Estimating Household Space Heating Consumption of Natural Gas Using Billing and Weather Data

Shaofen G. Deng[1], Greg Lawson[1]

[1]U.S. Energy Information Administration, 1000 Independence Ave, Washington D.C. 20585

Abstract
This research explores using billing data and weather data to estimate household consumption of natural gas for space heating. The data sources for this analysis were collected as part of the Residential Energy Consumption Survey (RECS). RECS is a periodic study conducted by the U.S. Energy Information Administration (EIA) since 1978. The data collected from RECS include housing characteristics and energy billing data. EIA uses these data to estimate residential energy consumption by fuel, and consumption for individual household end uses. In past cycles of RECS, space heating consumption was estimated based on either non-linear statistical models or engineering models using housing characteristics data, annualized billing data, and annual heating degree days. Since the end uses of natural gas in a household are limited (i.e., many fewer than those for electricity), and heating consumption is assumed to be highly correlated with weather, this research looks into a more direct method for natural gas space heating consumption. Without relying on housing characteristics data, this method is based on a regression approach using natural gas monthly bills and daily temperature data.

Key Words: RECS, natural gas, space heating, billing data, weather data, heating degree days

1. Introduction/Background

The Residential Energy Consumption Survey (RECS) is a periodic study conducted by the U.S. Energy Information Administration (EIA) since 1978. The study has two major phases of data collection: the first phase is a household questionnaire that collects housing characteristics and energy usage behaviors of occupied homes, such as home type, frequency of cooking, etc. The second phase is a request to energy suppliers for the billing data of the survey respondents from phase one, EIA collects the electricity, natural gas, propane, and fuel oil bills of the respondents when applicable. The billing data collected for electricity and natural gas are typically on a monthly basis, but for the bulk fuels, they are collected on an on demand basis.

EIA then uses the housing characteristics data and the billing data to perform various data analyses including but not limited to: estimation of total residential consumption at various geographic levels, ownership of various appliances, as well as estimation of various end-use consumptions in a home.

Unlike the estimation of total consumption, which is mainly calculated from the billing data, the end-use estimation requires more complex techniques such as submetering or disaggregation modeling. Submetering is a direct measurement of the household end uses with real-time energy monitoring, but this effort is very expensive to execute because it requires proper installation of equipment technologies, continuous monitoring of energy consumption, plus proper collection and storage of the data. Therefore, instead of using submetering measurements, the RECS program has been using modeling strategies such as nonlinear regression models or engineering models for end-use disaggregation. These models depend on housing characteristics as inputs. Figure 1 is an illustration of end-use disaggregation for electricity and natural gas. In the billing statements, we only get the total energy consumption of a house, we don't know how much is for space heating, water heating, etc., so the idea is to disaggregate the total consumption down to the level of individual appliances.

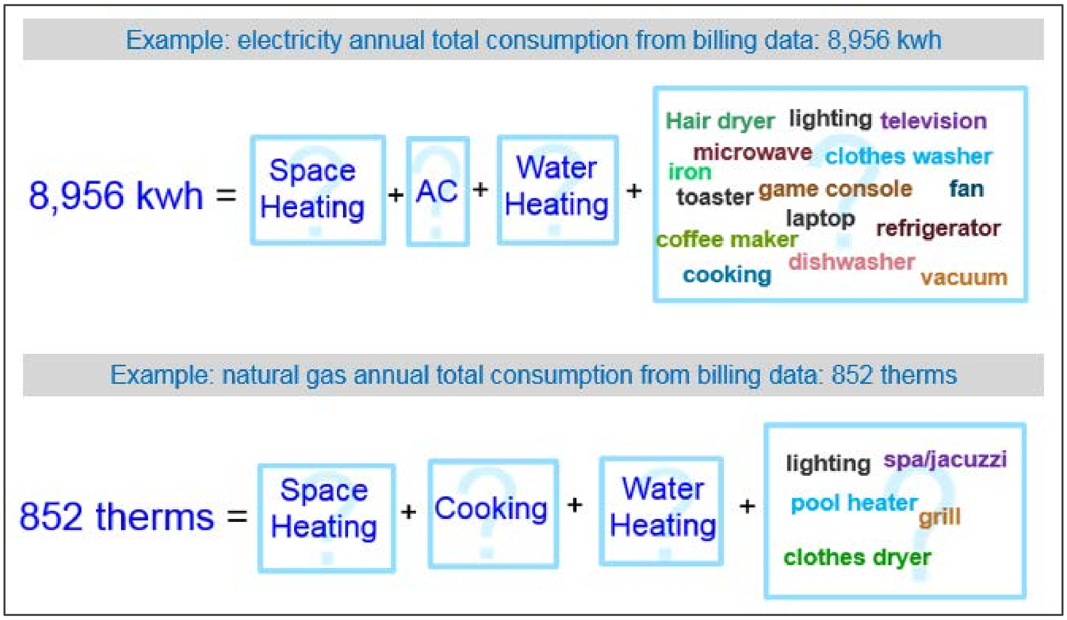

Figure 1: Example of end-use disaggregation for electricity and natural gas

For the RECS end-use modeling approach, nonlinear regression models were used in the RECS cycles prior to 2015. Starting with the RECS 2015, engineering models have been used, more detailed description of the engineering models approach can be found in the RECS methodology document 2015 Consumption and Expenditures Technical Documentation Summary. This paper focuses on the estimation of space heating as an end use for natural gas. As background information, the current engineering modeling approach for natural gas space heating has two major steps. The first step is to compute a heating load from the building characteristics, geography location, weather information, then followed by estimating the expected energy use to meet the load demand based on the efficiencies of the fuel and equipment in use. The second step is to obtain the final estimate through a calibration process, this is done by adjusting the initial space heating estimate with the other end-use estimates and their corresponding uncertainties so that the sum of all the calibrated end uses is equal to the whole-house billing total. Figure 2 outlines the two main steps of the engineering modeling approach for natural gas space heating.

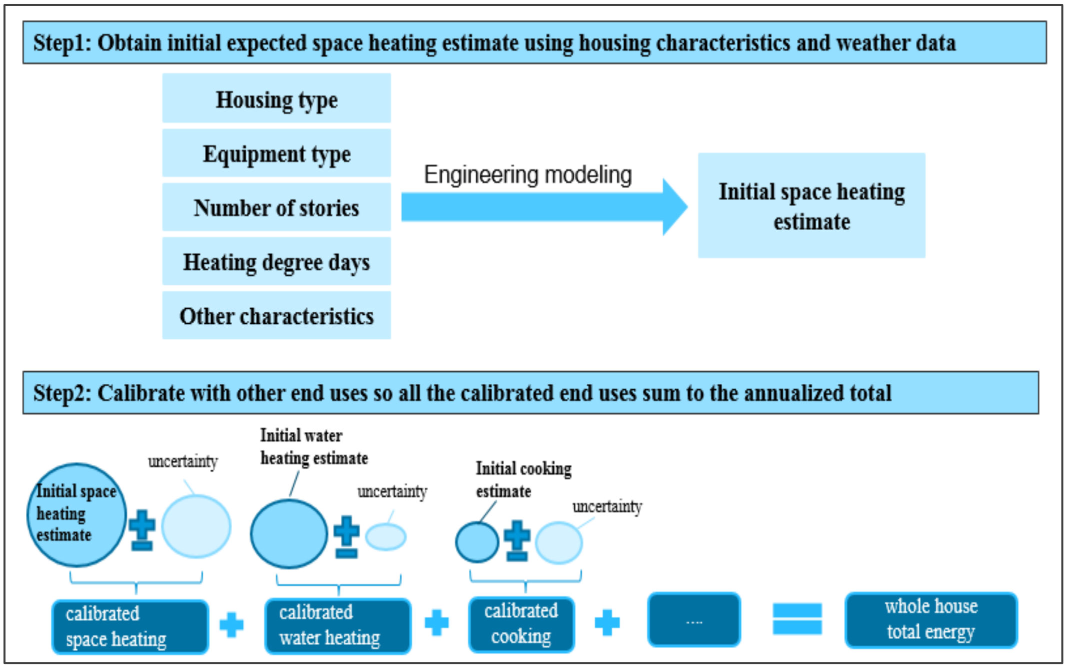

Figure 2: Estimation of natural gas space heating using the engineering model approach

The current engineering model approach relies on housing characteristics as inputs, and we are interested to see if using mainly weather and monthly billing data can provide us a robust estimate for natural gas space heating. The strong relationship between outdoor temperature and energy consumption has been long recognized. Since the 1980s, the Princeton Scorekeeping Method (PRISM) (Fels, 1985; Fels et al, 1986; Goldberg & Fels, 1986, Westman, 1986) was developed using billing and weather data to estimate a weather-adjusted index of consumption for energy savings (Fels, 1986). This index is called Normalized Annual Consumption. Unlike the purpose of PRISM, the goal of the billing and weather data methodology used in this paper is mainly to estimate the annual consumption of natural gas for space heating during the reference year. Therefore, the research questions in this paper are trying to address:

1. Can a methodology similar to the PRISM be used to properly estimate the annual consumption of natural gas space heating for the RECS households?
2. In addition to space heating, literature also shows that water heating is impacted by weather (Maguire et al, 2013; Goldner, 1994; Masiello & Parker, 2002) Therefore, how do we account for the seasonality effect(s) of water heating?

This paper focuses on space heating consumption of natural gas first, as opposed to space heating of other fuels, because 1) we have regular monthly billing data; 2) the number of end uses for natural gas in a home is limited, so space heating is likely to be the only major consumption. Even though the same approach can be applied to electricity, the modeling process would be a lot more complex with more end uses of electricity.

Finally, the submetered energy data from the Residential Building Stock Assessment (RBSA) (NEEA, 2014) study will be used as an example to illustrate the methodology, because the submetered data for space heating can be considered as the ground truth to validate the methodology. Some more details on the RBSA data are provided in the methodology section.

2. Methodology

The methodology for natural gas space heating consumption in this paper is defined in five steps outlined below. The details in each step will be described in the upcoming sections.

- Step 1: Find an optimal heating base temperature using regression models.
- Step 2: Identify the non-heating months and heating months based on the optimal heating base temperature.
- Step 3: Estimate average non-space heating baseload from the non-heating months.
- Step 4: Adjust the initial baseload for the seasonality effects of water heating.
- Step 5: Calculate final natural gas space heating consumption from the total consumption and adjusted baseload.

The most crucial step in this billing and weather data analysis is to obtain a reliable heating base temperature in step one. Traditionally, 65°F has been used as the standard heating base temperature in the United States, but realistically, the heating base temperatures of homes are likely to vary due to the climate, location, and household behaviors. Therefore, the billing analysis approach here is to obtain individualized heating base temperature for each home, which is based on the PRISM method.

What is a heating base temperature? A heating base temperature is the outdoor temperature of a building that signifies the heating system is starting to be in use. This base temperature is used to calculate heating degree days (HDDs), which provides an indication of how much energy is needed to heat a home. The formula for calculating HDDs is:

$$HDD_{Tb} = \max\,(0, T_{base} - T_{outdoor})$$

where HDD_{Tb} *is the daily heating degree days associated with a base temperature*
T_{base} *is a base temperature*
$T_{outdoor}$ *is the average daily outdoor temperature*

For example, on a particular day, suppose the heating base temperature of a home is 65°F. If the outdoor temperature is 50°F, then HDD_{65} is equal to 15, this index value informs us about the size of energy needed to heat a home. But if the outdoor temperature is 70°F, then HDD_{65} is equal to zero, which indicates no heating is required for the home.

2.1 Energy Consumption Data and Weather Data Used In This Study

Three datasets are used in this analysis. The first dataset is the submetered end-use data from the RBSA study: the submetered consumption data of the end uses and the corresponding outdoor temperature were recorded at 15-minute intervals for public use. The RBSA study was sponsored by the Northwest Energy Efficiency Alliance (NEEA) to measure residential end uses in the homes of Idaho, Montana, Oregon, and Washington during 2011 and 2013. About 100 houses had direct measurements of the energy consumption for various end uses, but only data of 31 homes with submetered natural gas space heating and water heating are used for the analysis in this paper. Since the RECS billing data of natural gas are available at monthly basis, the 15-minute submetered data were aggregated to monthly totals to simulate the RECS billing data. In addition, the RBSA natural gas data didn't have the whole-house consumption available, but only certain end uses of natural gas data were being monitored, therefore, for this analysis, the whole-house natural gas total consumption was simulated by adding up the metered space heating consumption and metered water heating consumption. As mentioned before, this RBSA dataset is used mainly for the validation of the methodology.

The second and the third datasets are in association with the 2015 RECS, they are used as a result of the methodology application. The second dataset is the monthly natural gas billing data from the 2015 RECS energy Supplier Survey (ESS). Because the monthly billing data do not always start on the first day or end on the last day of a calendar month, the days of the end months might need to adjusted for the reference year. The third dataset is the 2015 daily weather data pulled from an independent source – the Quality Controlled Local Climatological Data (QCLCD) of the National Oceanic and Atmospheric Administration (NOAA). The weather data will need to be processed and properly linked to the billing cycles.

2.2 Methodology Details Using an Example From the RBSA Dataset

2.2.1 Step 1: Find an optimal heating base temperature using regression models

As mentioned in the previous section, the most crucial step in the methodology is this first step. To determine the optimal base temperature of a home, the idea is to first set a range of possible base temperatures, calculate the corresponding HDDs of each base temperature, and then use regression modeling to find the base temperature that has the best fit – the highest r^2. This base temperature would then be picked as the optimal base temperature. The model used is just a simple regression model, the dependent variable (daily_therms) is the average daily consumption within a billing cycle, and the independent variable (daily_HDD_{Tb}) is the average daily HDDs associated with a base temperature. The model is defined as the formula below:

$$daily\_therms = \beta_0 + \beta_1 * daily\_HDD_{Tb} + e$$

where *daily_therms* is the average daily consumption of total energy in a billing cycle
β_0 = intercept parameter
β_1 = slope parameter
e = error
daily_HDD_{Tb} is the average daily heating degree days associated with a base temperature

2.2.1.1 RBSA Example Data

To illustrate the process of finding the heating base temperature in a home, an RBSA sample case is used. Table 1 consists of the simulated monthly consumption, and the average daily HDDs for base temperature candidates at 45ºF, 50ºF, 57ºF, and 65ºF, respectively. The average daily consumption and the corresponding average daily HDDs of each billing cycle can be calculated from the monthly total consumption and the monthly total HDDs. Each billing cycle of the RBSA data is simply the number of days in a calendar month. If we plot the monthly total consumption by month, as shown in Figure 3, the consumption is higher in colder months, which indicates space heating is in use. In Figure 4, the average daily consumption is plotted against the average daily outdoor temperature, we can see that the daily consumption drops linearly as temperature increases, until to a certain point, then consumption becomes more stable, which is an indication of baseload consumption.

Table 1: Monthly simulated consumption data of a RBSA example

month	*Billing days*	*Average outdoor temperature*	*"Total" therms*	*Average daily therms*	*Average daily hdd45*	*Average daily hdd50*	*Average daily hdd57*	*Average daily hdd65*
Jan	31	37.3	132.7	4.38	7.88	12.7	19.7	27.7
Feb	28	43.5	87.4	3.12	1.83	6.54	13.5	21.5
Mar	31	46.1	76.1	2.45	1.57	4.46	10.9	18.9
Apr	30	51.3	42.2	1.41	0.09	1.28	5.95	13.7
May	31	55.5	25.4	0.82	0	0.14	2.72	9.50
Jun	30	58.3	17.5	0.58	0	0	1.01	6.68
Jul	31	64.5	8.8	0.28	0	0	0	1.47
Aug	31	66.5	7.1	0.23	0	0	0	0.89
Sep	30	60.0	7.9	0.26	0	0	0.21	5.09
Oct	31	52.1	44.3	1.43	0.18	1.13	5.07	12.9
Nov	30	45.3	81.0	2.70	2.13	5.47	11.7	19.7
Dec	31	40.8	112.6	3.63	4.45	9.24	16.2	24.2

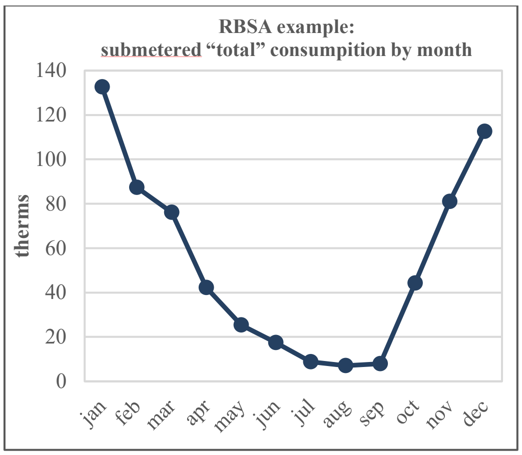

Figure 3: RBSA example plotting monthly total consumption by month

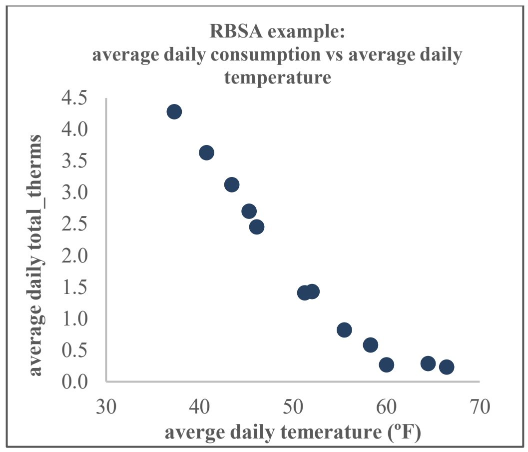

Figure 4: RBSA example plotting average consumption vs average temperature

2.2.1.2 Determining the optimal base temperature

For each base temperature, the regression model defined previously is run, and each model outputs the corresponding intercept, slope, and r^2. The base temperature with the model that has the highest r^2 is considered to be the optimal base temperature. Figures 5-8 are the individual regression models with the corresponding r^2 for the base temperature candidates at 45ºF, 50ºF, 57ºF, and 65ºF. In the comprehensive analysis, the regression models were run with base temperature candidates ranged from 40ºF to 75ºF, Figure 9 shows the corresponding r^2 of each base candidate from 40ºF to 75ºF, and the maximum r^2 occurs when the base candidate is equal to 57ºF, which will be set as the optimal base temperature. This modeling process was programmed in SAS language. In Figure 10, the daily consumption is plotted against daily temperature, the heating base temperature 57ºF can be viewed as the change point that separates when a home has heating consumption from when it does not. The RETScreen[1] software, a Clean Energy Management software developed by the Government of Canada for energy savings and efficiencies, which uses

[1] The RETScreen software is developed by the Government of Canada. For more information, visit https://www.nrcan.gc.ca/energy/retscreen/7465

the same statistical procedure in finding an optimal base temperature, was used for the example case to verify the proper programming procedures by SAS.

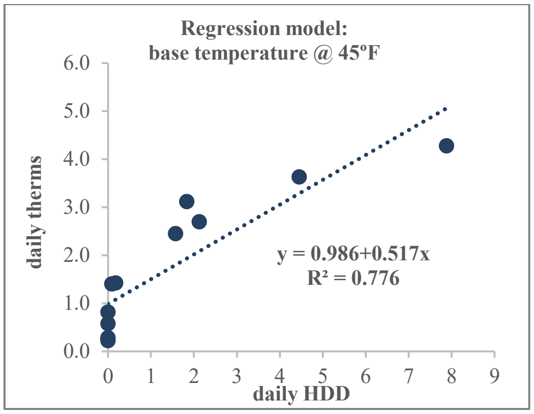

Figure 5: Regression model for RBSA example when base temperature @ 45ºF

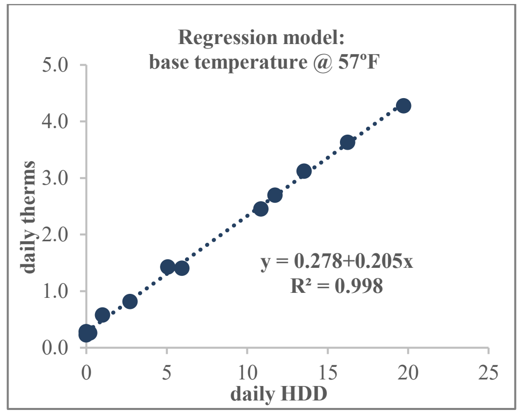

Figure 6: Regression model for RBSA example when base temperature @ 57ºF

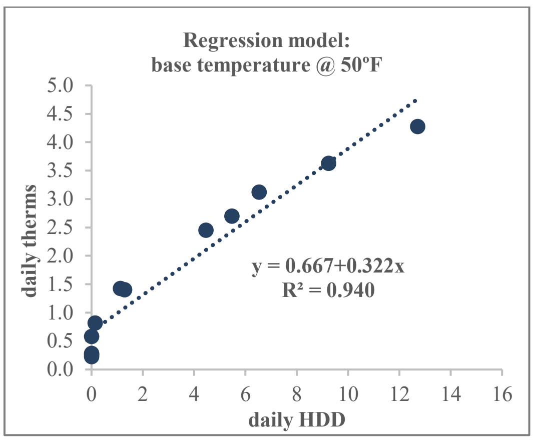

Figure 7: Regression model for RBSA example when base temperature @ 50ºF

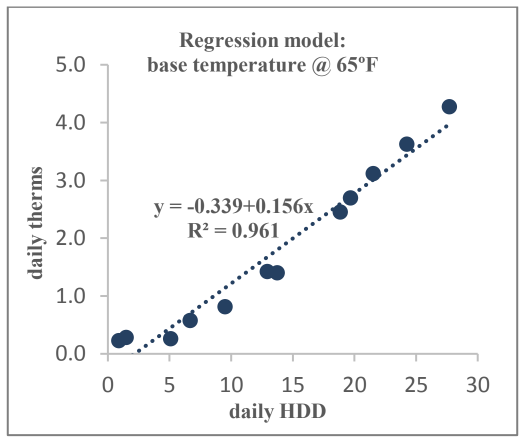

Figure 8: Regression model for RBSA example when base temperature @ 65ºF

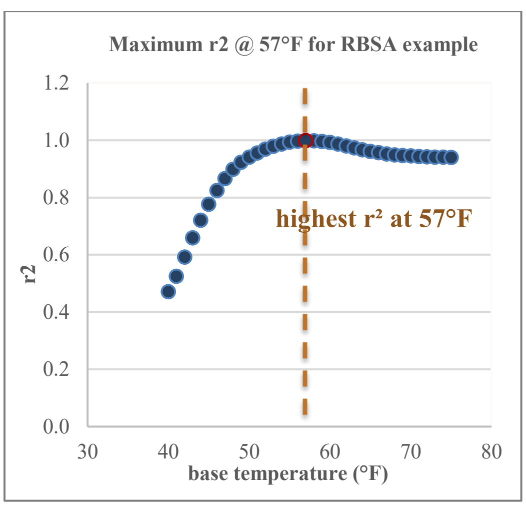

Figure 9: r^2 plot for RBSA example with base temperature ranged from 40ºF to 75ºF

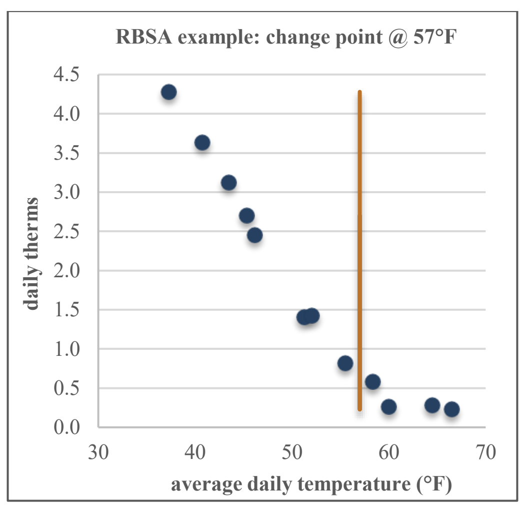

Figure 10: 57ºF is shown as the change point

2.2.2 Step 2: Identify the non-heating months and heating months based on the optimal heating base temperature

Once the optimal heating base temperature of a home is found, we compare it to the average outdoor temperature of a billing cycle to determine if space heating is in use. For each billing cycle, if the average outdoor temperature is higher than the heating base for a billing cycle, then the billing cycle is considered to be a non-heating month, implying no space heating used. Otherwise, if the average outdoor temperature is lower than the heating base for a billing cycle, the billing cycle is considered to be a heating month with space heating consumption. For the example case, the months of June, July, August, and September have average outdoor temperature above 57ºF, therefore these are the non-heating months. The rest of the months are heating months, as illustrated in Figure 11.

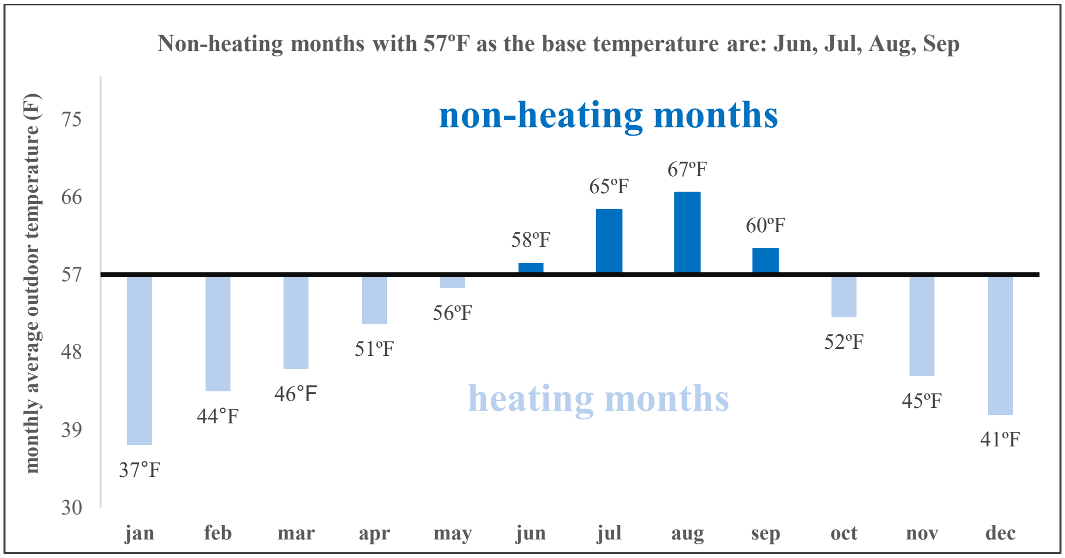

Figure 11: Identifying non-heating and heating months for the RBSA case

2.2.3 Step 3: Estimate average non-space heating baseload from the non-heating months

After we have identified the non-heating months, we can calculate the average daily baseload from the total consumption and total billing days of these months, and then use the daily baseload to estimate the average monthly baseload for the heating months. But for the non-heating months, the baseload is just the billing consumption of the billing cycles. Table 2 shows the total consumption *(Total_therms)* and billing days of each month *(Billing days)*. For the four non-heating months (Jun, Jul, Aug, and Sep), the monthly consumption values sum to 41.3 therms, and billing days account for 122 days, thus giving a daily baseload of 0.338 therms/day. This daily baseload is then used to estimate the average baseload of the heating months (Jan, Feb, Mar, Apr, May, Oct, Nov, Dec). Figure 12 shows the estimated baseload of each month. For the non-heating months, the baseload is just the original billing consumption, no adjustment is needed because we assume no space heating consumption in any of these months. If this is all we have to do with the baseload, then the space heating consumption estimate would simply be the total consumption minus this initial baseload estimate. But as mentioned previously, water heating is also impacted by weather, therefore, adjustment will need to be done on the initial baseload to account for the seasonality effects of water heating.

Table 2: Daily baseload calculation of an RBSA example

month	*Billing days*	*Total_therms*	*Non-heating month consumption*	*Non-heating month billing days*	*Non-heating billing*
Jan	31	132.7			
Feb	28	87.4			
Mar	31	76.1			
Apr	30	42.2			
May	31	25.4			
Jun	30	17.5	17.6	30	30
Jul	31	8.8	8.79	31	31
Aug	31	7.1	7.08	31	31
Sep	30	7.9	7.93	30	30
Oct	31	44.3			
Nov	30	81.0			
Dec	31	112.6			

Average daily baseload: 0.338

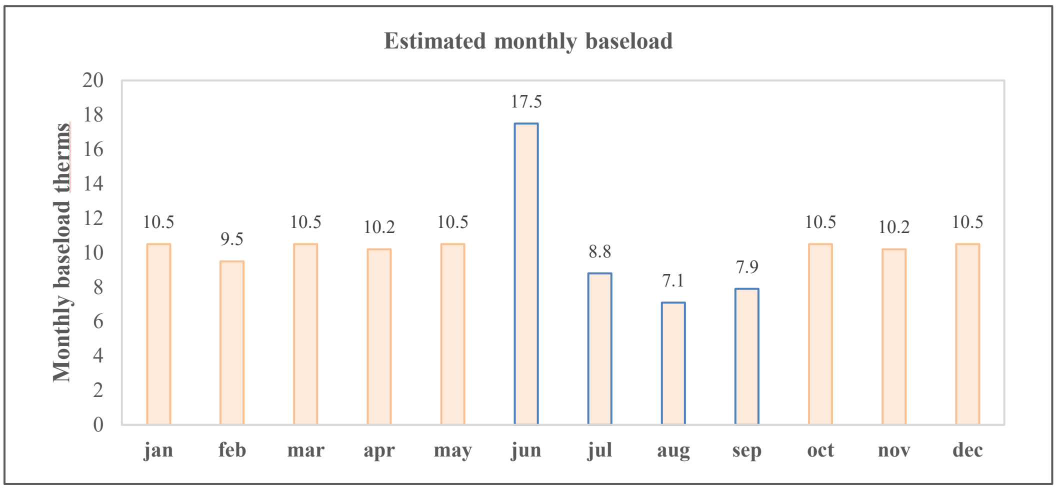

Figure 12: Estimated monthly baseload for the RBSA example

2.2.4 Step 4: Adjust the initial baseload for the seasonal effects of water heating

In a study of 171 homes in Central Florida, Masiello et al. (2002) found that "although water heating is not totally dominated by weather like space heating and cooling, these loads are still sensitive to temperature conditions". The side-by-side water heating systems study by Colon also shows seasonal effects on water heating (Colon, 2017). We also used the RBSA cases with water heating consumption to examine the relationship, and found that the cases also show a linear relationship between water heating energy consumption and temperature. Figure 13 and Figure 14 are regression model plots showing the linear relationship of two cases for example illustration.

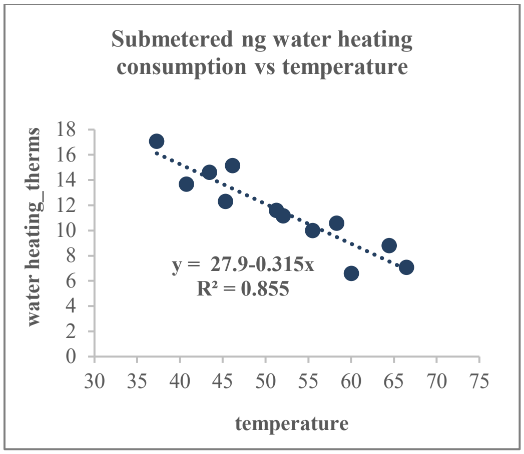

Figure 13: RBSA regression model case1 for water heating and temperature

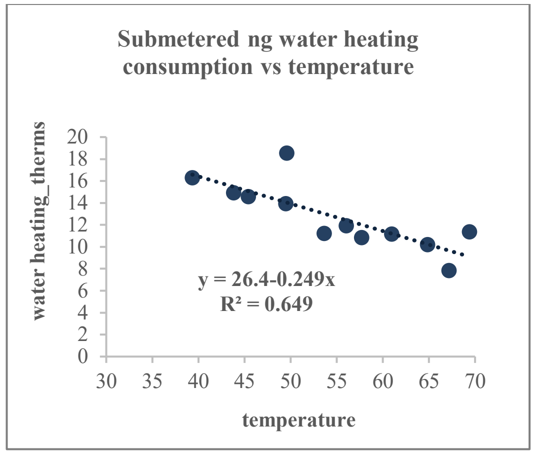

Figure 14: RBSA regression model case2 for water heating and temperature

To account for the seasonality effects of water heating, a methodology was developed to adjust for the seasonal effects:

- For the non-heating months, no adjustment is needed because all water heating is already accounted for in the initial baseload.
- For the heating months, the adjustment is done by multiplying a proportion of the initial baseload by the temperature difference between the heating base and outdoor temperature, as expressed in the formula:

$$rate \times initial\ baseload \times (T_{heating\ base} - T_{outdoor})\ .$$

Different methods were experimented with using the RBSA cases to find a reasonable value for the rate. One method is to use the individualized slopes of the corresponding regression models of the RBSA cases, this would be the ideal solution, but we cannot use this approach and apply individualized slopes for the RECS cases because we do not have water heating consumption in the RECS billing. Another method is to use the overall average slope among all the RBSA cases as a universal value and apply the same value to the RECS cases, but this would assume all the homes have the same seasonality effects of water heating, which is not likely given the variations seen in the literature. A third method is to assume that if a home has higher baseload consumption, then the seasonality effects of water heating consumption would be higher. Essentially, a percentage of the baseload might be more reasonable because the effect would vary from home to home based on household location and behaviors. This is depicted by the formula above. After some experimentation, it was found that 2% was a reasonable estimate. Therefore, the final methodology for the adjustment is: $2\% \times initial\ baseload \times (T_{heating\ base} - T_{outdoor})$. Figure 15 shows the initial baseload estimate and the estimate after water seasonality effects adjustment. Even though the adjustment size is small in relative to the total consumption, but it can be relatively large comparing to the initial baseload.

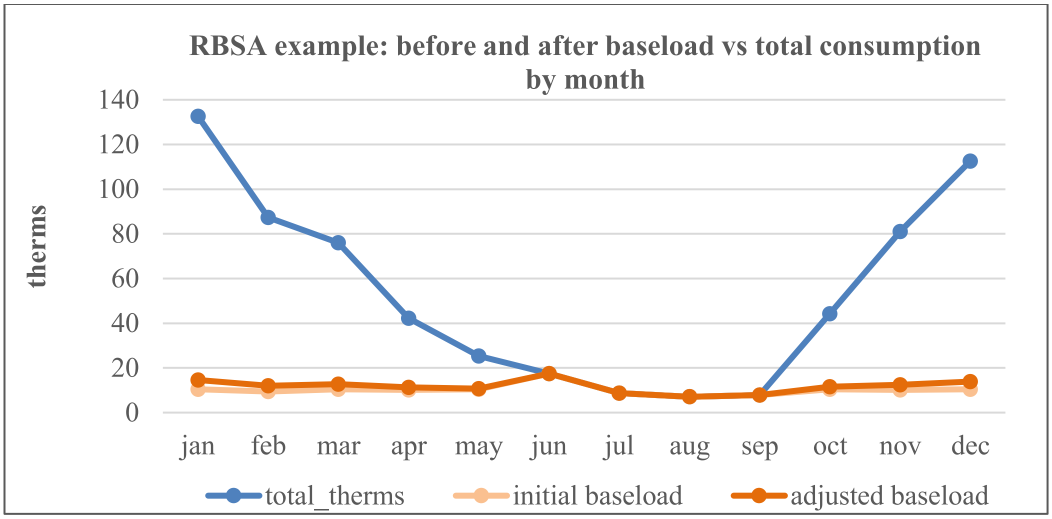

Figure 15: RBSA regression model example for water heating and temperature

2.2.4 Step 5: Calculate final space heating consumption of natural gas from total consumption and adjusted baseload

After estimating the initial baseload and adjusting for the seasonality effects of water heating, the final step is simply to obtain final space heating consumption by subtracting the initial baseload and water heating seasonality effects from the monthly total. Table 3 shows the monthly total consumption, the initial baseload, the seasonal effect of water heating, the adjusted baseload (the sum of the initial baseload and water heating seasonality adjustment), and the final estimated space heating consumption by month. Figures 16 and 17 are a bar chart and an area chart showing the estimated space heating consumption compared to the baseload and the total consumption. As seen in the charts, the adjusted baseload consumption is relatively small compared to the size of the total consumption or space heating consumption.

Table 3: Monthly billing estimation for space heating of an RBSA example

month	*Total_therms*	*Initial baseload*	*Water heating seasonality effect*	*Adjusted baseload*	*Estimated space heating*
1	132.7	10.5	4.13	14.6	118.0
2	87.4	9.5	2.56	12.0	75.4
3	76.1	10.5	2.28	12.8	63.3
4	42.2	10.2	1.16	11.3	30.9
5	25.4	10.5	0.31	10.8	14.6
6	17.5	17.5	0.0	17.5	0.0
7	8.8	8.8	0.0	8.8	0.0
8	7.1	7.1	0.0	7.1	0.0
9	7.9	7.9	0.0	7.9	0.0
10	44.3	10.5	1.03	11.5	32.7
11	81.0	10.2	2.37	12.5	68.5
12	112.6	10.5	3.41	13.9	98.7

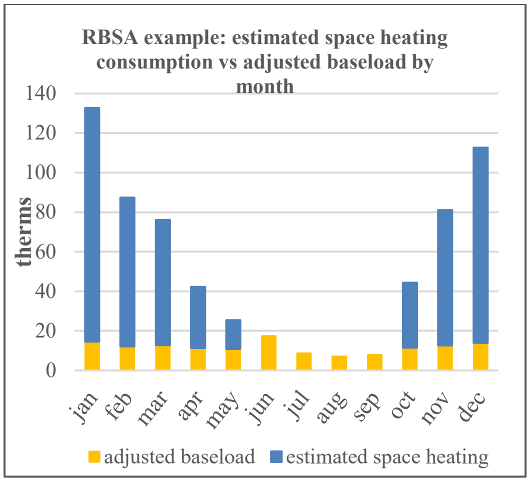

Figure 16: RBSA example: adjusted baseload vs estimated space heating

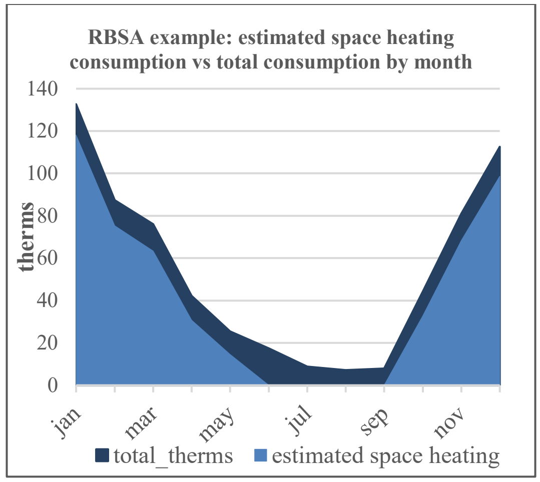

Figure 17: RBSA example: total consumption vs estimated space heating

3. Results

3.1 Comparison of the estimated space heating to the ground truth for the RBSA example

Figure 18 shows the estimated space heating to the ground truth by month for the RBSA example case. The darker colors are the estimated values, and the lighter colors are the ground-truth values. The results are quite close to each other except for the month of June, where the estimated space heating consumption shows zero while the ground truth shows partial space heating consumption was actually used. This is likely due to a few days in the month that were cold enough so that the heating system was on, but the "monthly billing" approach did not capture it because the average outdoor temperature of the month was compared to the base temperature. However, in terms of the consumption amount, it is relatively small.

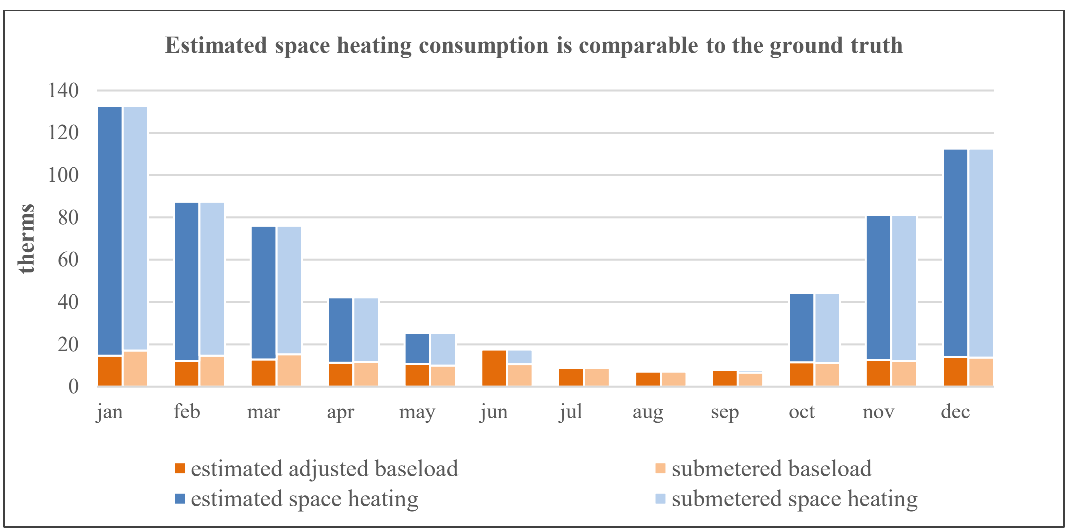

Figure 18: RBSA example: Estimated space heating vs submetered ground truth

3.2 Comparison of the estimated space heating to the ground truth for all the 31 RBSA cases

Based on the overall results comparing the 31 RBSA cases, we have high confidence in the robustness of the methodology. Percent error is used as the metric for comparison [(estimated space heating – ground truth)/ground truth] x 100. In Figure 19, if the blue dots are on the orange dashed line, it means the estimated values match the ground truth values. As it is shown, most of the cases are either on or very close to the orange line, which means most of the estimated results closed to the ground truth. In fact, among the 31 cases, 21 cases (68%) have their percent error within ±10%, 4 cases (13%) have their percent error between 10% and 15%, and the remaining 6 cases (19%) have their percent error within ±15%, as indicated by the darker dots in Figure 19.

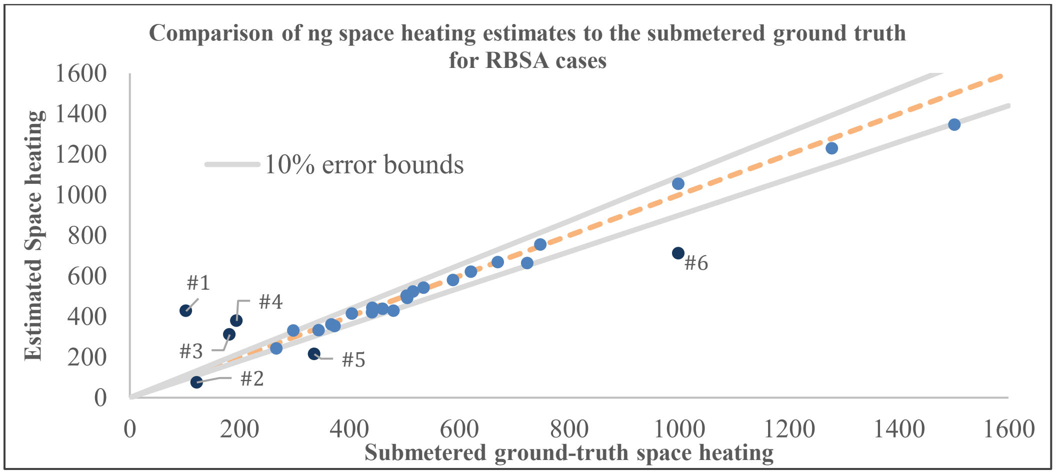

Figure 19: RBSA 31 cases: Estimated space heating vs submetered ground truth

After examining the ground truth of the six "bad" cases, there seems to be a general trend that we can identify. Five out the six cases have low space heating consumption, the size is relatively comparable to the size of the baseload, as shown from the first five cases in Figure 20. This might be an indication that if space heating consumption is small, relative to the size of the baseload, then the "monthly billing" approach is not as robust as those cases with relatively high space heating consumption.

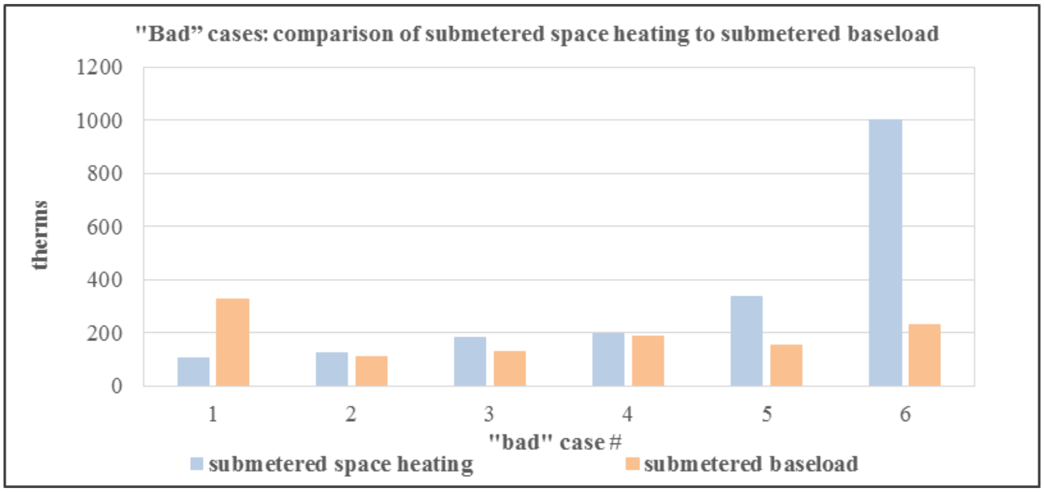

Figure 20: RBSA "bad" cases showing space heating consumption is comparable to that of baseload consumption

3.3 Comparison of "monthly billing" approach and "daily billing" approach

In the "monthly billing" approach, each case has 12 data points for regression modeling and space heating estimation. The same approach can be followed using "daily billing". As mentioned previously, the RBSA data were available in 15-minute intervals, and thus the data can be aggregated to daily totals for analysis. Instead of having 12 data points in the methodology, there will be 365 data points, which will necessitate using daily average outdoor temperatures instead of the average outdoor temperature of a month (billing cycle). Since this is capturing daily consumption, we assume it should improve the estimate for the "shoulder" months where there are some days cold enough to require space heating in a home. Figure 21 shows the estimated space heating consumption of the "daily billing" approach compared with the "monthly billing" approach from earlier. As is clear, the daily approach performs slightly better. Most of the "daily billing" results (in red circle) are closer to the ground truth than the "monthly billing" results (in blue circle); however, the "monthly billing" approach still performs relatively well, as evidenced by many of the red circles and blue circles overlapping one another.

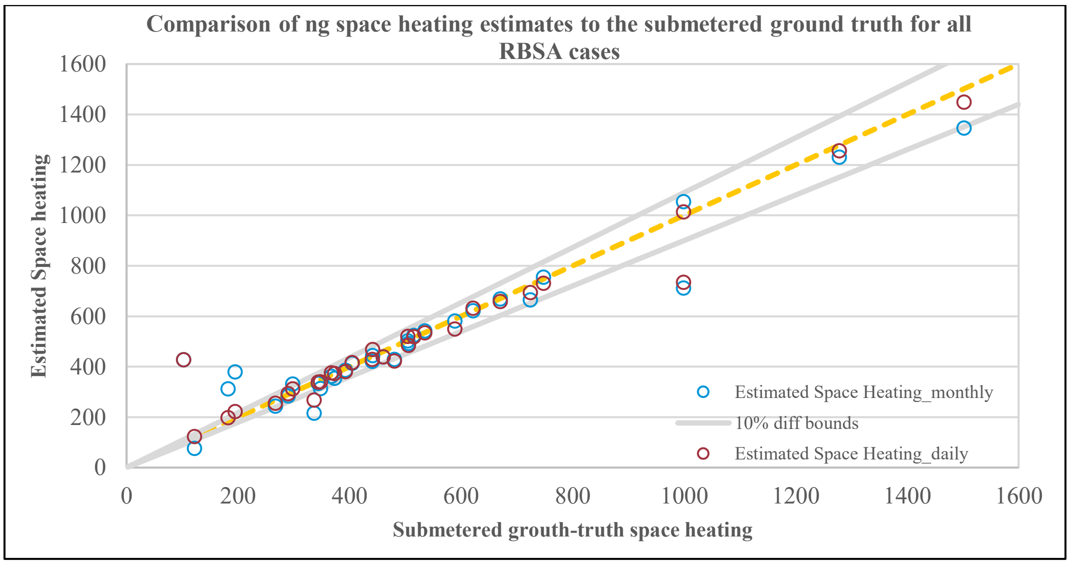

Figure 21: RBSA "monthly billing" and "daily billing" results compared to ground truth

3.4 Application to RECS

The "monthly billing" methodology described above can be applied to RECS cases that use natural gas for space heating. However, there are some aspects of the RECS cases that need to be handled differently, such as:

1) The usage indicators of natural gas space heating and water heating comes from survey respondents. Unlike in a submetering project where a measurement guarantees that an end use was in fact present and used in a home, in a survey we do not know the ground truth on the end uses of a home, because survey response errors exist (e.g., a respondent may not know whether his or her heating system is powered by natural gas or electricity). One way we can try to identify survey response errors is by checking the slopes and r^2 of the regression models---usually negative slopes or low r^2 values are an indication that the consumption data are not reasonably correlated with temperature, and therefore, that the reported

indicators of using natural gas space heating or water heating might be false when the respondents reported the use of natural gas, or vice versa.

2) There are cases that have partially missing bills or bimonthly issued bills, and therefore, regression modeling might not have enough data points to fit a good model, or might not perform well in finding a reliable base temperature. Billing data availability necessitates that these cases will have to be handled differently. Fortunately, there are not many of these kinds of cases. The methodology is sound for the majority of the cases.

3) Finally, there are RECS cases that do not have any billing data. Therefore, the estimation for these cases might rely on imputation strategy.

4. Further Research

As discussed in the previous sections, for some aspects of the RECS cases, such as for the cases where space heating consumption of a home is low, cases with partial missing billing data, or cases with complete missing billing data, modeling or imputation methodologies will need to be developed for these cases on the estimation of space heating consumption.

In addition, the model methodology to adjust for the seasonality effects in this paper is based on the submetered results of the 31 RBSA cases, which are mainly homes in the Pacific Northwest region, but the RECS is a national study. Hence, we should be careful extrapolating the method as it stands to the whole U.S., particularly the seasonality effects of water heating. If we become aware of additional natural gas submetered dataset we can study in the future, we will certainly do so and update our methodology as necessary. However, because the consumption of water heating is relatively small compare to that of space heating, any refined adjustments on the seasonality effects are assumed to be minor.

Finally, we are also interested in estimating electric space heating using mainly weather and billing data, however, based on some preliminary results using the "monthly billing" approach, it was found that the approach is not robust in finding a reliable base temperature. It is most likely due to: 1) there are more electric end uses that would influence the variation in the baseload, and 2) electric space cooling is another major source of consumption that is influenced by weather. Though, the "monthly billing" approach does not appear to be robust, it is still a possibility that the "daily billing" approach might be robust enough for the estimation of electric space heating consumption and space cooling consumption. This will be looked into further in the future when submetering data for homes with electric space heating and electric space cooling are available.

Acknowledgements

The authors would like to thank Chip Berry, Ian Mead, Hiro Minato, and Eileen O'Brien for their reviews and comments, special thanks to Chip Berry for providing valuable feedback on the structure and content of the initial draft; we also thank colleagues in the Office of Energy Demand and Integrated Statistics at EIA for giving useful feedback along the process of finalizing the JSM presentation.

References

Aguilar, C., White, D.J., and Ryan, D.L. 2005. *Domestic Water Heating and Water Heater Energy Consumption in Canada.*

Colon, C. 2017. *Side-by-side Testing of water Heating Systems: Results from the 2013-2014 Evolution*. Energy Efficiency & Renewable Energy, Department of Energy.

Fels, M.F. 1985. *The Princeton Scorekeeping Method: An Introduction.* The Center for Energy and Environmental Studies, Princeton University.

Fels, M.F., Hurvich, C.M. and Reynolds, C.L. (1986). *PRISM Research: Yesterday, today, and tomorrow*. ACEEE 1986, Panel 9.

Maguire, J., Fang, X. and Wilson, E. (2013). *Comparison of Advanced Residential Water Heating Technologies in the United States*. National Renewable Energy lab, Golden Co.

Goldberg, M.L. and Fels, M.F. (1986). *Taking PRISM One Step Further: The Refraction Method to Decompose Energy Savings*.

Goldner, F.S. 1994. *Energy Use and Domestic Hot Water Consumption*. The New York State Energy Research and Development Authority.

Masiello, J.A. and Parker, D.S., 2002. *Factors Influencing Water Heating Energy Use and Peak Demand in a Large Scale Residential Monitoring Study*. In ACEEE Proceedings, Panal1. Residential Buildings: Technologies, Design, Performance Analysis, and Building Industry Trends.

Northwest Energy Efficiency Alliance, 2014. *Residential Building Stock Assessment: Metering Study*. Report #E14-283. https://neea.org/img/documents/residential-building-stock-assessment-metering-study.pdf

Westman, E.D. and Monroe, W.H. 1986. *Performance of the Princeton Scorekeeping Method for Estimating Annual End-use Energy Shares as measured with Metered End-use Load Data*.

U.S. Energy Information Administration (EIA), 2018. *2015 Consumption and Expenditure Technical Documentation Summary*. https://www.eia.gov/consumption/residential/reports/2015/methodology_c&e/