0.99	196	146	-49.9	0.75	1
Oven	0.26	1187	866	37	1.98	1.47	0.6	0.04	73	45	-27.9	0.62	1
Vendor B, Post-Survey, HAN 10s, Daily, Included Homes Aggregated													
EV	0.88	789	1236	53	12.11	1.47	0.07	0.93	642	615	-27.4	0.96	1
Pool Pumps	0.74	991	1645	267	7.43	1.92	0.13	0.73	1983	1755	-227.7	0.89	1
Refrig.	0.78	186	205	246	1.28	0.28	0.15	0.11	316	329	12.6	0.96	1
HVAC	0.47	921	688	151	2.19	1.15	0.42	0.3	330	230	-100.3	0.7	1
WH	0.59	3277	1580	70	8.95	3.64	0.37	0.92	627	397	-229.4	0.63	1
Cook & W/D	0.35	994	1169	111	2.35	1.53	0.42	0.68	261	175	-85.4	0.67	1
Vendor C, Post-Survey, HAN 10s, Daily, Included Homes Aggregated													
EV	0.73	2242	2409	38	12.38	3.29	0.18	0.77	470	433	-37.6	0.92	1
Pool Pumps	0.74	1186	1622	141	7.73	2.01	0.15	0.68	1090	952	-138.3	0.87	1
Refrig.	0.54	458	489	168	1.46	0.67	0.31	0.02	244	168	-76.2	0.69	1
HVAC & WH	0.25	3319	3005	106	5.94	4.48	0.56	0.63	630	298	-331.9	0.47	1
Dryer	0.69	2168	1553	23	8.54	2.67	0.25	0.99	196	146	-49.9	0.75	1
Oven	0.26	1187	866	37	1.98	1.47	0.6	0.04	73	45	-27.9	0.62	1
Vendor C, Pre-Survey, GB, Daily, Included Homes Aggregated													
Pool Pumps	0.67	1871	2086	107	8.47	2.8	0.22	0.16	906	940	34.2	0.96	1
HVAC & WH	0.36	2241	2144	124	4.88	3.1	0.46	0.65	605	668	63.4	0.9	1
Oven	0.05	1004	721	42	1.3	1.24	0.78	0.04	54	89	34.9	0.36	1
Vendor D, Pre-Survey, GB, Daily, Included Homes Aggregated													
EV	0.57	3125	3357	106	10.58	4.59	0.3	0.55	1121	1145	23.7	0.98	1
Pool Pumps	0.66	1928	1872	230	7.86	2.69	0.25	0.46	1807	1627	-179.6	0.9	1
Dryer	0.69	2368	2269	17	10.54	3.28	0.22	0.76	179	167	-12.1	0.93	1
Vendor C, Post-Survey, GB, Daily, Included Homes Aggregated													
EV	0.57	4229	3128	26	12.15	5.26	0.35	0.77	316	206	-109.5	0.65	1
Pool Pumps	0.81	1115	1425	81	9.71	1.81	0.11	0.37	786	734	-52.7	0.93	1
HVAC & WH	0.36	2241	2144	124	4.88	3.1	0.46	0.65	605	668	63.4	0.9	1
Oven	0.05	1004	721	42	1.3	1.24	0.78	0.04	54	89	34.9	0.36	1
Vendor B, Post-Survey, HAN 10s Improved, Daily, Included Homes Aggregated													
Pool Pumps	0.74	991	1645	267	7.43	1.92	0.13	0.73	1983	1755	-227.7	0.89	1
Refrig.	0.78	186	205	246	1.28	0.28	0.15	0.11	316	329	12.6	0.96	1
HVAC	0.63	639	507	118	2.23	0.82	0.29	0.64	264	211	-52.5	0.8	1
WH	0.67	2429	1571	71	8.86	2.89	0.27	0.91	629	463	-165.8	0.74	1
Cook & W/D	0.49	660	979	108	2.32	1.18	0.28	0.83	251	200	-51.4	0.8	1
Solar	0.83	1427	936	121	9.9	1.71	0.14	0.83	1198	1035	-163.4	0.86	1

Table 4: All metrics for aggregated daily predictions with values near zero ignored

End Use	a	$\bar{X}$	σ	n_{obs}	$\bar{x}$	rmse	NEP	r^2	Σx_i	$\Sigma \hat{x}_i$	ETE	a_T	F
Vendor B, Pre-Survey, HAN 10s, Daily, Included Homes Aggregated													
EV	0.65	1450	2951	98	9.37	3.29	0.15	0.79	918	871	-47.6	0.95	0.87
Pool Pumps	0.6	1959	2252	151	7.48	2.99	0.26	0.77	1130	845	-285	0.75	0.82
Refrig.	0.72	375	270	89	1.63	0.46	0.23	0.68	145	128	-17.3	0.88	1
HVAC	0.12	1000	801	221	1.45	1.28	0.69	0.45	321	123	-198.2	0.38	0.63
WH	0.43	2932	1866	48	6.11	3.48	0.48	0.89	293	153	-140.7	0.52	0.7
Cook & W/D	0.07	531	674	172	0.93	0.86	0.57	0.69	159	131	-28.2	0.82	0.85
Solar	0.34	4713	2953	33	8.47	5.56	0.56	0.33	280	415	135.5	0.52	1
Vendor C, Pre-Survey, HAN 10s, Daily, Included Homes Aggregated													
PoolPump	0.55	2110	2528	178	7.3	3.29	0.29	0.48	1299	952	-346.7	0.73	0.89
Refrig.	0.54	458	489	168	1.46	0.67	0.31	0.02	244	168	-76.2	0.69	1
HVAC & WH	-0.04	2334	1958	433	2.92	3.05	0.8	0.64	1263	273	-990.5	0.22	0.46
Dryer	0.47	744	1363	68	2.9	1.55	0.26	0.99	197	147	-50.6	0.74	0.98
Oven	-1.1	624	1184	406	0.64	1.34	0.98	0.09	259	156	-103.1	0.6	0.56
Microwave	-0.05	130	123	499	0.17	0.18	0.76	0.11	85	33	-51.5	0.39	0.6
Vendor B, Post-Survey, HAN 10s, Daily, Included Homes Aggregated													
EV	0.86	710	1103	68	9.53	1.31	0.07	0.97	648	615	-33.8	0.95	0.88
Pool Pumps	0.74	988	1636	270	7.35	1.91	0.13	0.75	1985	1755	-229.8	0.88	0.99
Refrig.	0.78	186	205	246	1.28	0.28	0.15	0.11	316	329	12.6	0.96	1
HVAC	0.01	921	819	450	1.24	1.23	0.74	0.2	560	322	-237.7	0.58	0.7
WH	0.4	2964	2203	126	6.12	3.69	0.48	0.85	771	397	-373.5	0.52	0.72
Cook & W/D	-0.02	593	871	332	1.04	1.05	0.57	0.67	344	217	-126.8	0.63	0.81
Vendor C, Post-Survey, HAN 10s, Daily, Included Homes Aggregated													
EV	0.14	5518	6077	96	9.49	8.21	0.58	0.28	911	438	-472.8	0.48	0.6
PoolPump	0.55	2110	2528	178	7.3	3.29	0.29	0.48	1299	952	-346.7	0.73	0.89
Refrig.	0.54	458	489	168	1.46	0.67	0.31	0.02	244	168	-76.2	0.69	1
HVAC & WH	-0.01	2226	1930	433	2.92	2.95	0.76	0.66	1263	319	-944.1	0.25	0.64
Dryer	0.47	744	1363	68	2.9	1.55	0.26	0.99	197	147	-50.6	0.74	0.98
Oven	-1.1	624	1184	406	0.64	1.34	0.98	0.09	259	156	-103.1	0.6	0.56
Microwave	-0.05	130	123	499	0.17	0.18	0.76	0.11	85	33	-51.5	0.39	0.6
Vendor C, Pre-Survey, GB, Daily, Included Homes Aggregated													
Pool Pumps	0.65	1988	2083	115	8.12	2.88	0.24	0.28	934	940	5.8	0.99	0.96
HVAC & WH	0.06	2162	1668	433	2.92	2.73	0.74	0.54	1263	668	-595.1	0.53	0.45
Oven	-1	779	890	343	0.59	1.18	1.32	0.02	202	194	-8.5	0.96	0.44
Vendor D, Pre-Survey, GB, Daily, Included Homes Aggregated													
EV	0.49	3290	3348	123	9.27	4.69	0.35	0.58	1140	1209	69.1	0.94	0.99
Pool Pumps	0.66	1923	1870	231	7.83	2.68	0.25	0.47	1808	1627	-180.4	0.9	1
Dryer	0.29	900	1787	72	2.83	2	0.32	0.86	204	167	-36.7	0.82	0.79
AC	-0.25	1605	1195	538	1.6	2	1	0	862	5	-856.7	0.01	0.01
Vendor C, Post-Survey, GB, Daily, Included Homes Aggregated													
EV	0.48	3819	3552	33	10.06	5.22	0.38	0.74	332	206	-125.6	0.62	0.88
Pool Pumps	0.63	2069	2155	115	8.12	2.99	0.25	0.67	934	734	-200.4	0.79	0.83
HVAC & WH	0.06	2162	1668	433	2.92	2.73	0.74	0.54	1263	668	-595.1	0.53	0.45
Oven	-1	779	890	343	0.59	1.18	1.32	0.02	202	194	-8.5	0.96	0.44
Vendor B, Post-Survey, HAN 10s Improved, Daily, Included Homes Aggregated													
Pool Pumps	0.74	988	1636	270	7.35	1.91	0.13	0.75	1985	1755	-229.8	0.88	0.99
Refrig.	0.78	186	205	246	1.28	0.28	0.15	0.11	316	329	12.6	0.96	1
HVAC	-0.02	944	850	450	1.24	1.27	0.76	0.21	560	296	-264	0.53	0.58
WH	0.46	2495	2177	126	6.12	3.31	0.41	0.84	771	463	-307.8	0.6	0.72
Cook & W/D	0.11	511	769	332	1.04	0.92	0.49	0.76	344	224	-120.2	0.65	0.63
Solar	0.82	1455	981	122	9.86	1.75	0.15	0.83	1203	1035	-168.2	0.86	1

Table 5: All metrics for aggregated daily predictions with values near zero included

End Use	a	$\bar{X}$	σ	n_{obs}	$\bar{x}$	rmse	*NEP*	r^2	Σx_i	$\Sigma \hat{x}_i$	*ETE*	a_T
Vendor A, Pre-Survey, HAN 10s, Monthly, Included Homes Aggregated												
Refrig.	0.29	22776	20897	6	43.5	30.91	0.52	0.11	261	124	-136.6	0.48
HVAC	-1.09	77786	74161	10	51.43	107.5	1.51	0.05	514	1181	667	-0.3
Microwave	0.32	3029	1952	12	5.27	3.6	0.58	0.13	63	33	-29.8	0.53
W/D	-0.25	19886	16745	10	20.79	26	0.96	0.09	208	380	172.3	0.17
Dish.	-0.54	8677	7864	10	7.61	11.71	1.14	0.07	76	148	71.4	0.06
Vendor A, Pre-Survey, HAN 1-minute, Monthly, Included Homes Aggregated												
Refrig.	0.43	21041	13118	6	43.5	24.8	0.48	0.05	261	174	-87.4	0.67
HVAC	-2.35	2E+05	73990	10	51.43	172.1	3.02	0.09	514	2069	1554	-2.02
Microwave	0.36	3296	2363	8	6.31	4.06	0.52	0.02	50	27	-23.2	0.54
W/D	-0.12	19872	9970	12	19.86	22.23	1	0.17	238	476	237.8	0
Dish.	-0.45	9572	5525	10	7.61	11.05	1.26	0.05	76	168	91.7	-0.21
Vendor A, Pre-Survey, HAN 15-minute, Monthly, Included Homes Aggregated												
Refrig.	0.36	23337	15268	6	43.5	27.89	0.54	0.18	261	121	-140	0.46
HVAC	-2.27	1E+05	82574	10	51.43	168.4	2.85	0.03	514	1982	1468	-1.85
W/D	-0.15	20434	10047	12	19.86	22.77	1.03	0.17	238	483	245.2	-0.03
Dish.	-0.47	9682	5658	10	7.61	11.21	1.27	0.07	76	171	94.9	-0.25
Vendor A, Pre-Survey, GB, Monthly, Included Homes Aggregated												
Refrig.	0.39	22831	13525	6	43.5	26.54	0.52	0	261	382	120.5	0.54
HVAC	-3.76	2E+05	92657	8	43.89	208.9	4.27	0.14	351	1849	1498	-3.27
Vendor B, Post-Survey, GB, Monthly, Included Homes Aggregated												
Pool Pumps	0.79	37591	26853	9	217	46.2	0.17	0.81	1953	1683	-269.5	0.86
HVAC	0.34	25737	19908	19	49.4	32.54	0.52	0.16	939	712	-226.9	0.76

Table 6: All metrics for aggregated monthly predictions

Vendor B, Pre-Survey, HAN 10s, Daily, Itemized Per Home													
End Use	a	$\overline{X}$	σ	n_{obs}	$\overline{x}$	rmse	*NEP*	r^2	Σx_i	$\Sigma \hat{x}_i$	*ETE*	a_T	*F*
Cook-W-D 3	0.63	1252	1120	13	4.55	1.68	0.27	0.83	59	43	-16.3	0.73	1
Cook-W-D 7	0.52	928	585	18	2.28	1.1	0.41	0.48	41	34	-7.1	0.83	1
Cook-W-D 8	0.45	500	562	14	1.38	0.75	0.36	0	19	14	-5.4	0.72	1
EV 5	0.75	1372	2607	76	11.58	2.95	0.12	0.77	880	871	-9.7	0.99	1
HVAC 3	0.51	794	591	44	2	0.99	0.4	0.49	88	62	-26.2	0.7	1
HVAC 6	0.45	1554	780	23	3.16	1.74	0.49	0.69	73	37	-35.5	0.51	1
Pump 5	0.66	1116	1705	53	5.94	2.04	0.19	0.33	315	260	-55.1	0.82	1
Pump 8	0.91	604	886	53	11.52	1.07	0.05	0.85	611	585	-25.2	0.96	1
Refrig. 4	0.68	289	306	22	1.32	0.42	0.22	0	29	35	5.5	0.81	1
Refrig. 7	0.65	423	196	48	1.32	0.47	0.32	0.04	63	43	-20.3	0.68	1
Refrig. 9	0.82	355	350	19	2.76	0.5	0.13	0.19	52	50	-2.5	0.95	1
Solar 11	0.34	4713	2953	33	8.47	5.56	0.56	0.33	280	415	135.5	0.52	1
WH 7	0.59	3452	1606	26	9.32	3.81	0.37	0.91	242	153	-89.8	0.63	1
Vendor C, Pre-Survey, HAN 10s, Daily, Itemized Per Home													
Dryer 10	0.69	2168	1553	23	8.54	2.67	0.25	0.99	196	146	-49.9	0.75	1
HVAC-WH 3	0.25	1214	947	28	2.07	1.54	0.59	0.17	58	42	-16.1	0.72	1
HVAC-WH 7	0.34	4949	3420	53	9.16	6.02	0.54	0.62	485	225	-260.3	0.46	1
Oven 1	0.15	1366	1008	22	1.99	1.7	0.69	0.01	44	30	-14.1	0.68	1
Oven 7	0.47	923	494	15	1.97	1.05	0.47	0.52	30	16	-13.8	0.53	1
Pump 5	0.58	1328	1702	62	5.17	2.16	0.26	0.2	321	265	-55.6	0.83	1
Pump 8	0.81	1075	1548	79	9.74	1.88	0.11	0.45	769	687	-82.6	0.89	1
Refrig. 4	0.59	518	138	40	1.31	0.54	0.4	0.03	52	32	-20.7	0.6	1
Refrig. 7	0.77	248	177	82	1.3	0.3	0.19	0.1	106	87	-19.6	0.82	1
Refrig. 9	0.41	781	784	46	1.86	1.11	0.42	0.05	86	50	-35.9	0.58	1
Vendor B, Post-Survey, HAN 10s, Daily, Itemized Per Home													
Cook-W-D 3	0.59	1472	1590	20	5.24	2.17	0.28	0.85	105	78	-27.2	0.74	1
Cook-W-D 7	0.32	1263	939	41	2.3	1.57	0.55	0.23	94	51	-43.6	0.54	1
Cook-W-D 8	0.07	583	990	50	1.23	1.15	0.47	0	61	47	-14.7	0.76	1
EV 5	0.88	789	1236	53	12.11	1.47	0.07	0.93	642	615	-27.4	0.96	1
HVAC 3	0.61	567	438	51	1.85	0.72	0.31	0.19	94	81	-13.3	0.86	1
HVAC 6	0.44	1307	684	72	2.65	1.47	0.49	0.28	191	113	-77.3	0.59	1
HVAC 8	0.52	576	518	28	1.62	0.77	0.36	0.67	45	36	-9.7	0.79	1
Pump 5	0.56	1145	1884	135	5.03	2.21	0.23	0.16	679	542	-137	0.8	1
Pump 8	0.84	834	1339	132	9.88	1.58	0.08	0.65	1304	1214	-90.8	0.93	1
Refrig. 4	0.67	331	302	63	1.37	0.45	0.24	0.07	86	104	18	0.79	1
Refrig. 7	0.84	152	131	126	1.24	0.2	0.12	0.13	156	148	-8.2	0.95	1
Refrig. 9	0.89	103	97	57	1.29	0.14	0.08	0	74	77	2.9	0.96	1
WH 7	0.59	3277	1580	70	8.95	3.64	0.37	0.92	627	397	-229.4	0.63	1

Table 7: All metrics for itemized daily predictions with values near zero ignored (1 of 2)

Vendor C, Post-Survey, HAN 10s, Daily, Itemized Per Home													
End Use	α	$\overline{X}$	σ	n_{obs}	$\overline{x}$	rmse	*NEP*	r^2	Σx_i	$\Sigma \hat{x}_i$	*ETE*	α_T	*F*
Dryer 10	0.69	2168	1553	23	8.54	2.67	0.25	0.99	196	146	-49.9	0.75	1
EV 5	0.73	2242	2409	38	12.38	3.29	0.18	0.77	470	433	-37.6	0.92	1
HVAC-WH 3	0.25	1214	947	28	2.07	1.54	0.59	0.17	58	42	-16.1	0.72	1
HVAC-WH 6	0.32	2223	837	25	3.48	2.38	0.64	0.49	87	31	-55.6	0.36	1
HVAC-WH 7	0.34	4949	3420	53	9.16	6.02	0.54	0.62	485	225	-260.3	0.46	1
Oven 1	0.15	1366	1008	22	1.99	1.7	0.69	0.01	44	30	-14.1	0.68	1
Oven 7	0.47	923	494	15	1.97	1.05	0.47	0.52	30	16	-13.8	0.53	1
Pump 5	0.58	1328	1702	62	5.17	2.16	0.26	0.2	321	265	-55.6	0.83	1
Pump 8	0.81	1075	1548	79	9.74	1.88	0.11	0.45	769	687	-82.6	0.89	1
Refrig. 4	0.59	518	138	40	1.31	0.54	0.4	0.03	52	32	-20.7	0.6	1
Refrig. 7	0.77	248	177	82	1.3	0.3	0.19	0.1	106	87	-19.6	0.82	1
Refrig. 9	0.41	781	784	46	1.86	1.11	0.42	0.05	86	50	-35.9	0.58	1
Vendor C, Pre-Survey, GB, Daily, Itemized Per Home													
HVAC-WH 3	0.17	1409	731	55	1.91	1.59	0.74	0.1	105	167	61.8	0.41	1
HVAC-WH 7	0.46	2903	2616	69	7.24	3.91	0.4	0.56	500	501	1.5	1	1
Oven 1	0.04	1246	885	19	1.6	1.53	0.78	0	30	47	16.7	0.45	1
Oven 8	0.11	805	463	23	1.04	0.93	0.77	0.06	24	42	18.2	0.24	1
Pump 5	-0.02	4229	2059	26	4.59	4.7	0.92	0.08	119	206	86.9	0.27	1
Pump 8	0.81	1115	1425	81	9.71	1.81	0.11	0.37	786	734	-52.7	0.93	1
Vendor D, Pre-Survey, GB, Daily, Itemized Per Home													
Dryer 10	0.69	2368	2269	17	10.54	3.28	0.22	0.76	179	167	-12.1	0.93	1
EV 5	0.57	3125	3357	106	10.58	4.59	0.3	0.55	1121	1145	23.7	0.98	1
Pump 5	0.6	1734	1447	123	5.6	2.26	0.31	0.01	688	660	-28.5	0.96	1
Pump 8	0.7	2151	2244	107	10.45	3.11	0.21	0.94	1118	967	-151.2	0.86	1
Vendor C, Post-Survey, GB, Daily, Itemized Per Home													
EV 5	0.57	4229	3128	26	12.15	5.26	0.35	0.77	316	206	-109.5	0.65	1
HVAC-WH 3	0.17	1409	731	55	1.91	1.59	0.74	0.1	105	167	61.8	0.41	1
HVAC-WH 7	0.46	2903	2616	69	7.24	3.91	0.4	0.56	500	501	1.5	1	1
Oven 1	0.04	1246	885	19	1.6	1.53	0.78	0	30	47	16.7	0.45	1
Oven 8	0.11	805	463	23	1.04	0.93	0.77	0.06	24	42	18.2	0.24	1
Pump 8	0.81	1115	1425	81	9.71	1.81	0.11	0.37	786	734	-52.7	0.93	1
Vendor B, Post-Survey, HAN 10s Improved, Daily, Itemized Per Home													
Cook-W-D 3	0.59	1472	1590	20	5.24	2.17	0.28	0.85	105	78	-27.2	0.74	1
Cook-W-D 7	0.51	779	850	36	2.34	1.15	0.33	0.31	84	62	-21.8	0.74	1
Cook-W-D 8	0.66	266	314	52	1.19	0.41	0.22	0.2	62	60	-2.4	0.96	1
HVAC 3	0.61	567	438	51	1.85	0.72	0.31	0.19	94	81	-13.3	0.86	1
HVAC 6	0.66	855	584	41	3.05	1.04	0.28	0.63	125	95	-30.4	0.76	1
HVAC 8	0.67	441	356	26	1.69	0.57	0.26	0.86	44	35	-8.7	0.8	1
Pump 5	0.56	1145	1884	135	5.03	2.21	0.23	0.16	679	542	-137	0.8	1
Pump 8	0.84	834	1339	132	9.88	1.58	0.08	0.65	1304	1214	-90.8	0.93	1
Refrig. 4	0.67	331	302	63	1.37	0.45	0.24	0.07	86	104	18	0.79	1
Refrig. 7	0.84	152	131	126	1.24	0.2	0.12	0.13	156	148	-8.2	0.95	1
Refrig. 9	0.89	103	97	57	1.29	0.14	0.08	0	74	77	2.9	0.96	1
Solar 11	0.83	1427	936	121	9.9	1.71	0.14	0.83	1198	1035	-163.4	0.86	1
WH 7	0.67	2429	1571	71	8.86	2.89	0.27	0.91	629	463	-165.8	0.74	1

Table 8: All metrics for itemized daily predictions with values near zero ignored (2 of 2)

Vendor A, Pre-Survey, HAN 10s, Monthly, Itemized Per Home												
End Use	a	$\overline{X}$	σ	n_{obs}	$\overline{x}$	rmse	*NEP*	r^2	Σx_i	$\Sigma \hat{x}_i$	*ETE*	a_T
Dish 3	-1.87	10917	4762	2	4.16	11.91	2.63	1	8	30	21.8	-1.63
Dish 5	-0.87	16330	11531	2	10.71	19.99	1.52	1	21	54	32.7	-0.52
Dish 7	0.48	4662	2378	2	10.12	5.23	0.46	1	20	25	4.8	0.77
Dish 8	-0.54	7451	6689	2	6.5	10.01	1.15	1	13	26	13.4	-0.03
Dish 9	0.38	4025	589	2	6.54	4.07	0.62	1	13	12	-1.2	0.91
HVAC 2	0.61	31278	4570	2	81.58	31.61	0.38	1	163	172	9.1	0.94
HVAC 3	-1.26	1E+05	77149	2	56.6	128.1	1.81	1	113	318	204.4	-0.81
HVAC 6	0.29	48663	19950	2	73.85	52.59	0.66	1	148	188	39.9	0.73
HVAC 7	-17.16	1E+05	1E+05	2	9.65	175.2	14.05	1	19	290	271	-13.05
HVAC 8	-1.34	71287	42645	2	35.46	83.07	2.01	1	71	214	142.6	-1.01
Micro 2	0.48	3334	272	2	6.43	3.35	0.52	1	13	6	-6.7	0.48
Micro 3	0.58	881	14	2	2.12	0.88	0.42	1	4	6	1.8	0.58
Micro 4	0.25	5919	1769	2	8.22	6.18	0.72	1	16	5	-11.8	0.28
Micro 5	0.45	4600	578	2	8.45	4.64	0.54	1	17	8	-9.2	0.46
Micro 7	n/a	4000	0	1	0	4	inf	n/a	0	4	4	n/a
Micro 8	0.59	1640	527	2	4.24	1.72	0.39	1	8	5	-3.3	0.61
Micro 9	0.15	1804	299	2	2.15	1.83	0.84	1	4	4	-0.6	0.86
Refrig. 4	0.43	20322	11110	2	40.37	23.16	0.5	1	81	40	-40.6	0.5
Refrig. 7	0.45	17876	13987	2	41.13	22.7	0.43	1	82	47	-35.8	0.57
Refrig. 9	0.13	30131	30115	2	49.01	42.6	0.61	1	98	38	-60.2	0.39
W-D 3	0.74	8977	4287	2	38.34	9.95	0.23	1	77	68	-8.6	0.89
W-D 5	-0.59	34368	22737	2	25.89	41.21	1.33	1	52	121	68.7	-0.33
W-D 6	0.2	15325	8768	2	21.98	17.66	0.7	1	44	75	30.6	0.3
W-D 7	-0.73	16924	10388	2	11.48	19.86	1.47	1	23	57	33.8	-0.47
W-D 8	-3.73	23838	17549	2	6.26	29.6	3.81	1	13	60	47.7	-2.81
Vendor B, Post-Survey, GB, Monthly, Itemized Per Home												
HVAC 2	0.58	33992	6426	4	81.56	34.59	0.42	0.82	326	190	-136	0.58
HVAC 3	0.88	7121	730	3	61.11	7.16	0.12	0.95	183	192	8.6	0.95
HVAC 6	0.24	53162	11314	4	71.25	54.35	0.75	0.76	285	72	-212.6	0.25
HVAC 7	-2.68	17164	19093	4	6.97	25.67	2.46	0.92	28	91	62.7	-1.25
HVAC 8	0.54	12592	4679	4	29.01	13.43	0.43	0.99	116	166	50.4	0.57
Pump 5	0.7	34720	32715	5	160.8	47.7	0.22	0	804	664	-140.2	0.83
Pump 8	0.85	41180	16169	4	287.2	44.24	0.14	0.95	1149	1019	-129.3	0.89

Table 9: All metrics for select vendor monthly predictions itemized by home

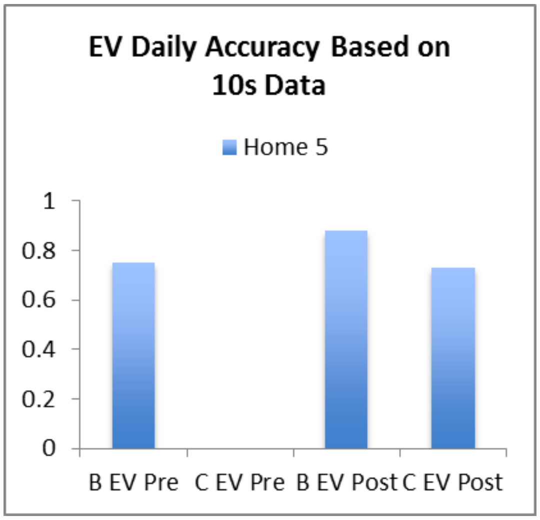
Vendor C, Pre-Survey, HAN 10s, Hourly, Included Homes Aggregated													
End Use	*a*	$\bar{X}$	σ	n_{obs}	$\bar{x}$	rmse	*NEP*	r^2	Σx_i	$\Sigma \hat{x}_i$	*ETE*	a_T	*F*
Pool Pumps	-0.06	815	877	444	1.13	1.2	0.72	0	504	799	295.1	0.41	1
Refrig.	0.38	29	27	4423	0.06	0.04	0.45	0	283	184	-98.9	0.65	1
HVAC-WH	-1.77	528	764	422	0.34	0.93	1.57	0.2	141	259	117.5	0.17	1
Dryer	0.27	928	603	73	1.51	1.11	0.62	0.01	110	95	-15.5	0.86	1
Oven	-0.3	541	537	40	0.59	0.76	0.92	0.02	23	13	-10.5	0.55	1
Microwave	-0.09	49	59	45	0.07	0.08	0.7	0	3	3	-0.6	0.8	1
Vendor C, Post-Survey, HAN 10s, Hourly, Included Homes Aggregated													
EV	0.17	1688	1469	86	2.7	2.24	0.63	0.02	232	229	-3.6	0.98	1
Pool Pumps	-0.06	815	877	444	1.13	1.2	0.72	0	504	799	295.1	0.41	1
Refrig.	0.38	29	27	4423	0.06	0.04	0.45	0	283	184	-98.9	0.65	1
HVAC-WH	-1.73	362	625	712	0.26	0.72	1.37	0.21	188	306	117.7	0.37	1
Dryer	0.27	928	603	73	1.51	1.11	0.62	0.01	110	95	-15.5	0.86	1
Oven	-0.3	541	537	40	0.59	0.76	0.92	0.02	23	13	-10.5	0.55	1
Microwave	-0.09	49	59	45	0.07	0.08	0.7	0	3	3	-0.6	0.8	1
Vendor C, Pre-Survey, GB, Hourly, Included Homes Aggregated													
Pool Pumps	0.85	145	214	403	1.74	0.26	0.08	0.88	700	734	33.4	0.95	1
HVAC-WH	-0.52	709	529	592	0.58	0.88	1.22	0.51	344	623	279	0.19	1
Oven	-0.54	615	288	59	0.44	0.68	1.39	0.09	26	61	35.3	-0.35	1
Vendor D, Pre-Survey, GB, Hourly, Included Homes Aggregated													
EV	0.61	1081	671	290	3.23	1.27	0.33	0.21	937	1037	99.1	0.89	1
Pool Pumps	0.8	276	172	850	1.63	0.33	0.17	0.69	1382	1452	70	0.95	1
Dryer	0.66	551	536	61	2.24	0.77	0.25	0.02	137	167	30.2	0.78	1
Vendor C, Post-Survey, GB, Hourly, Included Homes Aggregated													
EV	0.82	653	165	69	3.71	0.67	0.18	0.06	256	216	-40	0.84	1
Pool Pumps	0.85	145	214	403	1.74	0.26	0.08	0.88	700	734	33.4	0.95	1
HVAC-WH	-0.52	709	529	592	0.58	0.88	1.22	0.51	344	623	279	0.19	1
Oven	-0.54	615	288	59	0.44	0.68	1.39	0.09	26	61	35.3	-0.35	1

Table 10: All metrics for aggregated hourly predictions with values near zero ignored

Vendor C, Pre-Survey, HAN 10s, Minute Interval, Included Homes Aggregated													
End Use	*a*	$\bar{X}$	σ	n_{obs}	$\bar{x}$	rmse	*NEP*	r^2	Σx_i	$\Sigma \hat{x}_i$	*ETE*	a_T	*F*
Pool Pumps	0.91	61	154	29343	1.87	0.17	0.03	0.71	54906	54726	-180.3	1	1
Refrig.	0.78	21	19	69332	0.13	0.03	0.17	0.41	8814	7899	-914.4	0.9	1
HVAC-WH	0.68	537	991	4346	3.47	1.13	0.15	0.85	15095	13733	-1362	0.91	1
Dryer	0.69	771	480	3284	2.9	0.91	0.27	0.66	9511	8528	-983	0.9	1
Oven	0.38	727	637	2553	1.55	0.97	0.47	0.24	3955	3324	-631.4	0.84	1
Microwave	0.44	423	392	987	1.03	0.58	0.41	0.22	1014	873	-141.4	0.86	1
Vendor C, Post-Survey, HAN 10s, Minute Interval, Included Homes Aggregated													
EV	0.94	184	113	6332	3.8	0.22	0.05	0.19	24048	23071	-977.5	0.96	1
Pool Pumps	0.91	61	154	29343	1.87	0.17	0.03	0.71	54906	54726	-180.3	1	1
Refrig.	0.78	21	19	69332	0.13	0.03	0.17	0.41	8814	7899	-914.4	0.9	1
HVAC-WH	0.58	372	725	8962	1.95	0.81	0.19	0.89	17457	16448	-1010	0.94	1
Dryer	0.69	771	480	3284	2.9	0.91	0.27	0.66	9511	8528	-983	0.9	1
Oven	0.38	727	637	2553	1.55	0.97	0.47	0.24	3955	3324	-631.4	0.84	1
Microwave	0.44	423	392	987	1.03	0.58	0.41	0.22	1014	873	-141.4	0.86	1

Table 11: All metrics for aggregated minute-interval predictions with values near zero ignored

Now, to review some select individual home daily-interval accuracy results, Figure 1 and Figure 2 compare the daily accuracy of Vendors B and C for those three appliance categories where they overlap: electric vehicles, pool pumps, and refrigeration. In general, they do similarly well on the same homes and each do very well on at least one home per category.



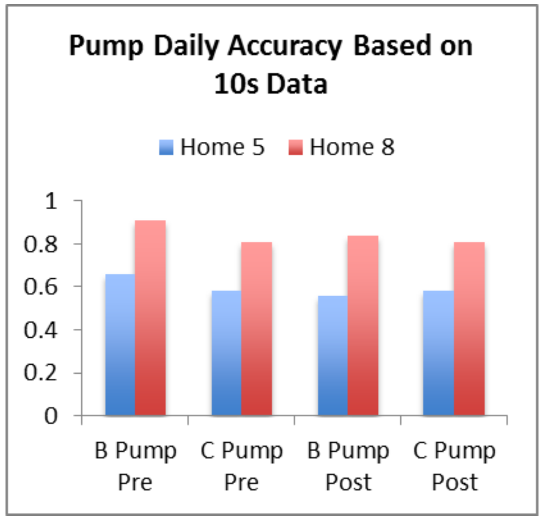


Figure 1: Vendor B and C daily accuracy per home for EV and pool pumps

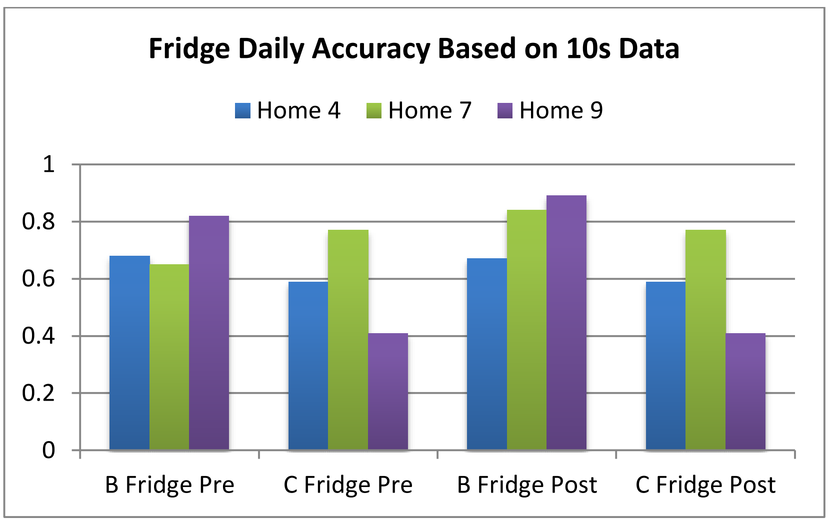


Figure 2: Vendor B and C daily accuracy per home for refrigeration

Figure 3, Figure 4, Figure 5, and Figure 6 show one timeline plot for each vendor for the electric vehicle category and the pool pump category. They each perform extremely well, even detecting a change in the regular operation of the pool pumps in late November. Please note that they each miss the spike in energy use in early December. However, this is likely inconsequential for most utility purposes.

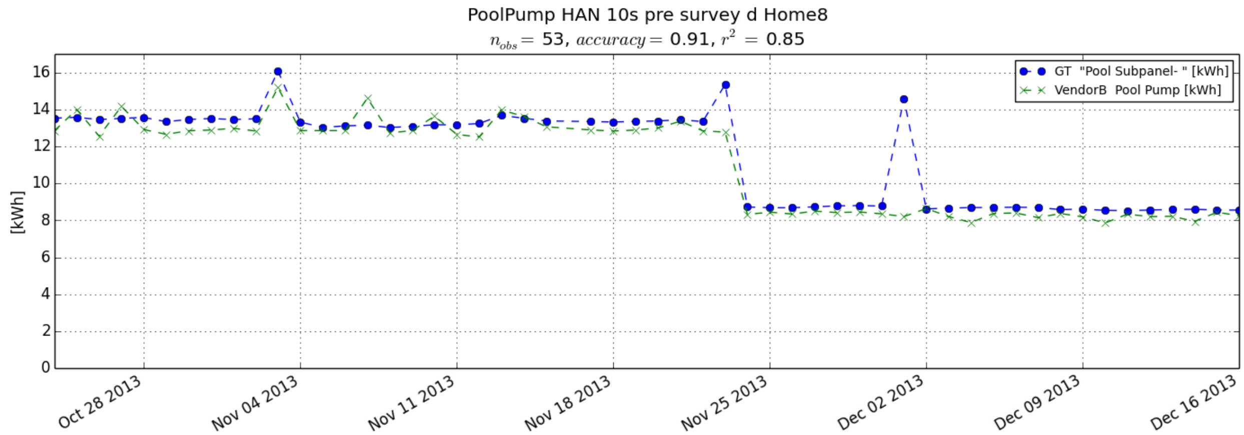


Figure 3: Vendor B accurate detection of pool pump

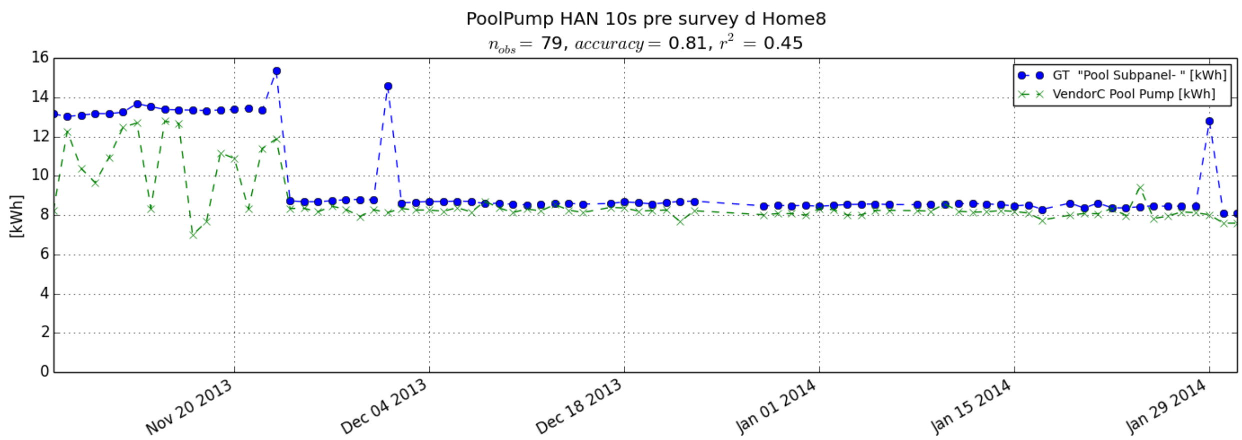


Figure 4: Vendor C accurate detection of pool pump

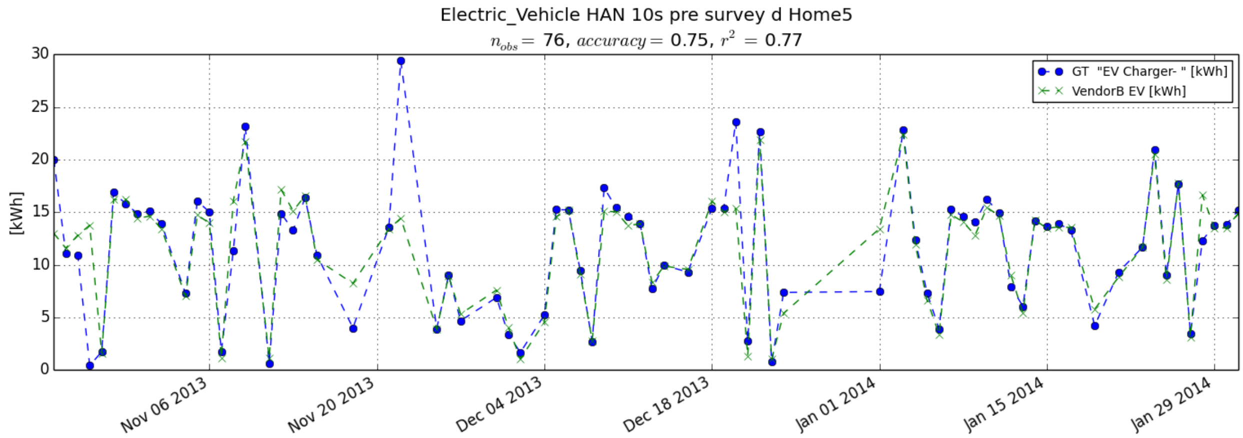


Figure 5: Vendor B accurate detection of EV

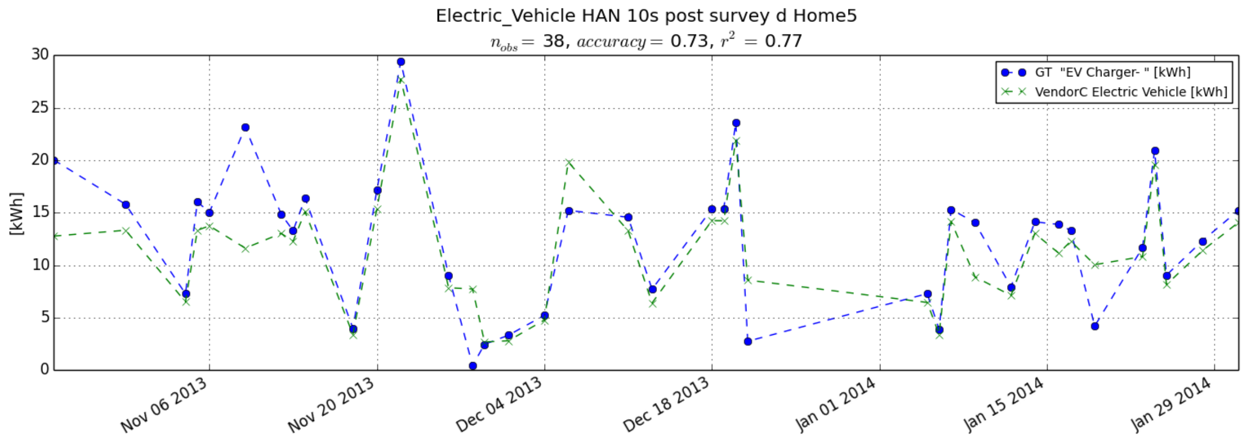


Figure 6: Vendor C accurate detection of EV

In contrast, Figure 7 and Figure 8 show one timeline plot for each vendor for a particular home where they vary dramatically on refrigeration accuracy. It is clear that the discrepancy is due the period of unusually high refrigeration use from mid-November to mid-December. For whatever reason, one vendor included these days and the other did not which led to the high discrepancy in accuracy. As indicated in the legend, it is possible that the outlet on the same breaker was used consistently for some extraneous appliance during that period only (and was not successfully omitted by the researchers' anomalous data filter).

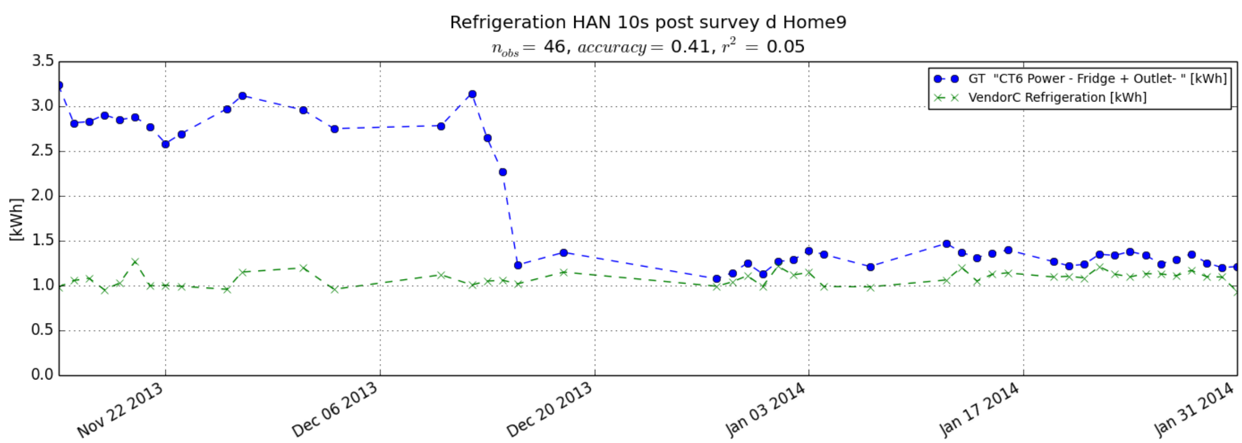


Figure 7: Vendor C poor accuracy on refrigeration due to unusual period of ground truth

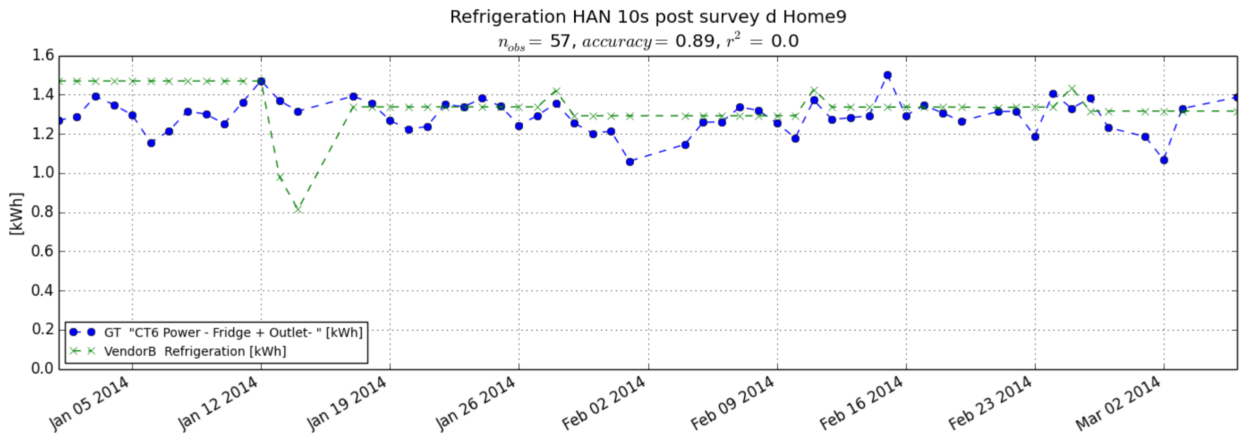


Figure 8: Vendor B good accuracy on refrigeration for more typical period of ground truth

Figure 9 shows an accurate detection by Vendor B of spa water heating. However, note that while R^2 is very good the accuracy metric is much lower. The scatter plot and individual daily errors plot in Figure 10 depict the reason. The vendor has very accurately found the variation in energy use from day to day but consistently slightly underestimated the energy use, especially for higher energy use days.

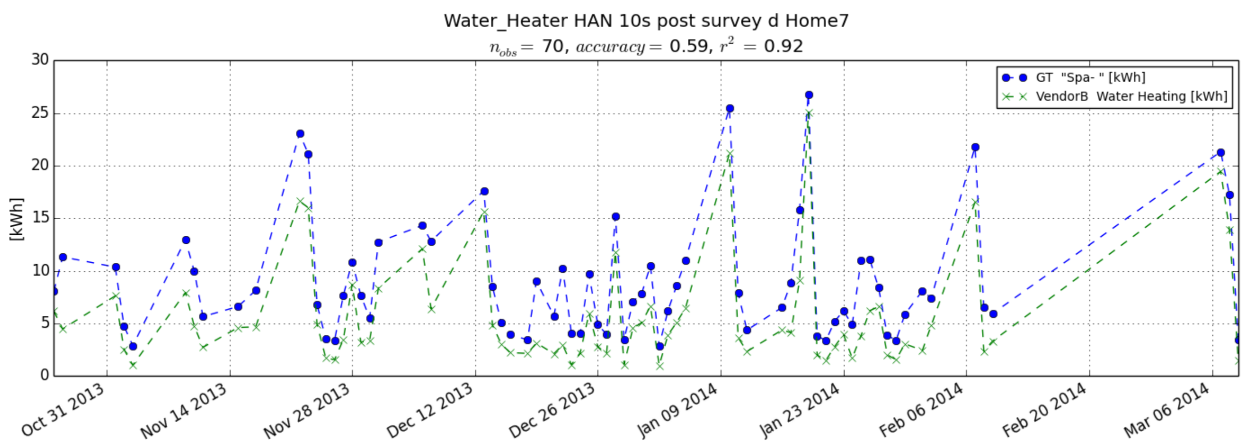


Figure 9: Vendor B good accuracy on water heating but consistently slightly low prediction

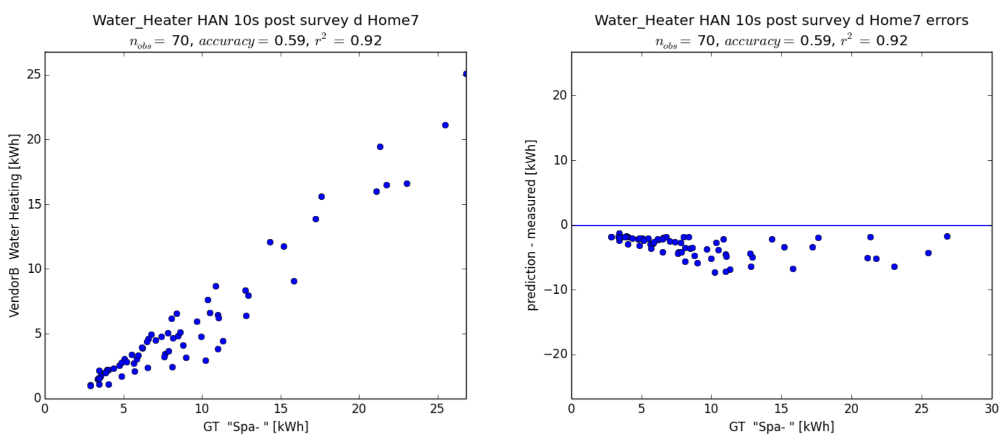


Figure 10: Vendor B water heating scatter plot and plot of errors

Figure 11 shows very accurate detection of a dryer by Vendor C. Please note that similar to the water heating example above for Vendor B, Vendor C has found the load shape more accurately than the magnitude.

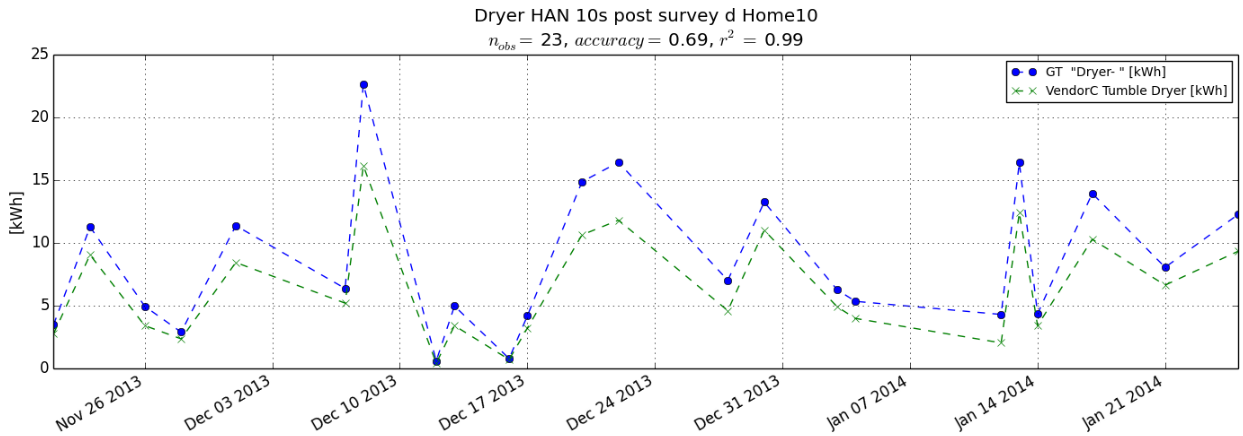


Figure 11: Vendor C good accuracy on dryer

Project Error Analysis

Project Plan Deviation

The major Project Plan deviations were that certain tasks were omitted. First, “Percent of total home energy use detected” was not calculated since numerous breakers in the ground truth data sets were omitted. Second, analysis of hourly and minute interval vendor predictions was abbreviated as opposed to being given equal footing with daily interval predictions.

Anomalous Data and Treatment

As described in the Accuracy Calculations section, the daily data was treated using the following methods and all systematically with Python scripts (other interval data was treated similarly):

- Filtered out daily energy use measurements and predictions less than 400 Watt-hours
- Ignored days where there were less than 7000 10-second whole building data points
- Removed extraneous loads from refrigeration data
- Ignored breakers with extraneous loads (except refrigeration)

Technical, Statistical, and Error Analysis

Mean and standard deviation of the errors between the ground truth and the predictions are given in section Accuracy Calculations. M&V equipment accuracy is given in Appendix B: Measurement & Verification Plan. The major issue of data quality in the breaker level ground truth data was handled by assigning a data quality index to every breaker. Aside for the refrigeration, any breaker that was not dedicated to the desired appliance was omitted from the analysis.

Conclusions

The researchers found that every vendor had something valuable to offer to SDG&E and/or its customers. Vendors B and C performed fairly well with high frequency HAN data. Vendor D performed fairly well with Green Button data and Vendor A's product is well suited for Green Button Connect. Across all vendors, high energy consuming devices like pool pumps and electric vehicles are easiest to detect. These can be detected with either Green Button or HAN data while successfully detecting smaller appliances is easier with HAN data. Home appliance survey data helps improve NILM accuracy but is not necessary.

Benefits

There are numerous benefits to accurate residential NILM. It provides insight into customer energy use that could yield energy savings opportunities. For instance, the vendor or utility could recommend replacing a refrigerator if the data showed that it was inefficient. NILM could also make home owners more engaged with SDG&E. It could potentially help resolve bill disputes. It could also help predict how a customer's utility bill would change if they switched to time of use rates.

Possible Risks

The risks include the following:

- Inaccuracies could discourage home owners from trusting the service
- Additional cost to hire the vendors
- Must purchase a HAN gateway to get the 10 second data that most of the vendors preferred

Technology Improvement Opportunities

The technology improvement opportunities include:

- Improve data quality issue with HAN gateways
- NILM vendors to provide data only for unique appliance categories instead of grouping them
- IOUs could provide higher quality and higher frequency data to NILM providers

Applicability of Case Study Findings to Other Load Types and Sectors

This case study is only applicable to residential customers with smart meters.

Considerations for Large-scale and Persistent Market Implementation

Large-scale implementation is worthwhile to pursue pending cost. All of the vendors are ready to provide a service to the IOUs and/or directly to customers. In particular, Vendor B and Vendor C performed well with high frequency HAN whole building power data. Vendor D performed well with Green Button data. Vendor A's product appears to be best suited for continued Green Button Connect participation at this time.

In regards to HAN data quality, the IOUs should continue their development efforts with HAN device manufacturers such as Rainforest Automation. The IOUs should consider allowing manufacturers to retrieve higher frequency power data and the manufacturers should continue their efforts to provide buffered data to fill in gaps of poor internet connectivity.

Possible future Study

Future study could include:

- Large scale testing on 100s to 1000s of homes
- Directly monitoring power on a subset of preferred end-uses and requiring vendor predictions at a specific interval in order to allow more targeted analysis
- Gas end-use NILM
- Higher frequency whole building data to NILM vendors
- Behavioral studies (e.g. review effectiveness of vendor’s energy efficiency recommendations)

Glossary and Acronyms

10s – 10 seconds

1h – 1 hour

1m – 1 minute

15m – 15 minutes

Cook – Electric cooking appliances (i.e. ovens and ranges)

CPUC – California Public Utilities Commission

Dish – Dishwasher

EV – Electric vehicle

GB – Green Button

HAN – Home Area Network

HVAC – Heating, ventilation, and air conditioning

IOU – Investor Owned Utility

IPMVP – International Performance Measurement and Verification Protocol

M&V – Measurement & Verification

Micro – Microwave

NILM – Non-Intrusive Load Monitoring

Pump – Pool pumps

Refrig. – Refrigerators and freezers

SDG&E – San Diego Gas & Electric

SEP – Smart Energy Profile

Solar – Solar photovoltaics

W/D – Washer and dryer

WH – Electric water heating

References

ASHRAE. (2002). *ASHRAE Guideline 14-2002 Measurement of Energy and Demand Savings.* Atlanta: ASHRAE.

California Public Utilities Commission. (2014, 02 20). *Smart Grid*. Retrieved from California Public Utilities Commission: http://www.cpuc.ca.gov/PUC/energy/smartgrid.htm

EPRI. (2013). *Non-Intrusive Load Monitoring (NILM) Technologies for End-Use Load Disaggregation.* EPRI.

NegaWatt Consulting. (2013). CircuitLevelPartialExample.csv.

NegaWatt Consulting. (2013a). *Residential Disaggregation Project Plan.* San Diego: SDG&E.

Nipun Batra, J. K. (2014). NILMTK: An Open Source Toolkit for Non-intrusive Load Monitoring. *Proceedings of the 5th International Conference on Future Energy Systems.* Cambridge: e-Energy '14.

RS Butner, D. R. (2013). *Non-Intrusive Load Monitoring Assessment: Literature Review and Laboratory Protocol.* Richland: PNNL, U.S. Department of Energy.

Zigbee Alliance. (2013). *ZigBee Smart Energy Profile 2*. Retrieved 12 3, 2013, from Zigbee Alliance: http://www.zigbee.org/Standards/ZigBeeSmartEnergy/SmartEnergyProfile2.aspx

Zoha A, G. A. (2012). Non-Intrusive Load Monitoring Approaches for Disaggregated Energy Sensing: A Survey. *Sensors*, 12(12):16838-16866.

Appendix A: Project Plan

[Intentionally blank page]

San Diego Gas and Electric Company

Emerging Technologies Program

Project ID TBD

Residential Disaggregation

Project Plan

Prepared For

Kate Zeng, Nathan Taylor
San Diego Gas and Electric Company
Emerging Technologies
8306 Century Park Court
San Diego, CA 92123

Prepared By

Bo White, Marc Esser
NegaWatt Consulting, Inc.
www.negawattconsult.com

03/31/2014

RESIDENTIAL DISAGGREGATION
NEGAWATT CONSULTING

Table of Contents

Document Change Tracking

Document Date	List of Changes
08/19/2013	Initial version
09/06/2013	Re-release after minor edits
09/11/2013	Milestone added to timeline per Kate's comments
09/17/2013	Edits in response to vendor comments
12/18/2013	Added one vendor; revised the schedule
03/31/2014	Grammatical edits

Description of the technology under investigation

Multiple vendors of non-intrusive load monitoring (NILM) technologies are studied here. They all disaggregate residential electrical utility smart meter data into constituent large appliance or appliance group energy use without any individually measured appliance data. High resolution (10 second interval) smart meter data is obtained using a Zigbee and internet enabled third party gateway installed in the home. Hourly green button electrical data is also provided to the vendors as supplemental information. Gas usage is completely ignored. All vendors disaggregate to the best of their ability using their own techniques and all focus on appliances which use more than 100 Watts. All vendors' results include energy saving recommendations but the focus of this study is the quality of the disaggregation itself.

Some vendors display their results to the occupants using a website, smartphone, or tablet app. Others primarily tailor their results directly to the utility and hope to gain access to all residential smart meter data from utility servers without the use of on-site gateways or smart meter upgrades.

Some vendors prefer that occupants complete a home survey but it isn't mandatory for their algorithms. Some prefer to also have home addresses in order to obtain demographic and climate information.

See Table 1 for an anonymized comparison of the basic features of the technologies. For more detailed information about NILM strategies currently in use, see (Zoha A, 2012).

Technology feature	A	B	C
NILM of residential electricity	X	X	X
Energy saving recommendations	X	X	X
User-facing graphical user interface	X	X	Capable by third party
Analytics targeted at home owner	X	X	X
Analytics targeted at the utility	X	X	X

Table 1: Technology comparison chart

Description of incumbent technology (or existing standard practice, etc.)

There is no incumbent disaggregation technology but there is an existing standard practice with regard to analyzing smart meter data: the federally supported Green Button program. This program gives residential utility customers access to their smart meter data. Without any additional hardware at the home, the customer can download or view plots of hourly electrical data and daily pulse gas data at 1 pulse/therm from their utility website. The user can also grant third-party access to this data. The standard practice is for those third parties to simply provide quick access and plotting within a smartphone or tablet app, sometimes augmented by rudimentary analysis or recommendations – the hourly resolution not permitting significant depth. With the purchase and installation of compatible third party hardware (supporting Home Area Network / SEP 1.x), the customer can obtain higher resolution electrical data (up to every 8 seconds) from the Smart Meter.

Goals of the pilot project

The goal of this project is to judge the efficacy of several commercially available NILM residential disaggregation technologies. SDG&E would like to understand how accurately these vendors can detect the use of individual appliances. While they may in the future also evaluate the energy saving recommendations the vendors can provide, the focus for now is to only *verify the accuracy of the disaggregation*.

Application and/or Generalization of project results to similar facilities in other locations, other types of facilities, etc.

The project results are expected to be qualitatively applicable to all residential homes with electrical smart meters within SDG&E territory. However, it is possible that the vendors will be more successful with certain home types and certain appliances than with others. The sample of homes selected for this study is comparatively small and not random in all respects. Therefore the results may not be extrapolated statistically with high confidence and accuracy. We will discuss extrapolation in detail in the results analysis and extrapolate our findings where sensible.

Measurement and Verification Plan

Please see (NegaWatt Consulting, 2013a).

Generic customer information (e.g. the type and geographic location of test sites, user demographics)

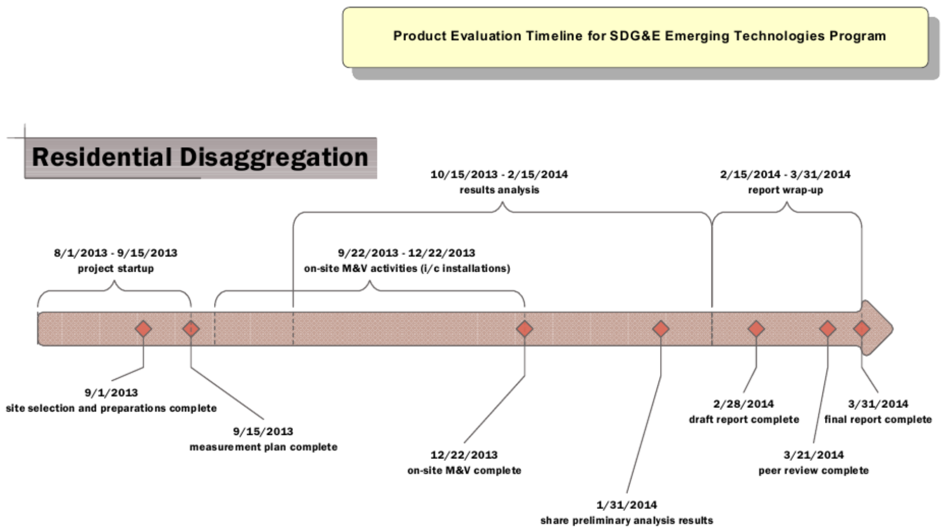
10 homes will be chosen from a group of about 27 homes owned by SDG&E employees (who can be expected to have a higher awareness about electrical energy than the average consumer). We will chose the homes so as to include as diverse an appliance pool as possible, without regard to whether the selection is representative of the market (at n=10 this would not be realistic).

One extra home owned by a non-SDG&E employee is also included.

Additional details about the test sites and selection can be found in the M&V plan (NegaWatt Consulting, 2013a).

RESIDENTIAL DISAGGREGATION
NEGAWATT CONSULTING

Project Milestones



Milestones are subject to change as project develops.

Etcetera

This assessment follows the scientific rigor protocol described in (SDG&E, 2010). The final report for this project will be made available as (NegaWatt Consulting, 2013b) on www.etcc-ca.org. Additional references will be contained therein.

This project will be tracked in NegaWatt's project management tools once the project plan has been approved. The document repository for this project is NegaWatt's secure file server. Please contact the authors of this project plan if you need access to these systems or to any of the referenced documents.

References

NegaWatt Consulting. (2013a). *Residential Disaggregation M&V Plan.* San Diego: SDG&E.

NegaWatt Consulting. (2013b). *Residential Disaggregation Final Report.* San Diego: SDG&E.

SDG&E. (2010). *Proposed Protocol for Scientific Rigor and Documentation In ETP Technology Assessments.* San Diego: SDG&E.

Zoha A, G. A. (2012). Non-Intrusive Load Monitoring Approaches for Disaggregated Energy Sensing: A Survey. *Sensors*, 12(12):16838-16866.

Appendix B: Measurement & Verification Plan

[Intentionally blank page]

San Diego Gas and Electric Company

Emerging Technologies Program

Project ID TBD

Residential Disaggregation

M&V Plan

Prepared For

Kate Zeng, Nathan Taylor
San Diego Gas and Electric Company
Emerging Technologies
8306 Century Park Court
San Diego, CA 92123

Prepared By

Bo White, Marc Esser
NegaWatt Consulting, Inc.
www.negawattconsult.com

12/18/2013

RESIDENTIAL DISAGGREGATION

NEGAWATT CONSULTING

Table of Contents

Document change tracking

Document Date	List of Changes
09/10/2013	Initial version
09/11/2013	Draft released to SDG&E, with small changes
09/17/2013	Edits in response to vendor comments
10/01/2013	Second round of edits in response to vendor comments (not published)
10/08/2013	Added Server and File information for data downloads
12/18/2013	Edits in response to vendor comments

Introduction

This measurement plan is an integral part of the project described in (NegaWatt Consulting, 2013a).

It follows the guidelines established in (SDG&E, 2010).

It has been designed to accurately assess both the baseline performance of the incumbent technology (or standard practice in the absence of an incumbent) and the performance of the technology under study.

It has been designed in compliance with one of the evaluation methods identified in the International Performance Measurement and Verification Protocol (IPMVP) except where site- or technology-specific circumstances dictated a deviation from one of these protocols. The Measurement Plan identifies selected IPMVP method to be used or the justification for any deviations from IPMVP.

All instrumentation under the control of evaluation staff shall be calibrated in accordance with guidelines established in the IPMVP as described in (Efficiency Valuation Organization, 2012).

For field evaluations, all reasonable efforts shall be made to calibrate or replace any customer-owned instrumentation or where this is not possible, to document the calibration status of such instrumentation.

Measurement uncertainty for each monitoring device will be documented. Note that error analysis evaluating other significant sources of uncertainty will be required for the Final Report.

All instrumentation will be commissioned prior to initiating data collection to ensure that measurement and logging systems are functioning properly, to minimize risk of unusable data sets.

Any anomalous data will be investigated and explained. Following investigation, careful consideration will be given to whether such data should be incorporated in the analysis or replaced by additional data collection.

Any events that occur at customer premises during the data collection period that are likely to compromise the validity of the assessment project and that are beyond the control of evaluation staff will be communicated to program management without delay.

Test site description

As stated in (NegaWatt Consulting, 2013a), "10 homes will be chosen from a group of about 27 homes owned by SDG&E employees". First, surveys were distributed to all 27 employees in order for them to self-report which major electrical appliances they owned. The appliance surveys were then reviewed and 10 homes were chosen with as diverse an appliance pool as possible, without regard to whether the selection is representative of the market (at n=10 this would not be realistic).

Please note that one extra home was subsequently added for a total of 11 homes.

Here is a non-exhaustive list of electrical appliances that may be present (i.e. zero, one, or more) at each home:

- Water Heater
- Pool Pump
- Air conditioner (and in some cases with compressor-based or resistance-based heating)
- Fridge/Freezer Combo
- Freezer
- Dishwasher
- Hot Tub
- Electric Vehicle
- Washer/Dryer
- Stove (range and/or oven)
- Stand-alone Resistance Heater
- Microwave
- Recreational Vehicle (RV) Hookups
- Aquarium
- Miscellaneous small household, audio, video, computing, or entertaining appliances

Please note we do not know whether any of the homes has a grid-tied solar system at this time. However, we will determine this on-site and we would like for the vendors to "find it" if there is one.

As is to be expected, all of the homes are in the SDG&E region and all have smart meters.

Data collection procedures

Data points & recording intervals – Meter and Green Button Data

For a tabulated list of measured data points, please refer to Table 1.

Two SDG&E Home Area Network (HAN) approved gateways ("Eagle") made by Rainforest Automation will be installed at each home (one is customized for one of the vendors, the other is generic and will connect to a NegaWatt server). They will be connected to the home's internet router and configured to communicate with the smart meter via Zigbee.

The smart meter data will be shared as individually gz-compressed xml and csv files, each containing 1 hour of data, at 3 different resolutions. The file naming convention is as follows:

[MacID]-[Timestamp]-[resolution].[xml|csv].gz, where

- MacID is the MacID of the Rainforest eagle device collecting the data. Since we have one at each home, there will be 11 different MacIDs.
- The timestamp is the number of seconds since 1/1/2000 in decimal format for the GMT timezone with no daylight savings time corrections.
- The resolution is either of "10s", "1m" or "15m". The latter two are upsampled mean values from the 10s data.
- "xml" or "csv" describes the file type. The xml files contain raw data as obtained from the Rainforest Eagle, with outer tags (i.e. <xml>) removed. Similarly, the csv files contain no headers. This is so the files can be concatenated without modification. The csv files contain identical data than the xml files, with unused data and tags removed. csv columns (in this order): *timestamp* (same as in file name but hexadecimal, not decimal), and *instantaneous demand*, in Watts, also hexadecimal. Participants can choose to use either file.

Green Button data (1h resolution) will be made available for each home on the same server. The data will follow green button data standards; we receive this data from SDG&E and simply pass it on.

All data can be accessed by technology vendors on our project server **eagle.negawattconsult.com** using the **sftp** protocol (ftp over ssh). Username and password will be provided as needed. Rainforest eagle (meter) data will be placed on the server hourly, new green button data will be placed daily. Data samples for all the above will be provided on the server.

Vendors may provide disaggregation reports for any data resolution of their choice; SDG&E prefers a report off of *all* available resolutions of 10s, 1m, 15m and 1h, but vendors may choose to offer reporting only for a subset of the resolutions at their discretion. Please note, 10s and 1h data is available for all Smart Meter Customers as is today, with the 10s data requiring the purchase of a HAN gateway. The 1m and 15m resolutions are available for some customers today, and may be available for more or for all customers in the future.

Please note that occasionally there is missing data. The vendors shall address this issue as best they can. If there are extensive gaps of missing data, we will treat any possible vendor results within or around those periods separately.

Vendors may provide their own M&V hardware and work off their own data if they so choose. We will assist with any required coordination or installation at the homes in this case.

If vendors' algorithms are not fully automated but require upfront "learning", for example by means of a survey to the home owner, results should be grouped into "with" and "without" successful "learning". During a large scale deployment it is unlikely that "learning" will complete successfully for all homes.

Data points & recording intervals – Ground Truth Data

For a tabulated list of measured data points, please refer to Table 1.

Current transformers will be installed at up to 12 circuits at the home. In rare cases – when there are subpanels, or otherwise significantly more than 12 circuits, we may install 2 data loggers in each home. We will do our best to select the circuits with the highest monthly energy usage, albeit this may be difficult in homes with many different loads. The priority at install time will therefore be to identify and include a) *all* circuits with *greater* than 20A capacity, and b) circuits with systems that are used regularly and that have a *non-flat profile* such as appliances (as opposed to flat profile systems such as lighting, which we will *not* measure with great detail). The current transformers will be connected to a line-powered logger inside the electrical panel (see below for instrumentation details). The logger data is then made available for online readout with a username and password via power line communication. The logger saves 10 minutes of "volatile" 1 second data and 1 year of 1-minute data. The 1-second data will be referred to as needed but not saved. The 1 minute data will be downloaded remotely, stored securely on a NegaWatt server, and be used for analysis.

Whenever there are multiple loads on one circuit, all of those loads will be manually cycled and power data will be recorded in 1-second intervals with either a Fluke Power logger, Extech clamp-on meter, or the circuit-level logger, whichever is most appropriate based on site conditions. This one-time data ("pattern samples") will then be used to manually disaggregate the loads that share a single breaker. Approximate confidence intervals will be noted for each circuit. We anticipate fairly little such manual disaggregation, because the priority are large loads that typically have their own breakers. For those loads the logged data (per breaker) is usable directly as "ground truth".

Data Point	Units	Quantity	Sampling	Recording	Notes
Circuit current	A	6-24	1 second	1 minute	For select breakers with relevant systems, for ground truth
Voltage	V	2	1 second	1 minute	Phases A and B
Circuit real power	W	6-24	1 second	1 minute	The calculation is performed in the logger software; measurement is averaged over the given interval.
Circuit apparent power	VA	6-24	1 second	1 minute	Same as for real power

RESIDENTIAL DISAGGREGATION NEGAWATT CONSULTING

Circuit power factor	None	6-24	1 minute	N/A	Calculated later as needed using power measurements
Smart Meter power via Rainforest Eagle	W	1	~10 seconds	~10 seconds	Instantaneous measurement (i.e. not averaged over the given interval)
Green Button electrical use	kWh	1	1 hour	1 hour	

Table 1: Data points, sampling and recording intervals (per home)

If desired we can save a limited amount of 1-second ground truth data per home or for certain appliances and systems for use by the vendors at their discretion. This may help for example to establish profiles for systems that are not widespread. We will share any such data with all vendors to ensure a level playing field. Please contact us if you need any 1-second data during the project.

Please see attached example file CircuitLevelPartialExample.csv (NegaWatt Consulting, 2013) for a partial example of the raw data from the circuit-level logger. Vendors may use formatting similar to the circuit-level logger raw data file to allow for a direct comparison of their results with our ground truth, if they find the format appropriate considering their algorithms, products and services. Different result presentations are just as acceptable. Note our ground truth data will be available in 1-minute, 1-hour, daily and monthly intervals, and will be used as such to evaluate the results (see also next section).

At the conclusion of the project, the circuit-level loggers and the generic one of the two gateways per home will be uninstalled. The home owners may be given the opportunity to keep the gateway that is customized for one of the vendors and continue using it at SDG&E's discretion. The Field Demonstration Agreement contains further details about equipment left behind.

Instrumentation

For a tabulated list of the measurement and verification instrumentation used, please refer to Table 2.

Make/Model/Name Link(s)	Specifications	Notes
eGauge EG3010 Real-time Web-based Energy Monitor. https://www.egauge.net/docs/eg30xx-ds.pdf	Device Power Usage: 3W typical. Measurement Voltage: Up to 480 VAC, 3 channels Measurement Current: 12 channels. Measurement Power: any voltage/current combination	Purchased new in Fall 2013. "Homeplug AV" included for power line communication.
Magnelab SCT-0400 Split-core AC Current Sensor. https://www.egauge.net/docs/sct_0400.pdf	Output: 0.333V at rated AC current Accuracy: ±1% Range: 10% to 130% of rated current Frequency: 50Hz – 400Hz	Purchased new in Fall 2013. Multiple sizes from 10 to 50 amps.
Rainforest Automation RFA-Z109 EAGLE Ethernet Gateway. http://rainforestautomation.com/eagle	Wireless Link: IEEE 802.15.4 MAC, 2.4GHz ISM Band Receiver sensitivity: -99dBm, Transmit Power: +20dBm Wireless Range: Up to 50m (150ft) Ethernet Link: 10/100Base-T, RJ-45 connector Standards: ZigBee SEP 1.1; HTTP/TCP/IPv4/Ethernet Power: External AC adapter (included)	Purchased new in Fall 2013.
Extech 380976 1-phase/3-phase 1000 Amp True RMS Power Clamp-On Meter http://www.extech.com/instruments/product.asp?catid=27&prodid=705	Function: Max Range/Resolution, Basic Accuracy True Power (W): 600kW/10W, ±5% Apparent Power(kVA): 600kVA/100VA, ±2% Reactive Power (kVAR): 600kVAR/10VAR, ±5% Phase Angle (f): -60 to +60° / 0.1°, ±6° AC Current (Trms): 1000A/10mA, ±2% uA Current (AC+DC) (Trms): 1000µA/10nA, ±1%	Calibrated on 10/11/2012.

RESIDENTIAL DISAGGREGATION

NEGAWATT CONSULTING

	AC/DC Voltage (Trms): 600V/0.1mV, ±1%	
Fluke 1735 Three Phase Power Logger http://www.fluke.com/fluke/usen/power-quality-tools/three-phase/fluke-1735.htm?PID=56028	V-RMS WYE MEASUREMENT Resolution: 0.1 V Intrinsic error: ± (0.2% of measured value + 5 digits) Operating error: ± (0.5% of measured value + 10 digits) V-RMS DELTA MEASUREMENT Resolution: 0.1 V Intrinsic error: ± (0.2% of measured value + 5 digits) Operating error: ± (0.5% of measured value + 10 digits) A-RMS CURRENT MEASUREMENT Resolution: 0.01 A Intrinsic error for 15A range: ± (0.5 % of m. v. + 20 digit) Operating error for 15A range: ± (1 % of m. v. + 20 digit)	Calibrated on 3/30/2011.

Table 2: M&V Instrumentation

Data analysis procedures

Data manipulation (aggregation, statistical analysis, etc.)

The data measured at the homes will be aggregated into hourly, daily, and monthly kWh values per appliance, in some cases with manual disaggregation applied as described earlier. We will use this data as follows:

1. Each vendor's approach will be discussed qualitatively to clarify what should and should not be expected of their results.
2. The percentage of total home energy use that was disaggregated correctly will be calculated. The Green Button or gateway data will be aggregated to obtain the whole home energy use.
3. The ground truth appliance-specific data will be compared to each vendor's appliance-specific data at the respective intervals, as available. The results will be compared a) **visually** by means of charts similar to Figure 1, and b) **statistically** by calculation of "Energy Share Accuracy" per appliance per page 23 of (RS Butner, 2013):

$$Energy\ share\ accuracy = 1 - \frac{measured\ energy\ share - NILM\ energy\ share}{measured\ energy\ share}$$

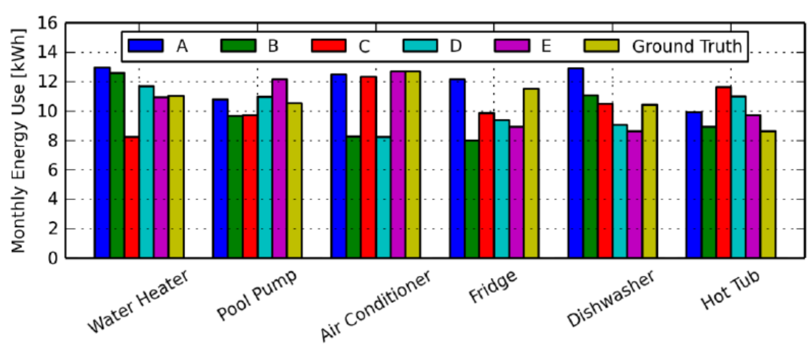


Figure 1: Example bar plot of disaggregation results from one home

4. The ratio of correctly detected appliances to the total number of regularly used major appliances per home will be calculated.
5. The accuracy of detecting each appliance run occurrence on an hourly basis will be calculated. The vendor's hourly disaggregated data and the measured hourly data will be converted to binary signals and compared for this purpose. In this scope, we will also confirm which appliances have been identified correctly or missed, overall.

Calculation of energy and demand savings

No energy or demand savings calculations will be performed. A qualitative review of vendors' energy saving recommendations (where applicable) may be provided.

Calculation of cost savings

No cost savings calculations will be performed.

References

Efficiency Valuation Organization. (2012, 01 01). *IPMVP Volume I EVO 10000-1:2012.* Retrieved 04 17, 2013, from Products - IPMVP: http://www.evo-world.org/index.php?option=com_docman&task=doc_view&gid=641&lang=en

NegaWatt Consulting. (2013). CircuitLevelPartialExample.csv.

NegaWatt Consulting. (2013). GatewaySample.zip.

NegaWatt Consulting. (2013a). *Residential Disaggregation M&V Plan.* San Diego: SDG&E.

NegaWatt Consulting. (2013a). *Residential Disaggregation Project Plan.* San Diego: SDG&E.

NegaWatt Consulting. (2013b). *Residential Disaggregation Final Report.* San Diego: SDG&E.

RS Butner, D. R. (2013). *Non-Intrusive Load Monitoring Assessment: Literature Review and Laboratory Protocol.* Richland: PNNL, U.S. Department of Energy.

SDG&E. (2010). *Proposed Protocol for Scientific Rigor and Documentation In ETP Technology Assessments.* San Diego: SDG&E.

SDG&E. (2013). *Additional information for developers and third parties*. Retrieved from San Diego Gas & Electric: http://www.sdge.com/customer-service/green-button/additional-information-developers-and-third-parties

Zoha A, G. A. (2012). Non-Intrusive Load Monitoring Approaches for Disaggregated Energy Sensing: A Survey. *Sensors*, 12(12):16838-16866.

Appendix C: Peer Review Certificate
