Table of Contents

1 Introduction 1
2 Experimental Data 4
2.1 Data Challenges 5
3 Algorithm Methodology 6
3.1 Literature Survey 6
3.2 Model Development 8
3.3 Restricted Grey-Box Model With Infiltration and the "Static" Approach 9
3.4 Air Leakage Characterization 14
3.5 Results for Homes With One CT 15
4 Homes With Two CTs 20
4.1 Results for Homes With Two CTs 22
5 Energy Saving Predictions 24
5.1 Overall Post-Retrofit R-Value 24
5.2 Post-Retrofit ACH_{50} 27
5.3 Energy Saving Calculations 28
6 Randomized Controlled Trial 30
6.1 RCT Design 30
6.2 RCT Results 35
7 Use Cases 37
7.1 Cooling Season Considerations 37
7.2 EM&V for Retrofits 39
7.3 Cases Without Fuel Bills 42
8 Conclusions 45
References 47

List of Figures

Figure 1. Energy use intensity versus physical parameters characterizing retrofit opportunity 2
Figure 2. Correlations predicted by Eq. (5) and calculated from CT data for a furnace-heated home 11
Figure 3. Correlations predicted by Eq. (6) and calculated from CT data for a furnace-heated home 12
Figure 4. Correlations, Eq. (5) calculated by CT data for a boiler-heated home 14
Figure 5. Estimated and HEA-based ("ground truth") overall R-values for homes with single CT and either gas furnace or boiler 16
Figure 6. Estimated and HEA-based ("ground truth") overall R-values for homes with single CT and either gas furnace or boiler, by CT vendor 17
Figure 7. Estimated and HEA-based ACH_{50} values for 16 single-CT homes with blower-door test results available 17
Figure 8. Estimated and HEA-based ACH_{50} values for 16 single-CT homes with blower-door test results available 18
Figure 9. Averaged daily runtime prediction errors (absolute values shown) over the entire heating system for 25 homes with a single CT and furnace 19
Figure 10. Calculated average decay rate for single-CT homes with furnace from vendor #1, versus their overall R-value 20
Figure 11. Calculation of equivalent whole-building power and runtime under two alternative assumptions 21
Figure 12. R-values estimated from CT data versus HEA data for homes with two CTs 23
Figure 13. ACH_{50} values calculated from CT data versus blower-door test results for eight homes with two CTs 23
Figure 14. Runtime prediction errors over entire heating season for all homes with two CTs (boiler or furnace) 24
Figure 15. Experimental data on pre/post-retrofit ACH_{50} as measured in blower-door tests 27
Figure 16. Predictions of ΔACH_{50} modeled by Eq. (3) versus estimated ACH_{50} for 24 homes that have both estimated and blower-door ACH50 available 28
Figure 17. Estimated and HEA-based (GT) predictions for percent energy savings 29
Figure 18. Estimated and HEA-based (GT) predictions for percent energy savings from air sealing 30
Figure 19. Estimated heating to cooling capacity ratio for homes with both heating and cooling data available 39
Figure 20. Correlations for pre/post ECM implementation for a home from vendor #3, downstairs CT 40
Figure 21. Correlations for pre/post ECM implementation for a home from vendor #3, upstairs CT 41
Figure 22. Correlations for pre/post ECM implementation for a home from vendor #2 42
Figure 23. Calculated heat supply, Q, for a sample of two-story homes 43

Figure 24. Distribution of calculated heat supply for a sample of two-story homes with conditioned area of about 2,000 ft^2 44
Figure 25. Distributions of potential values of R-values and ACH_{50} for a home with unknown heat supply 45

List of Tables

Table 1. CT Data Fields Reported by Vendor 4
Table 2. Numbers of Homes With Complete Data Sets 5
Table 3. Distribution of Post-Retrofit Attic R-Values as Reported in ~350 HEAs 25
Table 4. Prediction of Estimated Post-Retrofit Overall R-Value Increase 27
Table 5. Randomized Controlled Trial Design 33
Table 6. Updated RCT Design 34
Table 7. RCT Results 36

1 Introduction

Many states have aggressive energy efficiency goals. Given that space heating and cooling account for over 40% of residential primary energy consumption (DOE/EIA 2018) and many residential buildings have relatively poor thermal performance, insulation and air sealing retrofits of existing homes can make major contributions to reaching these goals.

In the United States, utility energy efficiency programs subsidize and deliver a large portion of the retrofits implemented each year. These programs face two daunting challenges. First, due to their past success, many programs have growing energy savings goals. For example, in Massachusetts the gas savings goal increased from 1.12% of sales for 2013–2015 to 1.24% in 2016–2018 (Mass DPU 2016). Second, they need to deliver these energy-saving retrofits cost-effectively—i.e., the discounted value of the energy saved by a retrofit over its lifetime must exceed its first cost.

Realizable retrofit opportunities vary appreciably among homes. For example, the U.S. Department of Energy (DOE)/U.S. Energy Information Administration (EIA) Residential Energy Consumption Survey (DOE/EIA 2009) and Massachusetts Residential Appliance Saturation Survey (Opinion Dynamics Corporation 2009) indicate that about 20% to 25% of Massachusetts homes have significant insulation and/or heating system retrofit opportunities. Assuming typical savings of approximately 10% to 30% from basic insulation and air sealing retrofits indicates an energy savings potential on the order of 4% of total space heating and cooling energy consumption.

Delivery of insulation, air sealing, and heating system retrofits follows a multistep process that is often costly and challenging to scale. Customer acquisition occurs primarily through energy bill mailers, mass media and online advertising, and customer-initiated home energy assessment (HEA) requests that lack information about home-specific energy savings opportunities, expected energy savings, and cost-effectiveness. Once a customer requests an on-site HEA (which is free for ratepayers in some states, such as Massachusetts), the HEA must be scheduled and take place. Currently, an HEA involves a home visit by an energy service professional who conducts an extensive survey and a critical analysis of the household's conditions to identify and characterize energy savings opportunities (S3C 2019, BPI 2012). These assessments can be inconvenient to many homeowners, expensive (approximately $250–$500) for the program and are of variable accuracy. After the HEA, the homeowner must then decide whether to implement the recommended energy conservation measures (ECMs), and a majority do not. For example, in the nation's top-ranked residential energy efficiency programs by ACEEE (Massachusetts), only about 35% of HEAs ultimately result in major retrofits (Klint 2018). In other programs, the conversion rates can be significantly lower, e.g., typical rates in Minnesota are 7%–15% (Mark et al. 2016). This low closure rate increases the effective cost of program delivery. Finally, customers rarely get feedback on realized savings from ECMs beyond energy bills, while utility energy efficiency programs do not learn of potential large-scale field problems with ECMs until

after costly evaluation, measurement, and verification (EM&V) studies, years after ECM implementation.

In some cases (Blasnik 2018), utility energy efficiency programs select candidate retrofit homes by identifying homes with high space heating energy use intensity (EUI, in kBtu/ft^2) based on gas bills and the conditioned floor space. Using HEA data provided by our utility partners (see Section 2), we calculated the EUI and compared it with objective metrics indicating insulation retrofit (overall building envelope R-value) and air sealing (ACH_{50}) opportunities. Figure 1 shows the results: no noticeable correlations between EUI and the objective retrofit metrics.

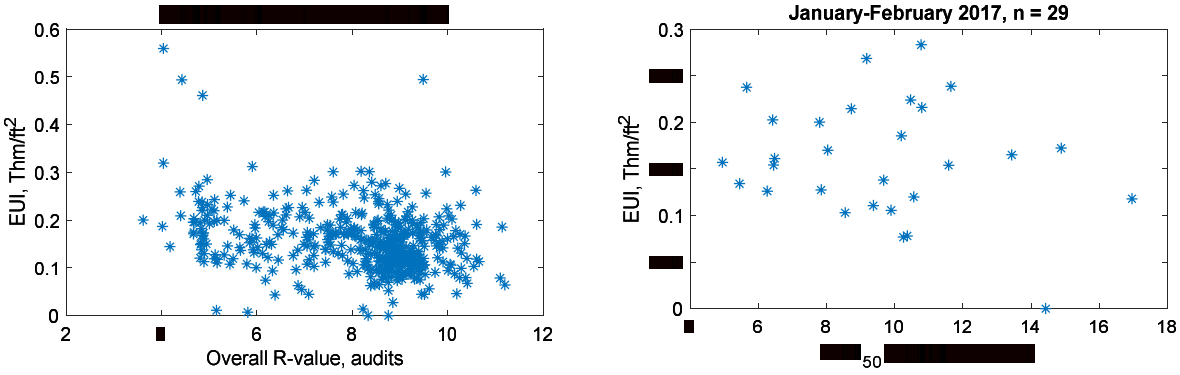

Figure 1. Energy use intensity versus physical parameters characterizing retrofit opportunity

The results shown were calculated for CT vendor #3 homes (see Table 2). Blower-door tests were performed for 29 homes.

In sum, the current approach makes it challenging to cost-effectively deliver the numerous retrofit opportunities that do exist; for example, around 1% of households in Massachusetts implement insulation; air sealing; or heating, ventilating, and air-conditioning (HVAC) retrofits each year.

Energy efficiency programs would benefit significantly from tools that improve each step of the retrofit delivery process, and emerging data sources provide the opportunity to develop such tools. Specifically, communicating thermostats (CTs) provide insights into heating system operations and building thermal response that, in turn, reflect building physical parameters corresponding to retrofit opportunities, i.e., R-value, air leakage, and heating system efficiency. CTs account for a significant portion of thermostat unit sales, and one market research firm projects that approximately 16.6% of U.S. homes with broadband service will have at least one CT by the end of 2021 (Barbour 2021). Moreover, many utilities provide energy efficiency incentives for purchasing CTs and, as a condition for providing the incentive, some obtain access to the CT data.

Therefore, there is an opportunity to significantly improve energy efficiency program effectiveness by using CT data combined with additional data available to utilities. What is necessary for this opportunity to materialize is a set of validated and scalable algorithms that automatically analyze CT data to accurately characterize home retrofit opportunities and predict expected retrofit energy savings.

In this project, Fraunhofer USA teamed with Eversource and National Grid to develop algorithms that analyze CT data combined with gas consumption data and basic home characteristics (floorspace, number of stories, and gas bills) available to utilities. The algorithms use these data to remotely and accurately identify the aforementioned energy efficiency opportunities to reduce residential space heating energy consumption for each individual home, as well as to estimate the home-specific energy savings potential. The team expects that using the algorithms in energy efficiency programs to provide targeted, customized, and actionable outreach to customers that are likely to have target retrofit opportunities will be significantly more compelling to customers than generic messages. In particular, the expected ultimate project outcomes are:

- Increasing the number of HEAs requested in homes with the target retrofit opportunities;
- Increasing the number of HEAs that result in implementation, thus, increasing the number of target retrofit measures implemented; and
- Ensuring that the retrofits deliver the expected savings (remote quality control).

The quantitative, measurable project objectives were to:

1. Identify the approximately 20% homes (per Opinion Dynamics Corporation 2009, DOE/EIA 2009) that would most benefit from at least one of the target ECMs to reduce space heating energy consumption: insulation and air sealing.
2. Predict the household-specific energy savings of the target ECMs within ±25% as compared to either the predicted energy savings from the energy audits, adjusted as appropriate for realization rates, or the actual energy savings obtained from implemented ECMs.
3. Double the participation in on-site energy audits through partner utility programs for the target households identified by the tool.

This report is organized as follows. Section 2 describes the data sets we obtained along with the associated data issues and challenges. Section 3 explains the physics-based algorithmic methodology that evolved from a simple inverse problem to a more elaborate approach that is successfully applied to homes with a single CT. This approach is extended further to incorporate homes with two CTs in Section 4. In Section 5, we develop and partially validate our approach to estimate prospective savings from ECM implementation. To measure the effect of the developed algorithms on HEA requests and implementations, we designed and conducted a randomized controlled trial (RCT) that is discussed in Section 6. In Section 7, we consider three immediate use cases extending the scope of this project: (1) analysis of cooling season data, (2) evaluation, measurement, and verification (EM&V) of ECM performance through comparison of pre- and post-ECM implementation CT data, and (3) application of the algorithms to homes heated with delivered fuels (i.e., evaluating algorithm effectiveness without gas bills). In the final section, we draw overall project conclusions and provide recommendations for further work.

In addition, a companion Best Practices Guide provides recommendations for effectively integrating the CT algorithms with energy efficiency programs (Roth and Zeifman 2020).

2 Experimental Data

We used anonymized interval CT and HEA home characteristics data from the Eversource and National Grid Home Energy Services program. The Home Energy Services program is the residential Mass Save program that delivers free home energy assessments, heavily subsidized CT installations, insulation and air sealing retrofits, and incentives for high-efficiency HVAC, water heating, and appliance retrofits (Home Energy Services 2018). To receive a CT rebate, households must sign a data waiver that grants the utilities access to the CT data.

Each data set for a home includes three pieces of information anonymized by the utility:

- CT data collected by the CT vendor (one of three) over a heating season, in CSV format
- HEA report in Excel spreadsheet format, performed by the HEA vendor (same vendor for all homes)
- Monthly utility gas bills in Excel spreadsheet format, overlapping with the CT data in time (number of bills varies from 3 to 24 per home depending on availability).

The CT data characteristics[1] differ among vendors. For example, vendor #1 records all data fields every 5 minutes, whereas vendors #2 and #3 record all data fields as soon as there is a change in at least one of the fields. Table 1 describes some of the data fields in detail. Additional data fields that were not used in this work (e.g., indoor humidity) are not shown. Vendor #1 also has a data field for "stage," to report what furnace stage(s) are running, although that field often remained unpopulated (i.e., for single-stage device).

Table 1. CT Data Fields Reported by Vendor

CT Vendor/Data Field	HVAC Status	Room Temp	Outdoor Temp	Wind Speed	Time Stamp
Vendor #1	Duration of "on" time over 5-min interval, 1-s resolution	Average over 5-min interval, 0.1°F resolution	Average over 1-h interval (from nearest weather station), 0.1°F resolution	Average over 1-h interval (from nearest weather station), 1 km/h resolution	Every 5 min
Vendor #2	"On" or "off," reported only when any	Unclear if averaged, 1°F resolution, reported only	Average over 1-h interval (from nearest weather	N/a	Reported for every change in any data

[1] Note that due to a long chain of communications, not all our inquiries were answered by the CT/HEA vendors. The process of data collection and transferring took almost two years to complete.

	data field changes	when any data field changes	station), 1°F resolution		field, 1-s resolution
Vendor #3	"On" or "off," reported only when any data field changes	Unclear if averaged, 1°F resolution, reported only when any data field changes	Average over 1-h interval (from nearest weather station), 1°F resolution	N/a	Reported for every change in any data field, 1-s resolution

To be useful for this project, combined data for each home must include at least one month of CT data collected during a heating season along with coincident gas bills and the HEA data. Because of this completeness requirement, we could only use some of the obtained CT data sets.

Table 2 lists the numbers of complete home data sets we received. In total, we obtained complete data sets for about 450 Massachusetts homes.

Table 2. Numbers of Homes With Complete Data Sets

CT Vendor	Furnace and 1 CT	Furnace and 2 CTs	Boiler and 1 CT	Boiler and 2 CTs
Vendor #1	17	10	4	2
Vendor #2, National Grid	84	42	24	15
Vendor #3, Eversource	79	69	28	83
Total	180	121	56	100

2.1 Data Challenges

Similar to other researchers (Siemann 2013), we noticed several problems with CT data. The most frequent problem is missing or unreported HVAC runtime. Unlike indoor or ambient temperature, missing runtime cannot be interpolated. Moreover, algorithmic identification of missing runtime can be a daunting task; for example, rising indoor temperature may or may not be indicative of missing runtime (it could indicate a supplemental heating system). Another problem is an apparent time lag of up to 10 minutes between the HVAC status and change in indoor temperature reported; such a lag seems excessive for furnace-heated homes. The outdoor temperature data were mostly missing for vendors #2 and #3, in those cases we replaced it with the temperature observed at the closest weather station (same for wind). Finally, the 1°F resolution of temperature data from vendors #2 and #3 is too coarse to be used for a conventional grey-box model identification technique (see Section 3).

HEA reports (also known as audit results) comprise numerous data fields; however, not all of them were populated in the anonymized digital files we received. The populated fields included

basic home sizing information (ceiling height, conditioned floor space, number of floors, total volume, attic area, wall area), overall R-values for walls and attic, U-factor for windows, infiltration rate in CFM_{50}, and basic HVAC information (heating system type, atmospheric or condensing heating system, fuel type, heating- and distribution-system efficiencies, and heating system capacity). The unpopulated fields in some HEA reports included, e.g., window area, and floor space.

We obtained a description of HEA procedure underlying these reports and also analyzed the HEA data and compared them to the corresponding interval data and gas bills. These analyzes suggest different levels of reliability for the audit results. In particular, we are confident in home size information and in overall R-values for walls and attics/roofs and in U-values for windows. In contrast, heating system capacity is an estimated parameter that is not consistently accurate, because the discrepancy between the audit value and the value calculated from gas bills/runtime information can differ by a factor of up to 10. The infiltration rate is also estimated based on a qualitative on-site assessment, making it highly unreliable as well (e.g., blower-door test results that we obtained for some audited homes varied by a factor of up to four from the audit values).

3 Algorithm Methodology

3.1 Literature Survey

Prior to this project, we reviewed the state-of-the-art for remote HEA approaches using interval (e.g., CT) data (Zeifman and Roth 2016). Although there are currently no widely accepted methods capable of characterizing the insulation, air sealing, and/or heating system retrofit opportunities for homes at scale, the potential approaches should be based on predictive models connecting the data inputs with retrofit-characterization outputs (Gaasch et al. 2014). Major retrofit opportunities can be characterized by physical home parameters such as the overall envelope R-value, ACH_{50}, and HVAC efficiency, and models capable of predicting these parameters can be loosely divided into white-box, grey-box, and black-box categories (Afram and Janabi-Sharifi 2015; Berthou et al. 2014).

White-box models are very detailed and accurate physics-based simulation tools, e.g., EnergyPlus™. Because these models typically require hundreds of parameters to describe a single building, both setting up the model and estimation of its parameters from experimental data (i.e., calibration) to characterize the retrofit opportunities are time-consuming and, sometimes, ill-posed tasks, making the white-box models difficult to scale.[2]

In this application, black-box models rely on large training data sets and machine learning techniques to estimate building physical parameters and/or classify buildings by their retrofit opportunities (e.g., Pathak et al. 2019). Because these models do not have a physical basis, their predictive ability is limited and restricted to homes whose characteristics are represented by

[2] We are aware of efforts to "autotune" EnergyPlus (see New et al. 2012), but this technology is still at early stage.

those in the training data set. Thanks to their simplicity, these models can scale fairly easily, but only if appropriate and large training data sets exist.

Grey-box models use relatively coarse-grained physical models (typically, lumped models) with just a few parameters. Although these models seem to combine the advantages of the other two model categories (i.e., the predictive ability of physics-based white-box models and the scalability of the black-box models), they are inherently coarse; consequently, the estimated building parameters may not precisely match the actual physical building parameters.

We concluded that grey-box models are most suitable for remote HEAs, as they combine the physics-based predictive ability of white-box models and the scalability of black-box models.

A first-order model is the simplest grey-box model, and, unsurprisingly, several research groups (Goldman 2014; Newsham et al. 2017; Chong and George 2018) proposed a retrofit-characterizing measure based on a first-order grey-box model. This measure is related to *cooling gradient* (i.e., temperature slope during thermostat setback periods) and its dependence on the ambient temperature.

The first-order grey-box model implies a single lumped heat balance equation for a home, e.g.,

$$M_t \frac{dT_r}{dt} = Q_{HVAC} + q_{\frac{int}{ext}} + A_w U(T_a - T_r) \qquad (1)$$

where variables T_r and T_a are indoor (room) and outdoor temperatures, A_w is overall area of the external surfaces like walls, windows, roof/attic, foundation (i.e., building envelope), U is integrated envelope heat loss characteristic due to conductive and infiltrative processes (Newsham et al. 2017), M_t is thermal mass characteristic (Newsham et al. 2017), Q_{HVAC} is HVAC heat supply, and $q_{\text{int/ext}}$ are additional internal (e.g., from appliances and people) and external (e.g., solar) heat gains.

When the heating system does not run for extended periods of time, such as during thermostat setbacks, $Q_{\text{HVAC}} = 0$ and Eq. (1) has a simple exponential decay function as a closed-form solution for indoor temperature T_r, if the additional heat gains are neglected (Chong and George 2018) and if the integral heat loss characteristic U is constant. The "time constant" of this solution (i.e., inverse of the decay rate) is proposed by Chong and George (2018) to serve as ultimate measure of "the leakiness of a building." Although this measure seems enticing for HEA, Newsham et al. (2017) are cautious in their analysis, suggesting that the differences in the cooling rates between different buildings, as well as the variability of the cooling rates of the same building, could be attributed to the differences in envelope insulation, air leakage rate, envelope area, and/or thermal mass.

First-order grey-box models are known to yield relatively poor accuracy in predicting building thermal response, whereas second- and higher-order models usually offer satisfactory accuracy (Mejri et al. 2011). Moreover, the meaning of some physical parameters of the first-order model is unclear and may be misleading (e.g., the "thermal mass characteristic" M_t in Eq. (1) lumps together the heat capacitances of building envelope and those of the internal space). Lastly,

although in principle the air leakage heat loss can be separated from the heat conductance loss by introducing an additional air leakage heat loss term in Eq. (1), incorporating the nonlinear stack effect for air leakage (see next section) will make the differential equation nonlinear, thus invalidating the simple exponential solution along with its "time constant."

Accordingly, we decided to use a second-order grey-box model as a basis for development of remote HEA methods.

3.2 Model Development

Our grey-box model incorporates characteristics of building insulation and airtightness as well as the building thermal mass by using two capacitances (indoor space and lumped envelope) and three thermal resistances (two identical resistances for external and internal surfaces of the lumped envelope and one for convection induced by air infiltration). The lumped envelope is a single value representing a building's entire envelope, including opaque walls, windows, roof/attic, and foundation. Note that in our earlier work (Zeifman and Roth 2016), we did not split the external wall into two resistances, which led to a nonphysical factor of two in matching the estimated and HEA-based R-values. The proposed balance equations are:

$$C_r \frac{dT_r}{dt} = Q_{HVAC} + q_{int} + \frac{A_w}{\frac{R_w}{2}}(T_w - T_r) + q_{inf} \qquad (2)$$

$$C_w \frac{dT_w}{dt} = (A_w/(R_w/2)) \times (T_r - T_w) + A_w/(R_w/2))(T_a - T_w) + q_{ext} \qquad (3)$$

where variables T_r, T_w, T_a are, respectively, indoor, lumped envelope, and outdoor temperatures, R_w and A_w are overall R-value and area of the lumped envelope, C_w is overall thermal capacitance of the exterior surfaces, C_r is overall heat capacitance of the internal space, Q_{HVAC} is HVAC heat supply, q_{int} is internal heat gains/losses affecting directly T_r, q_{ext} represents radiative (foremost solar) or rain/sleet/snow-related heat transfer between the outside of the building enclosure and the outdoor environment, and q_{inf} is heat loss due to air infiltration. The latter factor accounts for, on average, about 25–30% of U.S. single-family space-heating loads in heating-dominated climate zones (Huang et al. 1999). Among these variables, only T_r and HVAC on/off status (embedded in Q_{HVAC}) are directly sensed by a CT.

A conventional way to calibrate the model, i.e., to estimate the model parameters, is to discretize the Eqs. (2) and (3) and minimize the difference between the observed and modeled state variable, T_r(t) (Lin et al. 2012; Bacher and Madsen 2011; Siemann 2013; Harish and Kumar 2016). This inverse problem can be ill-posed in the sense that small variations of the observed variable lead to large changes of the parameter estimates. To alleviate this problem, we use the derived closed-form solution that does not require the discretization step (Zeifman and Roth 2016). In a more recent work, we applied this method to CT data sets from several homes and compared the estimated overall R-values with qualitative self-assessments of the home insulation levels (Zeifman, Roth, and Urban 2017). Although that comparison showed good

correspondence between the estimated and HEA data, several limitations make it difficult to scale-up this initial approach:

1. **Nighttime:** To reduce the effect of non-HVAC heat gains/losses that are difficult to model, we restricted the CT data to nighttime only. This restriction, however, can lead to an overfitting problem (Lin et al. 2012) because of the lack of system excitation during a given night. Combining numerous nighttime data segments together could potentially mitigate this problem, yet the number of parameters to be identified (e.g., initial lumped wall temperature for each nighttime segment) grows proportionally to the number of nights, making the problem computationally intractable.
2. **Zone solution:** Eqs. (2)-(3) are applicable to a single thermal zone in a residential building, but even a single-thermostat home does not necessarily comprise a single thermal zone. Accordingly, the second-order grey-box model, Eqs. (2)-(3), can be too coarse to describe the thermal response of actual homes.
3. **Nonlinear air leakage:** While the wind-driven component of the air infiltration phenomenon is linear with $(T_r - T_a)$, meaning its incorporation still allows for the closed-form solution of Eqs. (2)-(3), the stack effect varies nonlinearly with $(T_r - T_a)$ (Younes 2012). As a result, no closed-form solution can be derived for a second-order grey-box model that incorporates both wind-driven and stack components of air infiltration.
4. **On/off heating:** Originally, we modeled the HVAC heat supply Q_{HVAC} as a two-state variable that has a fixed value whenever the HVAC is called on, and is zero otherwise. This model is appropriate for a single-stage furnace, and in principle it can be extended to accommodate a multi-stage furnace. However, it is unclear if the on-off or discrete heating model also works for boilers, especially for condensing boilers that can modulate extensively. In addition, the *timing* of heat delivery can be delayed for boilers with more massive (e.g., cast-iron) radiators and for steam-based systems.

Due to the air leakage nonlinearity limitation, we used a commercial software package, the MATLAB Grey-box toolbox (Ljung 2017), that is specifically designed for implementation and identification of arbitrary grey-box models. Overcoming other challenges is explained in the next section.

3.3 Restricted Grey-Box Model With Infiltration and the "Static" Approach

Accurate physics-based modeling of air infiltration requires addressing both wind and stack effects (see previous section), which can lead to cumbersome mathematical terms with numerous fitting parameters (Walker and Wilson 1990). On the other hand, given the coarseness of the second-order grey-box model and limited experimental data available to us (e.g., no local home-specific weather data available), empirical relationships can be as accurate but more practical for implementation. Accordingly, in this work we use Walker's (2017) empirical model with just two fitting parameters:

$$q_{inf} = -\rho_{air} c_{p,air} (C_1 W^{2.6} + C_2 |T_a - T_r|^{1.3})^{0.5} (T_r - T_a) \quad (4)$$

where the first two terms designate air density and air heat capacitance, respectively, W the wind speed, and C_1 and C_2 are fitting parameters.

To address the zone challenge of the initial approach (see previous section), we extended the conventional grey-box model by restricting its search space. The main assumption is that Q_{HVAC} is evenly distributed over the interior floor space of the residential single-family building and that zone temperature dynamics follow those of the "average" indoor temperature. We think this assumption mainly applies to homes controlled by a single thermostat and heated with furnaces, but can also correspond to boiler-heated homes with a single thermostat. The basic modeling idea then is that the experimental indoor temperature curve is no longer considered to be the "best" solution to which a grey-box model's solution is conventionally fitted for parameter identification (Lazrak and Zeifman 2017). Rather, the parameters of Eqs. (2)-(4) are estimated by fitting the model prediction to an unknown yet "best" second-order solution, i.e., a hypothetical curve that may differ from the individual experimental ones. Although such a curve is unknown, we can assess some parameters that define this curve using overall approximated correlations. Such correlations can, in turn, yield confidence intervals for these parameters that we propose to use to restrict the search space in the conventional grey-box model identification.

How do we obtain those correlations? First, let us integrate Eqs. (2) and (3) over a relatively long period of time τ, so that the initial and final values of indoor (and wall) temperatures are approximately the same:

$$0 = \bar{Q} \sum t_{on} + \left(\overline{q_{int}} + \overline{0.5 q_{ext}}\right)\tau - \frac{A_w(\overline{T_r} - \overline{T_a})\tau}{R_w} + \overline{q_{inf}}\tau \quad (5)$$

where the bar designates averaging over time and Q is the HVAC heat supply at state "on" (zero heat supply in the "off" state). It is easy to see that this "static" Eq. (5) is an energy conservation equation and also is the well-known PRISM model (Fels 1986). The statistical confidence interval for the slope in this linear regression can be used to restrict the search space for $R_w \bar{Q}$.

Figure 2 shows an example of a correlation between the "on" time and indoor-outdoor temperature difference, predicted by Eq. (5), for a home. It can be seen in the figure that better correlations (scatter-wise) occur for τ, ranging from 24 hours to several days.

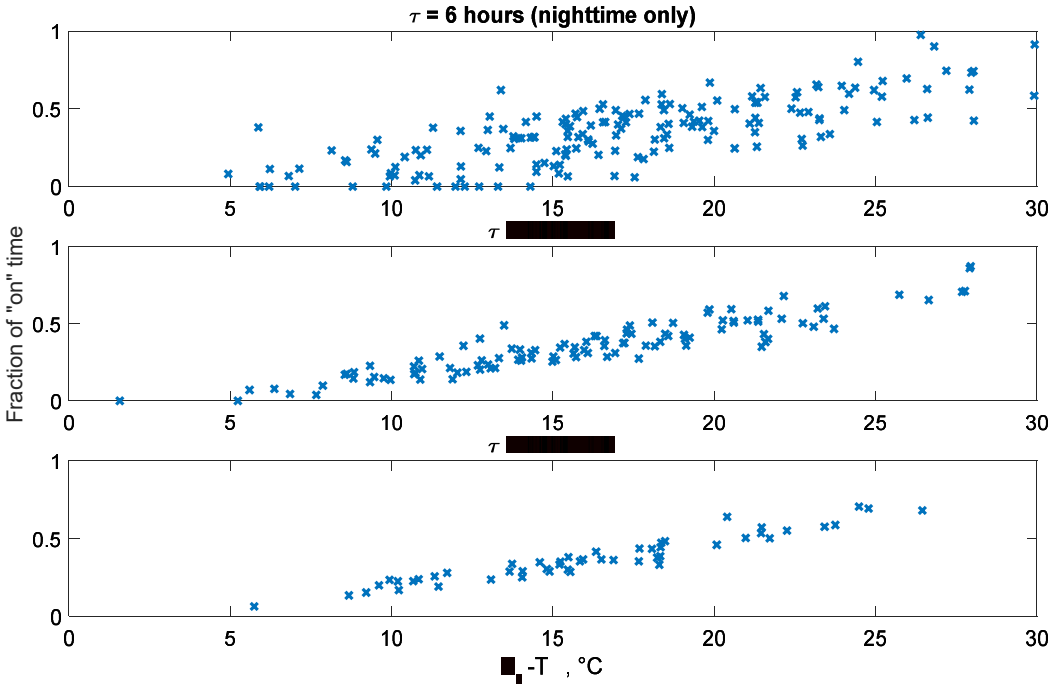

Figure 2. Correlations predicted by Eq. (5) and calculated from CT data for a furnace-heated home

We can also add a second, "dynamic" correlation by considering dependence of the room temperature gradient on ambient temperature. Unlike the cooling gradient discussed earlier (see section 3.1), this one is a *heating* gradient. Many U.S. residential heating systems, particularly furnaces, are sized to enable quick temperature recovery, meaning they have significantly more capacity than design loads (Brand and Rose 2012), so that the *heating* curves are often much more linear in time than the cooling curves. Accordingly, Eq. (2) can be approximated by a difference equation over the "on" portion of heating cycle

$$\frac{\Delta T_r}{t_{on}} \approx \frac{Q}{C_r} + \frac{q_{int}}{C_r} + \frac{q_{ext}}{2C_r} - \frac{A_w}{R_w C_r}(T_r - T_a + \Delta T_w) + \frac{q_{inf}}{C_r} \qquad (6)$$

where ΔT_r is the room temperature gain during the "on" portion of the heating cycle and ΔT_w is the difference between the actual lumped wall temperature and the steady-state lumped wall temperature, obtained from Eq. 3. The latter variable (ΔT_w) is a manifestation of building envelope's thermal mass that cannot be explained by the first-order model, Eq. (1).

Figure 3 shows experimental correlations of type Eq. (6), calculated for the same home used for Fig. 1. The significant scatter, observable in the figure, forms a characteristic parallelogram structure. Because the scatter does not go to zero during nighttime, we attribute it, at least in part, to the difference between the lumped wall temperature and the "equilibrium" lumped wall temperature (i.e., the thermal mass effect ΔT_w) in Eq. (6).

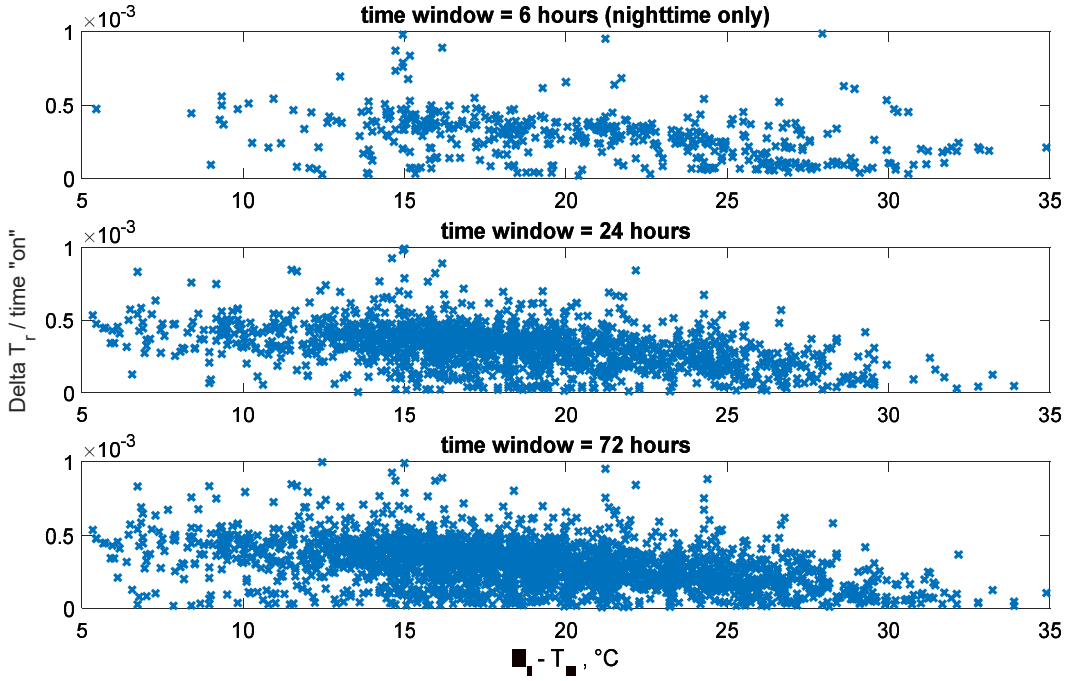

Figure 3. Correlations predicted by Eq. (6) and calculated from CT data for a furnace-heated home

To build each plot, we used a moving time window and included in each time window all incidences of furnace at state "on," each point represents the heating rate over one incidence; T_r-T_a is averaged over each time window.

In principle, if the overall heat capacitance of the internal space C_r were available or could be accurately estimated for a home, the slopes of the two correlations, i.e., $A_w/(R_wQ)$ for Eq. (5) and $A_w/(R_wC_r)$ for Eq. (6), would uniquely define the R-value (R_w) and HVAC heat supply (Q) for the case of negligible air leakage. Similarly, in case of the non-negligible air leakage, analogous estimations would be possible by nonlinear curve fitting using Eq. (4). Whereas Figure 2 suggests relatively narrow confidence interval for Eq. (5)'s slope, Figure 3 suggests that the confidence interval for Eq. (6)'s slope is rather large; yet further processing by MATLAB toolbox would yield reasonably accurate estimates for the home physical parameters.

However, because C_r includes the heat capacitance of the internal air as well as the furniture, carpets and other household contents and internal surfaces and structure,[3] its calculation is not straightforward. Although some semi-empirical formulas are available in the literature (e.g., Berthou 2013), our experimental results do not support them (Zeifman, Lazrak, and Roth 2018).

Accordingly, in this work, we limited the correlations to Eq. (5). We found, however, that having only one correlation, Eq. (5), available was not helpful for the proposed restricted grey-box model that used nighttime data only: The restricted grey-box model usually yielded parameter estimates that were very close to the starting parameter values sampled from the restricted parameter space and fluctuated drastically from night to night. The standard remedy to this

[3] Some authors refer to C_r as "internal thermal mass" (Lee and Hong 2017).

overfitting problem is to extend the time to larger periods, such as to several weeks to get the system sufficient excitation (e.g., Lin et al. 2012). Given that such extension (see above) implies accurate modeling of external and internal heat gains and that, among the experimental homes, there are relatively few equipped with "high-resolution" CTs (i.e., from vendor #1, see Table 1, Table 2), we decided not to pursue this approach.

For actual home HVAC systems, the range of efficiencies is relatively small. Hallinan et al. (2011) suggest the efficiencies range from 70% (worst case) to 95% (best case) for residential furnaces and boilers. Therefore, as a practical alternative, we assumed the same HVAC efficiency of 80%[4] for all experimental homes, which, coupled with an air infiltration estimation procedure, permits direct estimation of the overall R-value using an estimate for $\bar{Q}$ (from the fuel bills and HVAC runtime) and the value of slope from the correlation, Eq. (5). The air infiltration estimation procedure is based on an analysis of salient CT data points and will be discussed in the next section.

Therefore, unlike the conventional grey-box model calibration methods that estimate model parameters using a dynamic time series of state variables, e.g., T_r (Lin et al. 2012, Bacher and Madsen 2011), the proposed method is static. The method is suitable for homes equipped with a single thermostat and a furnace. To extend this method to boiler-based HVAC systems, we note that boiler systems are two-stage systems (Peeters et al. 2018). That is, water is heated to a target temperature (controlled by an aquastat) by a burner and then is pumped to the emitters (radiators or convectors) when a thermostat calls for heat. This two-stage process implies two additional heat balance equations, one for heating water with a boiler and one for the radiator heat exchange. Fortunately, under simplifying assumptions, integration of these four equations—similar to integration of Eqs. (2) and (3)—yields an equation similar to Eq. (5).

The main simplifying assumption is that for a boiler with fixed Q_{HVAC}, the time "on" as reported by a CT roughly equals a constant fraction, χ, of the burner on time .[5] Because $\bar{Q} \sum t_{on}$ is the only term with time "on" in Eq. (5), and because we calculate $\bar{Q}$ using the gas bills and also time "on," χ will cancel out in this term, and we can use the reported by CT time "on" as a proxy for the burner time "on" in Eq. (5). Violations of the simplifying assumptions would result in nonlinearity in the time "on"—temperature difference correlations, Eq. (5). Likewise, CT data from a properly configured modulating boiler (usually a condensing boiler) would also have a significant nonlinearity in these correlations, with a higher ratio of time "on" at warmer outdoor temperatures due to modulation of water circulation temperatures as a function of T_a.

Figure 4 shows an example of correlations, Eq. (5) for a boiler-heated home. The experimental correlations in this figure do not practically differ from those of a furnace-heated home (see

[4] In practice, estimated HVAC efficiencies reported in audits ranged from 78% to 82% for 67% of homes in this study. These values exclude distribution efficiency.
[5] This reflects that boilers are often oversized relative to peak loads to facilitate recover from temperature setbacks.

Figure 2). We did not observe significant nonlinearities in such correlations for other boiler-heated homes in this project, except those caused by missing data.

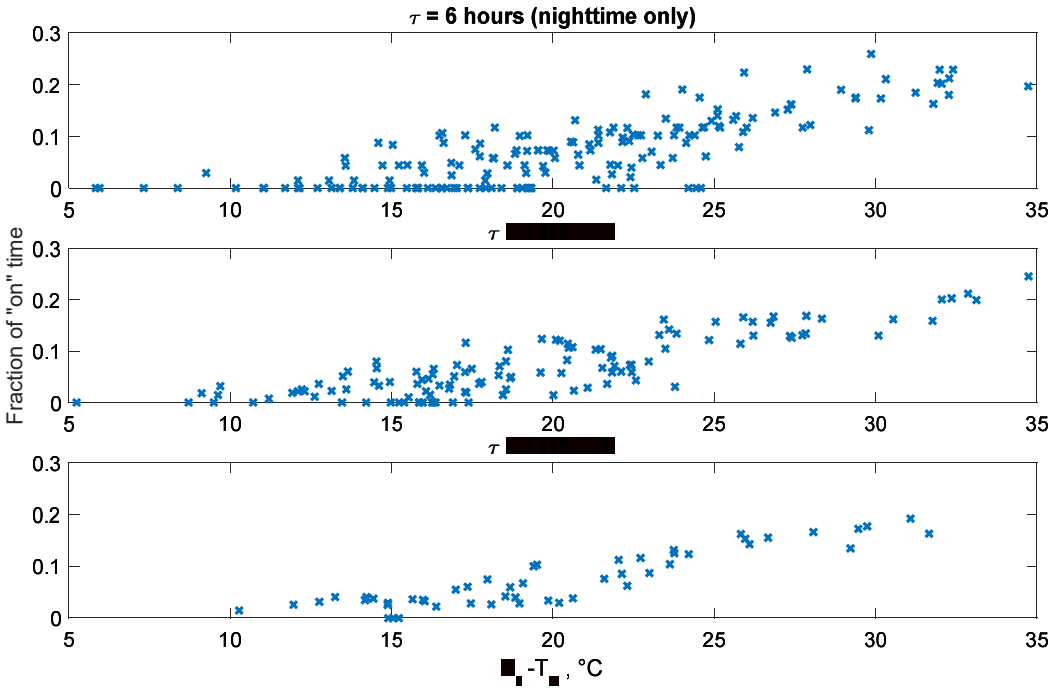

Figure 4. Correlations, Eq. (5) calculated by CT data for a boiler-heated home

3.4 Air Leakage Characterization

Our approach uses approximations and integrations to estimate the two air leakage parameters C_1 and C_2 in Eq. (4) and then to calculate the ACH_{50} for a home. The key idea is to compare HVAC runtimes for time windows with essentially different wind speeds but similar otherwise (e.g., in terms of time of the day, inside/outside temperatures).

Suppose we use two different time windows of same duration τ and with same average temperatures (T_r and T_a), same internal heat gains (usually this is approximately true if the time of the day is the same), but different wind speeds W_1 and W_2. Assume $W_2 \approx 0$ and W_1 to be close to the maximum wind speed value over the heating season. From Eqs. (4) and (5), we get

$$\frac{\sum t_{on1} - \sum t_{on2}}{\tau} = \frac{\left(C_1 W_1^{2.6} + C_2 |T_a - T_r|^{1.3}\right)^{0.5} (T_r - T_a)}{Q} - \frac{\left(C_1 W_2^{2.6} + C_2 |T_a - T_r|^{1.3}\right)^{0.5} (T_r - T_a)}{Q} \quad (7)$$

Walker (2017) indicated that, typically, wind-based and stack effect-based infiltration have similar energy impacts over the course of the heating season. Therefore, because the maximum wind speed is much higher than the average wind speed, $C_1 W_1^{2.6} \gg C_2 |T_a - T_r|^{1.3}$ and Eq. (7) takes the following approximate form:

$$\frac{\sum t_{on1} - \sum t_{on2}}{\tau} \approx \frac{\left(C_1 W_1^{2.6}\right)^{0.5} (T_r - T_a)}{Q} \left(1 - \left(\frac{W_2}{W_1}\right)^{2.6}\right) \quad (8)$$

We can then derive the parameter C_1 from Eq. (8):

$$C_1 \approx \left[\frac{(\sum t_{on1} - \sum t_{on2})Q}{\tau(T_r - T_a)}\right]^2 \frac{1}{W_1^{2.6}\left(1-\left(\frac{W_2}{W_1}\right)^{2.6}\right)^2} \quad (9)$$

In this way, wind-based infiltration is fully characterized. For the stack effect, we can assume the approximate equality, which is expressed by

$$C_1\left(\overline{W^{2.6}}\right) = \beta C_2\left[\overline{(T_r - T_a)^{1.3}}\right] \quad (10)$$

where the bar designates averaging over entire heating season and β is a fitting parameter of the order of unity.

In practice, we estimate the parameter C_1 for a home using all available pairs of the similar time windows with high/low wind speed in a home's CT data set.[6] Once C_1 is estimated, we use the overall daily correlations, Eq. (5), and a set of "possible" values of β = {0.5, 1.0, and 2.0} to calculate the corresponding set of values of parameter C_2 and then to estimate the value of R_w and the goodness-of-fit (e.g., the sum of errors squared) by least-square curve fitting to Eq. (5). We then select the set {C_2, R} with the best fit as our estimate for a given home. Lastly, we assume that the internal and external heat gains are not correlated with the indoor-outdoor temperature difference $T_r - T_a$, thus these quantities can be considered to be random noise for the curve fitting.

Once the estimates of the parameters C_1 and C_2 are available, the "natural" air leakage flow rate $\dot{V}_{natural}$ can be calculated by

$$\dot{V}_{natural} = \frac{\{C_1\left(\overline{W^{2.6}}\right) + C_2\left[\overline{(T_r - T_a)^{1.3}}\right]\}^{0.5}}{\rho_{air} c_{p,air}} \quad (11)$$

where ρ_{air} and c_p are density and heat capacitance of air. Finally, a conversion factor F ranging from 10 to 25 for U.S. homes (Krigger and Dorsi 2004) can be used to obtain ACH_{50}:

$$ACH_{50} = F\frac{\dot{V}_{natural}}{V}3600 \quad (12)$$

where V is the home volume and 3600 is the number of seconds in hour.

3.5 Results for Homes With One CT

3.5.1 Overall R-Value

We had 87 homes with acceptable quality CT data (i.e., with no or minimal missing data) with a single thermostat and furnace/boiler as well as available audit results and gas bills, of which 13 homes were from vendor #1, 41 from vendor #2, and 26 from vendor #3 (7 homes were excluded from consideration as outliers[7]). To get appropriate values for A_w and R_w, we complemented the HEA data with estimates for the heat-loss characteristics of the foundation and the window-to-

[6] If more than five such pairs are available, we can also test statistical significance of C_1.
[7] Homes with questionable reported ratios of surface area to conditioned floor space accounted for most outliers.

wall area ratio. For the floor R-value, we used a value of 9 (°F·h·ft^2/Btu) that corresponds to a typical value of 7 (Hallinan 2011), taking into account the lower floor-ground temperature difference, and we assumed the window area equaled 15% of the wall area.

Figure 5 and Figure 6 show a comparison between the overall R-value calculated by our method and the HEA R-value. We do not use a correlation coefficient to measure the goodness-of-fit as proposed by Goldman et al. (2018) for two reasons. First, some values from the HEAs are estimates that have varying degrees of uncertainty. Second, our ultimate goal is to identify homes with significant retrofit opportunities (i.e., classification) versus a precise estimate of R-value or ACH_{50}. It can be seen that, generally, our method accurately separates homes with poor insulation (R-values <8 in imperial units) from homes with adequate insulation (R-values ≥ 8). Quantitatively, the classification accuracy for these two classes is 88% overall (70 out of 80 classified correctly). The method tends to overpredict higher R-values; our initial assessment is that this overprediction is due to challenges identifying missing runtime data,[8] which we found is more challenging to detect for vendors #2 and #3. Given the sample sizes, there is no indication that the classification accuracy depends on the CT vendor.

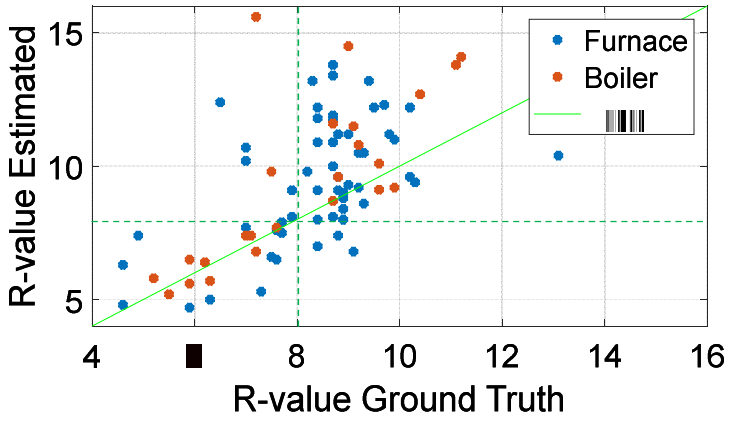

Figure 5. Estimated and HEA-based ("ground truth") overall R-values for homes with single CT and either gas furnace or boiler

R-values are given in imperial units (°F·h·ft²/Btu)

[8] Unaccounted runtime results in lower calculated heat supply from the heating system, making it appear that the building envelope has a higher R-value to maintain the indoor temperature set point.

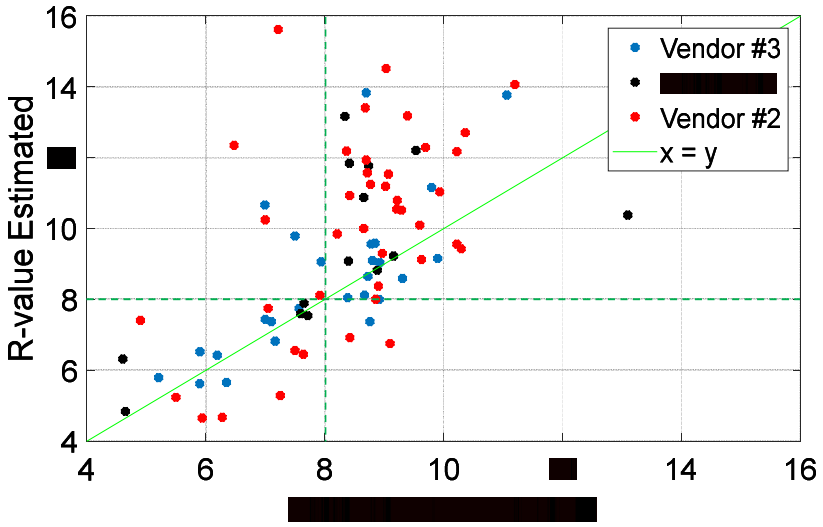

Figure 6. Estimated and HEA-based ("ground truth") overall R-values for homes with single CT and either gas furnace or boiler, by CT vendor

R-values are given in imperial units (°F·h·ft²/Btu)

3.5.2 Air Leakage

The results of our air leakage prediction for 16 homes with blower-door test results available are shown in Figure 7 and Figure 8. To estimate ACH_{50}, we used the conversion factor F of 14.8 in Eq. (12) for a two-story home located in Massachusetts (Krigger and Dorsi 2004). Although the discrepancy between the predicted and measured value can reach up to ~40%, we can effectively separate the homes with relatively low ACH_{50} from the leaky homes (with ACH_{50} >15). To the best of our knowledge, this is the first report of a successful ACH_{50} prediction based on CT data, limited home characteristics available to utilities, and weather station data only.

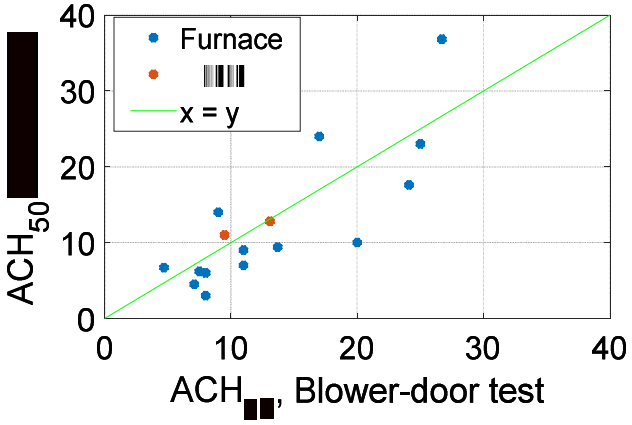

Figure 7. Estimated and HEA-based ACH_{50} values for 16 single-CT homes with blower-door test results available

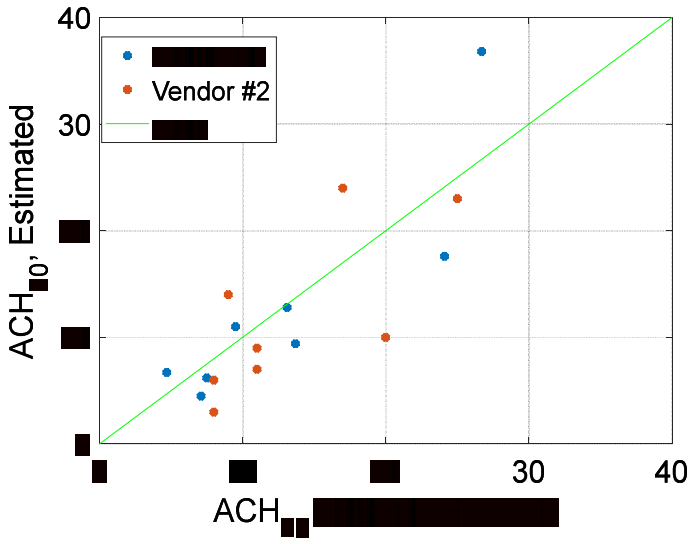

Figure 8. Estimated and HEA-based ACH_{50} values for 16 single-CT homes with blower-door test results available

3.5.3 Runtime

Accurately predicting runtime is important for estimation and verification of energy savings (see Section 5). Because our model is fitted to daily runtime data by minimizing the errors over the entire season, theoretically the difference between actual and predicted runtime over the entire season should approach zero. This is because when the sum of squared errors is minimized, the algebraic sum of errors (its derivative) tends to zero. In our case, the seasonal total error is *de minimus* but not exactly zero, mainly due to the numeric precision of MATLAB software (of the order of 10^{-15}).

It is still useful to look into daily runtime prediction errors as those essentially indicate the level of scatter in the correlation plots (see, e.g., Figure 2). Figure 9 shows an example of such average daily errors calculated over the entire heating season for the single-CT homes with furnaces. Although there are few homes with relatively large errors (primarily, these are homes with low furnace runtime and correspondingly high overall R-values), the average error among all these homes is 17% (absolute value).

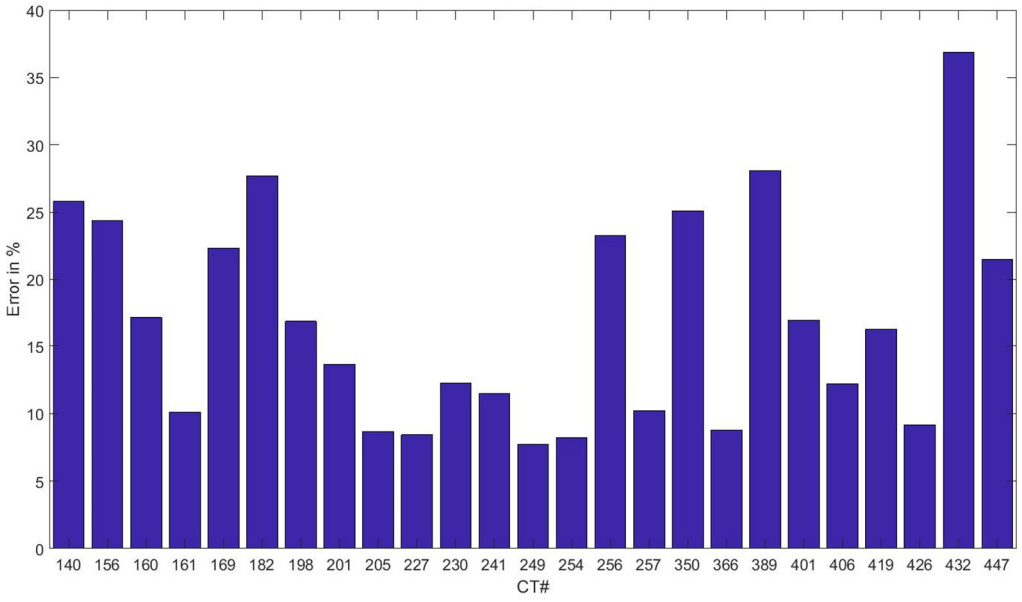

Figure 9. Averaged daily runtime prediction errors (absolute values shown) over the entire heating system for 25 homes with a single CT and furnace

Data for vendor #2 is shown. Average prediction error among all homes is 17%.

3.5.4 Comparison With State of the Art

For the sake of comparison, we also calculated the conventional "time constant" (Chong and George 2018) values for homes with significant nighttime thermostat setbacks. Because of the coarse resolution of temperature data from vendors #2 and #3, we used data from vendor #1 for single-CT homes with a furnace (see Table 1). For these calculations, we did not need gas bill data; therefore, we were able to process data for more homes (40) than reported in Table 2 (17). Figure 10 shows the average calculated decay rate (i.e., inverse "time constant") for each home versus the HEA-based R-value.

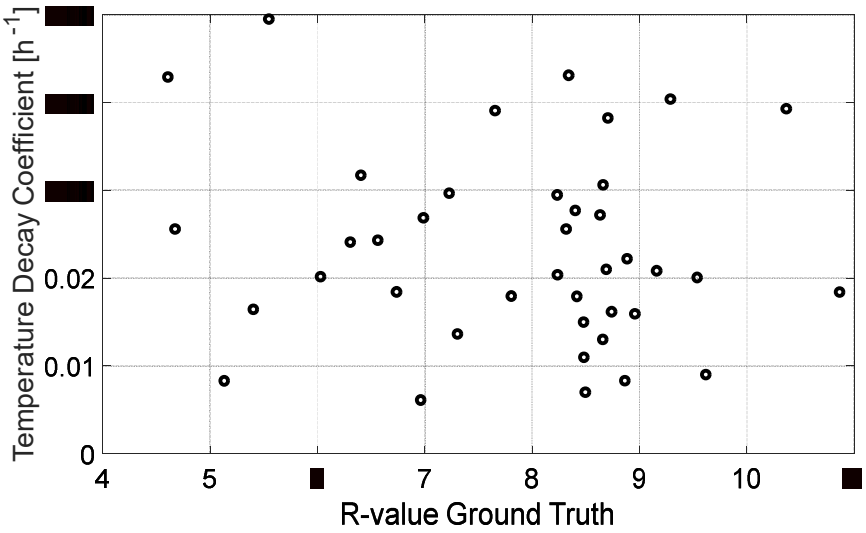

Figure 10. Calculated average decay rate for single-CT homes with furnace from vendor #1, versus their overall R-value

Chong and George (2018) suggest that the decay rate of 0.0167 h-1 corresponds to a thermally "tight" home, 0.025 corresponds to an average home, and 0.05 corresponds to a thermally "leaky" home. Whereas the homes with higher HEA-based R-value (≥ 8) may have significant air leakage and thus do not necessarily exhibit good thermal performance, the homes with low R-value (i.e., <8) certainly should exhibit below-average thermal performance. However, Figure 8 indicates that 9 out of 16 homes with low R-value (i.e., 56%) have decay rates below 0.025 which, according to Chong and George (2018), corresponds to above-average thermal performance. This observed inability of the "time constant" to serve as a reliable indicator of home thermal performance is consistent with our previous observation that the time constant depends on both thermal resistance and capacitance, and that effective thermal capacitance can vary appreciably among homes (see our discussion on first-order grey-box models in Section 3.1).

4 Homes With Two CTs

Homes with two CTs have more model uncertainties than one-CT homes. Although it is natural to assume that homes with two CTs have two major thermal zones (one CT per zone), the zone characteristics are usually not explicitly available in the CT/home data. Ideally, we would want the following information for modeling homes with two CTs:

- Zone/CT location and geometry (e.g., each per floor)
  - What CT corresponds to the upper/lower floor?
  - What is the zone external area?
- Heating system capacity per zone.

Because this information is not available, we decided to implement a whole-building equivalent approach for modeling the homes with two CTs. In this approach, we assume that the home is heated by an equivalent gas-fired system whose equivalent power and runtime can be calculated using the actual data. This approach overcomes the problem of matching the zone-specific estimates of R-value/ACH_{50} with the whole-home HEA-based values as well as the potential need to model interzone heat transfer. At the same time, the whole-home approach cannot isolate a retrofit opportunity to a specific zone.

Homes with two CTs can have either a single heating system (furnace or boiler) or two independent heating system (usually two furnaces or two boilers); in practice, we did not have that information available. Figure 11 shows how the equivalent power and runtime is proposed to be calculated in either case. For a single furnace, we assume that the full device's power is exercised if either CT calls for "on." Therefore, the equivalent power is the same as the device power (assumed to be constant), whereas the equivalent runtime is the union of the individual CT runtimes. For two separate devices (two furnaces or a boiler modeled as serving two separate zones), we assume that the devices have (generally different) powers q_1 and q_2, and that the equivalent power is either q_1, q_2 or q_1+q_2 depending on the "on" status as reported by the two CTs.

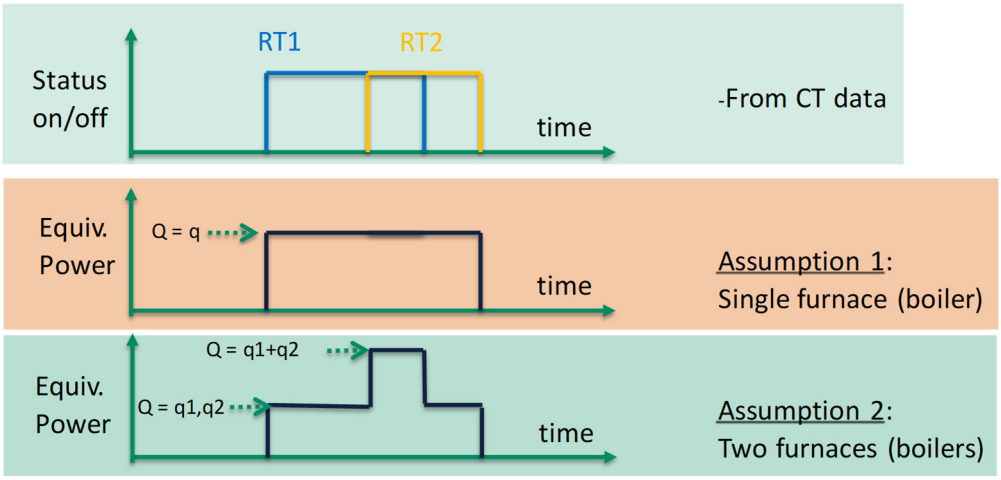

Figure 11. Calculation of equivalent whole-building power and runtime under two alternative assumptions

q = per/system, Q = per/home

To make this approach practical, we tested its sensitivity to a variety of potential heating configurations. In particular, we calculated the equivalent power and runtime by making empirical assumptions about a home heating system and tested the sensitivity of the end results (i.e., overall R-value and ACH_{50}) to these assumptions. Low sensitivity would imply high robustness of the proposed whole-building method, whereas high sensitivity would require modification of the approach.

To assess the sensitivity of the results to these assumptions, we used various ratios between q_1 and q_2 (ranging from 50:50 to 30:70 to 70:30) as well as those between the areas A_1 and A_2 (also

ranging from 50:50 to 30:70 to 70:30) and calculated, for each combination of q_1, q_2, A_1 and A_2, the overall building parameters (R-value and ACH_{50}) for all the two-CT homes with complete quality data sets. We found that the estimated overall building parameters only fluctuated within ±5% for 90+% of the homes, under the two assumptions and the aforementioned ranges.

Therefore, for homes with two CTs we implemented the single-device model that requires fewer underlying assumptions. In this model, we assume that the equivalent runtime is the union of zone runtimes (Assumption 1) and that the equivalent room temperature is the arithmetic average of the two zonal temperatures. Our reference (Zeifman, Lazrak, and Roth 2020b) provides additional theoretical framework to justify the above assumptions.

4.1 Results for Homes With Two CTs

We introduced additional heuristic data processing rules to address specific data issues we encountered. We found that chunks of missing runtime data often show up as outliers in the runtime-temperature delta correlations. Therefore, we removed such outliers using a statistical technique, "patched" these now missing data points with their estimates using regressions developed from the other, valid data applied to weather data for the outlier periods, and recalculated the heating power from gas bills using the updated runtime.

For homes with reported integer number of floors, we implemented consistency checks to correct potential audit-reported geometrical values. For homes with a fractional number of floors, e.g., 2.5 floors, we always used the audit-reported geometrical values because consistency checks can be more difficult for such homes. Finally, we observed several homes with unusual gas bills (e.g., homes with very high summer gas bills) and excluded those high bills in the gas baseline usage calculations. Some other homes had highly variable winter gas bills, even though the CT runtime and outside temperature did not show any anomalies. For these homes, we excluded periods with unusually low bills.

Figure 12 compares the HEA data and estimated overall R-value for 74 homes with two CTs and complete sets of CT interval data (i.e., data sets with missing data repaired or with no missing data). Using a threshold of R-value = 8 to separate homes with a significant insulation retrofit opportunity from those with no insulation opportunity, we obtain 87% for the overall classification accuracy. These results are comparable to those we obtained for homes with one CT (see Section 3.5.1).

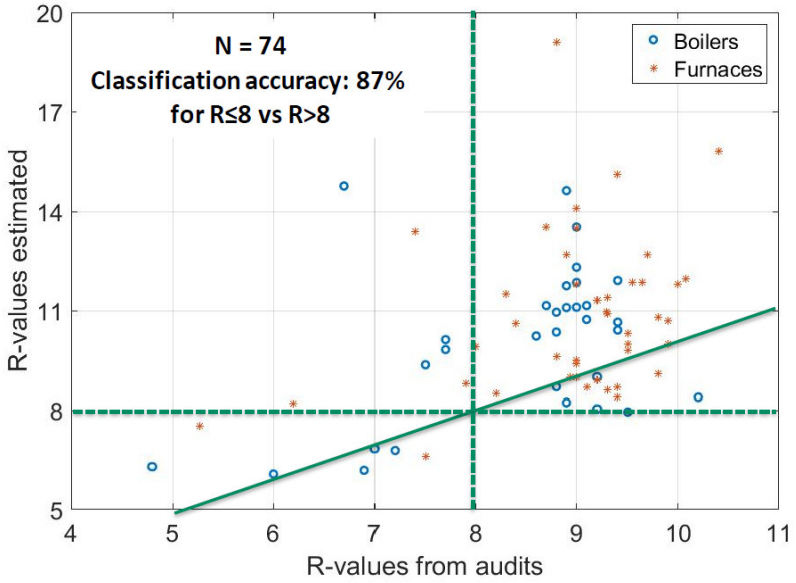

Figure 12. R-values estimated from CT data versus HEA data for homes with two CTs

Overall results for all homes with complete data sets

Figure 13 plots the ACH_{50} values for the model versus the HEA data calculated from the blower-door tests. Unfortunately, there were only eight homes with available HEA data ACH_{50} and complete sets of quality data, and none were leaky according to our threshold of $ACH_{50} \geq 15$. Yet, all these homes were correctly classified as homes with no air sealing opportunity. Note that our method appears to systematically underestimate ACH_{50}, which is consistent with Hales (2014). For a more stringent threshold of 7 ACH_{50}, the classification accuracy is 6/8 = 75%.

Although this technically meets the project objectives (see Section 1), we acknowledge that this finding has high uncertainty due to the very small sample size.

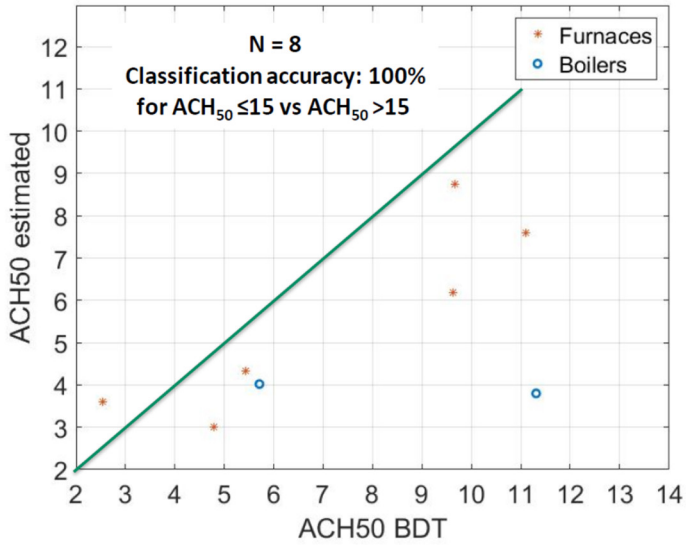

Figure 13. ACH_{50} values calculated from CT data versus blower-door test results for eight homes with two CTs

Finally, for homes with two CTs, we calculated the runtime prediction error over the season. Following our discussion in Section 3.5.3, the calculated overall runtime errors are small because of the way we perform our estimations.

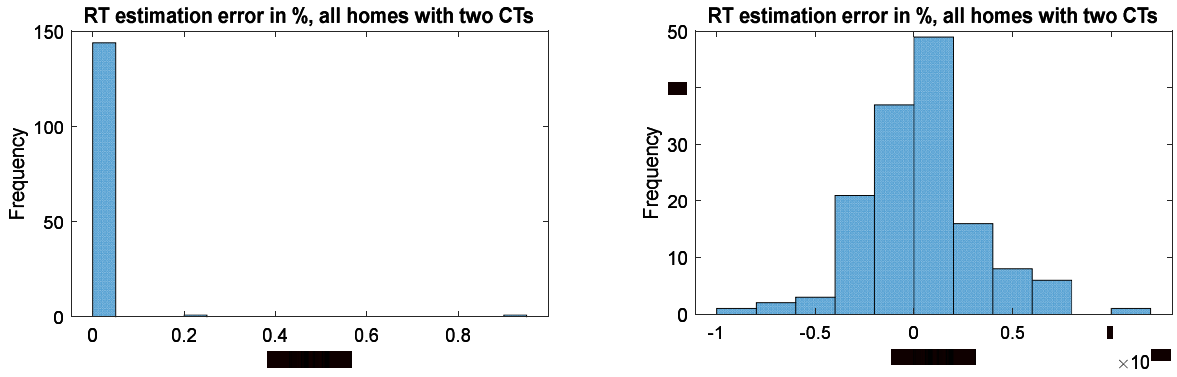

Figure 14. Runtime prediction errors over entire heating season for all homes with two CTs (boiler or furnace)

Left – all homes, right – outliers excluded

5 Energy Saving Predictions

The project's objective related to energy saving predictions (see Section 1) indicates that we can use either the predicted energy savings from the energy audits, or the actual energy savings obtained from implemented ECMs as a comparison value. Because we identified only a few homes that had implemented significant ECMs during the data collection period (see also Section 7.2), we used the HEA energy saving predictions as the comparison value.

HEA companies use proprietary software for energy saving prediction, and we reviewed a proprietary document describing the energy saving prediction algorithms for an energy audit company (Harley 2011). This document suggests that the state of the art uses a PRISM-like equation (Fels 1986, Hallinan et al. 2011) to calculate energy savings based on area-weighted pre-retrofit values of component R-values and average CFM along with their post-retrofit projections. However, this document did not disclose a methodology for prediction *post-retrofit* building characteristics.

Accordingly, we analyzed the data and verbose descriptions of energy audits and ECM implementations, part of the anonymized data sets transferred to us by our utility partners, to model post-retrofit building physical parameters (i.e., overall R-value and ACH_{50}).

5.1 Overall Post-Retrofit R-Value

We calculate the overall building R-value using the building geometry (external wall area, attic/roof area, window-wall area ratio) and the R-values for building major components (external walls, attic, basement/foundation and windows). Because the building geometry does not change in retrofits, we need to characterize the post-retrofit R-values for external walls and for attic (i.e., corresponding to the retrofits we consider in this project) to calculate the overall post-retrofit R-value.

5.1.1 External Walls' Post-Retrofit R-Value

The R-value of wall cavities filled with blown-in cellulose can vary due to differences in wall cavity thickness, exterior cladding properties, etc. Thus, we took two complementary approaches to estimate and model the post-retrofit R-value of walls that are "drilled and filled." First, we used building physics to estimate a typical wall R-values using area-weighted parallel heat transfer paths through wood framing and a 3.5-inch wall cavity filled with cellulose. Assuming a 25% framing factor (Lstiburek 2010) and a cellulose insulation R-value of 3.5 (Fisette 2005) yields a whole-wall R-value of around 11.9. Second, we reviewed the whole-wall R-values estimated for audited homes to understand how the audit software assessed the wall R-value of walls with cavity insulation. We considered all the homes with audit information available that did not have a wall-insulation retrofit implemented (i.e., these homes do not appear in the measures file). We found that of approximately 800 homes with external wall R-value greater than 6, more than 50% had the R-value in the range between 11 and 12 and only 10% had an R-value higher than 12. Based on these two approaches, we decided to use the midpoint, R-value of 11.5, as a practically achievable post-retrofit whole-wall R-value.

5.1.2 Attic Post-Retrofit R-Value

For the attic retrofits, there is less uncertainty in converting the verbose ECM description into R-value as compared to the wall retrofits. Accordingly, we calculated the post-retrofit R-value following energy audit standards (Energy Assessment Standards 2012) and the retrofit verbose descriptions for approximately 350 homes with significant (i.e., more than 25% R-value increase) attic retrofits. Table 3 shows the distribution of post-retrofit attic R-values in these homes. Although a majority (60%+) of retrofits achieve around R-40, the overall distribution is very broad. Hence, we believe an assumption of a single post-retrofit R-value of 40 is not adequate for all homes.

Table 3. Distribution of Post-Retrofit Attic R-Values as Reported in ~350 HEAs

Post-retrofit attic R-value	**<10**	**10-15**	**15-20**	**20-25**	**25-30**	**30-35**	**35-40**	**40-45**	**45-50**
Fraction of homes, %	3	4	7	5	10	7	35	27	4

As a working alternative, we calculated the difference between the post- and pre-retrofit attic R-value. We found an average difference of 15.6, a median of 14.8, and the three modes of a multi-modal distribution of 6.6, 15.3, and 22.0. Based on this, we use the average increase in R-value (15.6) corresponding to the statistical expectation to model the post-retrofit attic R-value.

5.1.3 Overall Post-Retrofit R-Value Prediction

With the developed models for post-retrofit wall and attic R-values, we can calculate the post-retrofit overall R-value as would be predicted for homes in HEAs. However, the *prediction* of the post-retrofit overall R-value for the proposed method, which is conditional on the estimated pre-retrofit R-value, is not straightforward. Although our R-value prediction is generally in line with the R-value based on the HEA data for binary classification of homes, there are notable discrepancies between our estimates and the HEA-based data (see Figure 5 and Figure 12). These discrepancies could be attributed to the assumptions underlying the HEA calculation (e.g., the assumed values of window-to-wall ratio and basement R-value) and/or to errors in CT data, gas bills and audit results, variability in HVAC system efficiency (particularly when considering duct losses), as well as to the coarseness of the underlying physics-based model.

Therefore, we need to calculate post-retrofit R-value in a way that minimizes the expected errors from the HEA-based values. To this end, we developed the following mapping procedure:

1. Divide the range of the overall pre-retrofit whole-home R-values for homes worthy of insulation upgrade ($R \leq 8$) into four intervals: $R<5$; $5 \leq R<6$; $6 \leq R<7$; and $7 \leq R \leq 8$.
2. For each interval i, $\{i = 1, 2, 3, 4\}$ select homes with the estimated from CT data (pre-retrofit) R-value falling within the interval. For each selected home j from interval i, calculate the post-retrofit overall R-value $R_{ij\_GT\_Post}$ assuming the wall post-retrofit R-value of 11.5 and the attic's R-value increase of 15.6. Because of the complex nonlinear relationship between the estimated and R-values based on the HEA data, the obtained values cannot be used directly to predict the estimated post-retrofit R-values.
3. Instead, for each interval i, find the post-retrofit value X_i for the estimated R-value that minimizes the difference between

$$\sum_j \left(1 - \frac{R_{ij}}{X_i}\right) \tag{13}$$

and

$$\sum_j \left(1 - \frac{R_{ij\_GT}}{R_{ij\_GT\_Post}}\right) \tag{14}$$

4. The physical meaning of each summand is a relative change of insulation heat loss in a building. Build a table mapping each interval onto the average R-value increase $\Delta R_i = X_i - \bar{R}_{ij}$, where bar designates the average over j value.

We built the mapping table using 34 homes with estimated R-values lower than 8 (with one or two CTs). Table 4 lists the details of the mapping table for R-value prediction. To calculate the post-retrofit prediction for an arbitrary pre-retrofit R-value, we need to find the proper interval and the corresponding ΔR. The post retrofit prediction equals the sum of R and ΔR. Using the proposed mapping methodology, we can extend this table for higher R-values should the need arise.

Table 4. Prediction of Estimated Post-Retrofit Overall R-Value Increase

Estimated R-value, pre-retrofit interval	Post-retrofit ΔR for estimated
R<5	2.4
5≤R<6	2.5
6≤R<7	2.2
7≤R≤8	2.1

5.2 Post-Retrofit ACH_{50}

Unlike the overall R-value, ACH_{50} is directly measured in a blower-door test.[9] Accordingly, we can directly project post-retrofit ACH_{50} based on the measured change in ACH_{50} from actual air sealing projects. However, we need to reduce the scatter in the HEA data to make meaningful predictions. An analysis of the blower-door test results suggests that the reduction in air leakage (i.e., Δ ACH_{50}) can be modeled as a linear function of the pre-retrofit ACH_{50} value. Figure 15 illustrates the analysis with experimental data on 85 homes with blower-door test results available. The following stepwise linear fit can be used to model the data:

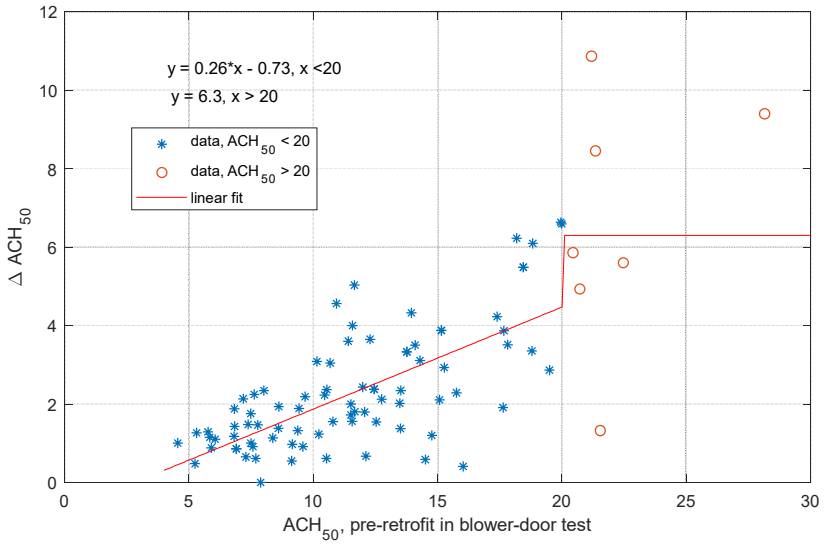

Figure 15. Experimental data on pre/post-retrofit ACH_{50} as measured in blower-door tests

$$\Delta ACH_{50} = \begin{matrix} 0.26 \times ACH_{50} - 0.73, if\, ACH_{50} < 20 \\ 6.3, if\ ACH_{50} \geq 20 \end{matrix} \quad (15)$$

Eq. (15) can be used to predict the expected reduction air leakage for homes. With this model, we mapped the estimated ACH_{50} to that predicted by Eq. (15) ΔACH_{50} for the 24 homes for

[9] However, Hales (2014) suggests that the blower-door test systematically overestimates ACH_{50} relative to tracer-gas tests.

which we have both estimated and HEA-based infiltration rates (ACH_{50}) available (see Figure 7 and Figure 13). Figure 16 shows the results.

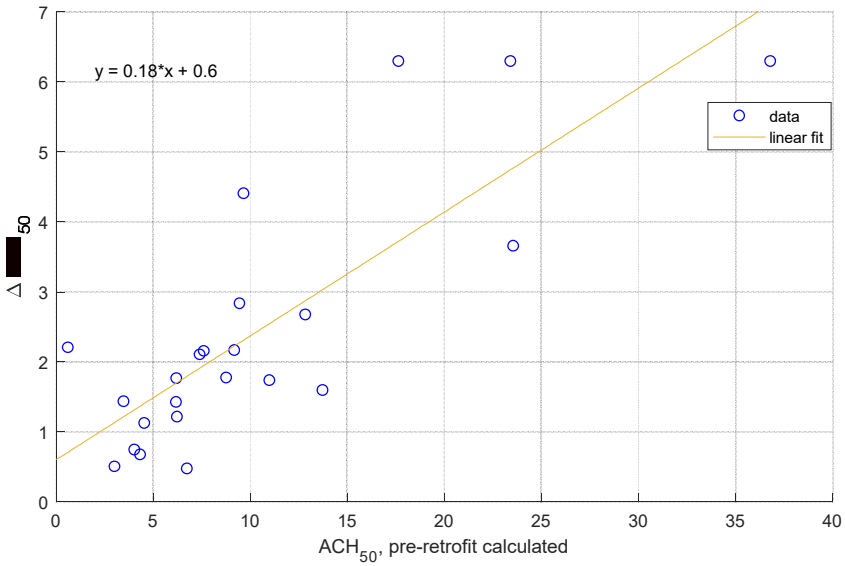

Figure 16. Predictions of ΔACH_{50} modeled by Eq. (3) versus estimated ACH_{50} for 24 homes that have both estimated and blower-door ACH50 available

The final equation for prediction of ΔACH_{50} given the *estimated* ACH_{50} value is

$$\Delta ACH_{50} = 0.18 \times ACH_{50} + 0.6 \qquad (16)$$

We use this equation to predict the post-retrofit parameter ΔACH_{50} values. Ultimately, we use this combined with the change in whole-home R-value (calculated using Table 4) to assess the combined energy impact of the two potential measures.

5.3 Energy Saving Calculations

Because our modeling approach can be considered as an extended version of PRISM (Fels 1986, Hallinan et al. 2011), we can apply the same methodology to calculate predicted savings for both the HEA-based values and our proposed method. In particular, for a given home, we can calculate the change in runtime caused by either a higher overall R-value or lower air leakage over the entire heating season. By using separate sets of HEA-based pre/post values and those estimated, we obtain runtime reductions for the HEA prediction and for the proposed method. We can also predict relative (i.e., %) savings by normalizing the runtime reduction over the overall season by the pre-retrofit runtime for a home.

We used this methodology to calculate HEA-based energy savings and those for the proposed method. For insulation savings, we calculated the savings for 34 homes with insulation retrofit opportunities (i.e., R <8). Figure 17 shows a comparison between the HEA prediction and our predicted savings. Note that the average energy saving for such homes is substantial, of the order of 30%. Out of 34 predictions, 29 (85%) are within ±25% from the HEA values, which suggests that we have exceeded the Project Objective #2 (see Section 1).

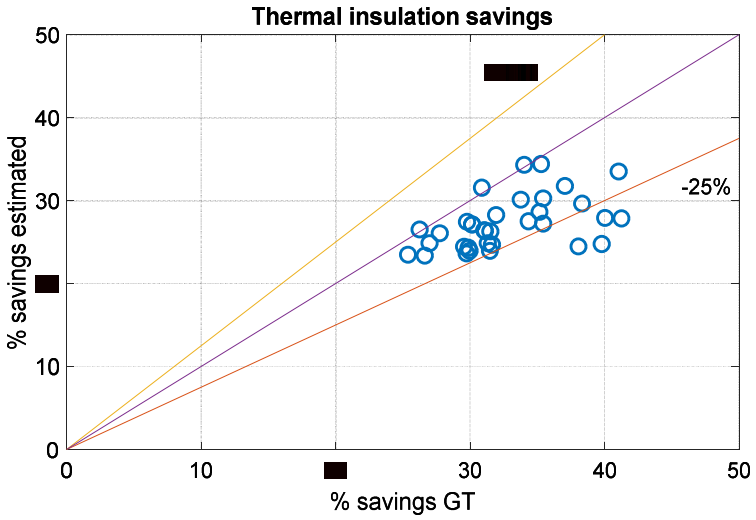

Figure 17. Estimated and HEA-based (GT) predictions for percent energy savings

The results are calculated for homes with one/two CTs, complete sets of quality data and insulation retrofit opportunity. Out of 34 homes, 29 have estimated predictions within ±25% of HEA predictions.

For the air sealing opportunities, we have 24 homes for which we have both ACH_{50} estimates and HEA data (i.e., blower-door tests). Out of these, we have only four homes with ACH_{50} >15, i.e., homes with significant retrofit opportunities. To increase the sample size and in accordance with the classification results discussed earlier, we consider homes with blower-door test results ACH50 exceeding 7,[10] i.e., 15 homes.

Figure 18 shows the results for these 15 homes. Even though the air sealing retrofits yield less savings than the insulation retrofits (see Figure 17), the savings are still significant. The predictions are within ±25% for 11 homes, i.e., for 73%. That said, this finding has high uncertainty due to the very small sample size.

[10] Looking at the pre-retrofit ACH_{50} blower-door test values, we could not identify a clear value (or even range) when air sealing is recommended (see Figure 15). We believe this likely reflects that air sealing recommendations are based on qualitative assessments of the expected benefit and ease of accessing potential leakage paths for air sealing.

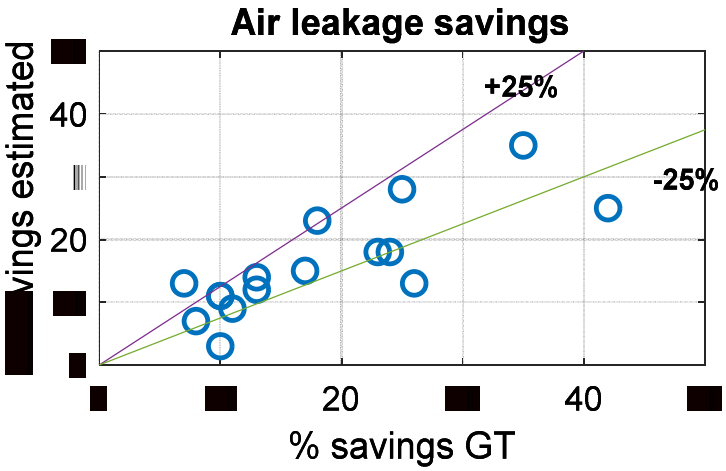

Figure 18. Estimated and HEA-based (GT) predictions for percent energy savings from air sealing

The results are calculated for homes with one/two CTs, complete sets of quality data and measured in blower-door tests ACH_{50} >7. Out of 15 of homes, 11 have estimated predictions within ±25% of HEA-based predictions

6 Randomized Controlled Trial

Successful development of algorithms to conduct remote HEA/predict energy savings provides an opportunity to test whether customized retrofit recommendations and savings potentials for individual homes ultimately can:

1. Significantly increase the uptake rate of on-site HEAs, and
2. Significantly increase the fraction of HEAs resulting in ECM implementation.

6.1 RCT Design

We designed and conducted a randomized controlled trial (RCT) to answer these two research questions. In an RCT, households meeting the test criteria are randomly assigned to the treatment and control groups. The criteria for homes to qualify for the treatment and control groups in the RCT are:

- Located in Eversource or National Grid's (also known as the program administrators) gas service territory in Massachusetts
- Gas is the primary space-heating fuel
- Detached single-family home
- Customer has received a rebate from Mass Save for a CT
- The program administrator has access to customer CT data for at least half of one heating season
- The program administrator has access to customer gas billing data for at least one year coincident with the period of the CT data

- Customer has not had a Mass Save HEA.

Later in the project, our utility partners found that the last criterion was violated for a majority of customers in some groups (see Section 6.1.1 for details).

In the RCT, the treatment group receives targeted outreach informed by the algorithms. That is, Fraunhofer applied the algorithms to the CT data from the treatment group customers to identify homes that are likely to have an insulation (attic or wall) or air sealing opportunity. For these homes, the algorithms identified the expected retrofit opportunity(ies) and calculated their expected post-retrofit energy savings (by updating the physical parameters in the algorithms to reflect the retrofit targets), and Fraunhofer provided those data to the program administrators. In turn, the program administrators incorporated the customer-specific offerings and energy savings into customer outreach and sent them to the appropriate treatment group customers.

Based on the data sets we received from our utility partners, we designed one treatment and four control groups. The control groups serve distinct purposes. Control groups 1 and 1A comprise customers who took a CT incentive and for whom we have CT data that are taken from the same population as the treatment group. Group 1 includes customers with low CT data quality that precluded identification retrofit opportunities. That said, we expect them to have the same rate of retrofit opportunities as the population of customers that we analyzed with acceptable CT data quality, i.e., control group 1A and the treatment group combined. In contrast, control group 1A comprises customers for whom the algorithms found no retrofit opportunities. Because both control groups will receive generic outreach, this should reveal if the customized feedback has a significant impact on HEA and/or retrofit uptake for people purchasing and requesting a rebate for a CT. This should enable us to quantify the impact of the customized outreach.

Control group 2 represents customers who received an incentive for installing a CT from another manufacturer, foremost Nest, that did not provide CT data to the PAs. Consequently, it will provide insight into if the frequency of customer outreach affects enrollment. Finally, we surmise that customers who install a CT and take an incentive may differ meaningfully from the large majority of customers who do not have a CT installed. Thus, control group 3 compares participation relative to customers who did not take a CT incentive. Because any outreach has the potential to increase program participation, the first control group will also receive generic Mass Save marketing outreach whenever the treatment group receives customized outreach.

Table 5 summarizes the RCT design. We calculated the statistical test power given the expected participation rates of 2%–3% and 10% and implementation rates of 30% and 60% for the control and treatment groups, respectively. We concluded that these characteristics exceed those commonly acceptable in experimental design (typically, test power of 0.8 at significance level of 0.1).

Table 5. Randomized Controlled Trial Design

Group Type	Description	Supplemental Outreach?	Sample Size, N (actual)
Control Groups	1. Took CT incentive, have CT data – Low CT data quality	Receive generic outreach four times, i.e., same as treatment group	902
	1A. Took CT incentive – Algorithm Identified NO retrofit opportunity	None	212
	2. Took CT incentive, other CT type	None	1,000
	3. Did not take CT incentive	None	1,000
Treatment Group	1. Took CT incentive, have CT data	Receive customized outreach four times	216

Table 5 shows a fraction of homes with high-quality CT data that is similar to that in the previous data sets in this project. Initially, we received CT data from vendors #1 and #2 (see Table 1), along with gas bills and publicly available home information (conditioned area and number of floors) for 1,332 homes. Of those, 430 homes or 32% had good-quality CT data. At the same time, Table 2 suggests that for algorithm development, we had 457 homes with one/two CTs with complete sets of data overall. Of these, we were able to process 87 homes with one CT (see Section 3.5.1) and 74 homes with two CTs (see Section 4.1), i.e., 161 homes or 35% overall. We also had to extend Table 4 to incorporate higher estimated overall R-values (up to 11) to boost the treatment group size; as expected, those homes with the higher estimated R-value commanded relatively small saving percentage that nonetheless often exceeded our targeted value of $50 per season.

6.1.1 The Repeat Customer Problem

Initially, we identified 216 customers for the treatment group ($50+ projected savings) and provided their anonymized customer IDs along with their projected seasonal savings (in therms and dollars) for insulation and air sealing opportunities to National Grid and Eversource.

However, subsequent communications with National Grid in June 2019 found that:

- A significant number of customers (up to 75%) had a previous home energy assessment (as long as the program administrator data goes back in time), and
- 49 out of the 216 customers initially selected for the treatment group and two customers of control group 1A had air sealing/insulation measures installed (note: the depth of those measures was not disclosed).

Consequently, eliminating customers with any prior home energy assessment from the RCT would shrink the treatment group and control groups 1 and 1A approximately fourfold, rendering these groups too small for a statistically meaningful inference. Accordingly, we decided to not remove such homes from participation, but have their (anonymized) IDs available for future study.

On the other hand, prior ECM implementation does preclude selecting customers for the other control groups. Table 6 lists the updated group sizes. For control groups 1 and 1A, we only received information on prior HEAs from National Grid.

Table 6. Updated RCT Design

Group Type	N, Original RCT design	Number of customers that had HEA in the past	Number of qualifying customers (N)
Treatment Group (opportunity found)	216	163 (75%)	53 (25%)

Control Group 1A, National Grid part (no opportunity found)	130	53 (41%)	77 (59%)
Control Group 1, National Grid part (poor quality CT data)	604	325 (54%)	279 (46%)
Control Group 2 (another CT vendor)	1000	0	999[11] (100%)
Control Group 3 (no CT incentive)	1000	0	999 (100%)

Another potential problem that our utility partners have discovered was that some of the homes were multifamily. The majority of such multifamily homes were detected during the initial screening and disqualified from RCT, yet some homes could be identified as multifamily only by a manual search using their physical address.

Note that the fraction of homes with previous HEAs is highest in the treatment group (75%), lowest in control group 1A (41%), and at a midpoint in control group 1 (54%). Given the group sizes, these differences are statistically significant. We believe that these differences can be attributed to the way we selected the homes with retrofit opportunities: Homes with the identified retrofit opportunities are likelier to have requested HEAs in the past than the homes with no found retrofit opportunity.

6.2 RCT Results

The printed materials were sent to the treatment group and to control group 1 as follows:

- 1st round: Sent on July 26
- 2nd round: Sent on week of September 27
- 3rd round: Sent on week of November 4
- 4th round: Sent on December 6.

For the treatment group, the results for the qualified customers (N = 53, see Table 6) are as follows. Five homes had requested an HEA as of June 30, 2020—on August 29, September 14, September 24, September 30, and January 15. Two of these homes also had insulation and air sealing measures installed a month and three months after the HEA (HEA on September 30, ECMs installed on November 1 for the first home, HEA on September 24, retrofits implemented on January 2 for the second home). The savings predicted by our algorithms for these homes ranged from 11%–22%; interestingly, the homes with the highest predicted savings among these

[11] We do not know why the numbers of customers in control groups 2 and 3 were 999 and not 1,000.

homes (21% and 22%) were the homes that decided to implement ECMs. No customer who had had an HEA within the past six years has requested an HEA.

For control group 1, which received four waves of generic mailers, the results were somewhat similar in terms of timing. For the pool of qualified customers (N = 279, see Table 6), 11 homes requested HEAs on the following dates: October 18, October 26, November 16, November 22, November 26, December 9, December 11, December 12, January 11, February 11, and February 26. One of these homes also had insulation and air sealing retrofits implemented (HEA on October 18, retrofits implemented on January 3).

That said, nine unqualified homes (i.e., those that requested HEAs in the past) also requested HEAs on the following dates: August 16, October 2, October 25, November 26, December 23, January 22, February 19, April 4, and April 18, and the homes that requested an HEA on October 2, December 23, and February 7, then installed ECMs (both insulation and air sealing) on December 12, February 27, and March 20, respectively. We are not clear why a significant number of homes that requested an HEA (9 out of 20) were the homes that already had an HEA in the past in control group 1. Our utility partner suggests that potentially, some homes could have changed an owner recently and/or were from a multifamily building.

For control group 1A (N = 77, see Table 6), one home requested an HEA on July 31 and then installed insulation and air sealing ECMs on October 29.

For control group 2 (N = 999), 28 homes requested HEAs, while 10 homes had implemented insulation and air sealing ECMs as of June 30, 2020. Notably, the time lag between the HEA and ECM was as long as five months for some homes.

Finally, for control group 3 (N = 999), 24 homes requested HEAs, and three homes had implemented insulation or air sealing ECMs as of June 30, 2020.

These results are summarized in Table 7. For control group 1, the numbers in parentheses show the results assuming that the nine homes with prior HEAs had new owners and thus were qualified to participate. **Crucially, the treatment group has an HEA rate approximately two to five times greater than the control group that received generic mailers, supporting the hypothesis that targeted, customized outreach can realize significant increases in energy efficiency program participation.**

Table 7. RCT Results

Group Type	N	Number of homes with HEAs	Number of homes that installed ECMs	HEA Rate: # HEAs/N, ±1σ	HEA Conversion Rate: HEAs/ECMs
Treatment Group	53	5	2	9±4%	2 out of 5
Control Group 1	279 (604)	11 (20)	1(4)	4±1% (3±0.5%)	1 out of 11 / 4 out of 20

Control Group 1A	77	1	1	1±1%	1 out of 1
Control Group 2	999	28	10	2.8±0.5%	36% (10 out of 28)
Control Group 3	999	24	3	2.4±0.5%	3 of 24

Generally, the RCT results follow our expectations. Control group 1A (no opportunity found, no messaging) was the poorest performing group, closely followed by control groups 3 (no CT incentive, no messaging) and 2 (CT incentive, no messaging). Control group 1 (CT incentive, generic messaging) performed better than those. Finally, the treatment group shows the highest HEA request rate (about six time higher than the no-messaging background, meaning control groups 1A, 2, and 3, and three times higher than control group 1). Given the time lag between an HEA and ECM installation, we expect additional ECM installs in all RCT groups over the next several months.

7 Use Cases

In this section, we consider three use cases that are potential extensions of the proposed technical approach, yet are somewhat beyond the scope of the original project.

7.1 Cooling Season Considerations

The technical approach developed is applicable to CT data collected from individual homes over the heating season. Discussions with other utilities revealed appreciable interest in using CT data collected over the *cooling* season, potentially combined with electric interval data. This would significantly enlarge the pool of candidate homes with retrofit opportunities.

The grey-box models, Eqs. (2)-(5) that provide a foundation of our approach, were derived for the heating season. In principle, they could be extended to a cooling season provided that the following two factors are incorporated:

- Air conditioner (AC) performance curve
- Moisture transport.

In addition, we expect that solar heat gains will play a larger role in AC loads than during the space heating season.

Unlike a fuel-burning heating system modeled by just two values (Q at state “on” and 0 otherwise) in Eqs. (2)-(5), the cooling supply of an AC system is not constant at state “on”; instead, the corresponding performance curve (e.g., Cutler et al. 2013) that models the dependence of AC cooling capacity and power as a function of outdoor and indoor temperatures will need to be incorporated in the grey-box equations. Latent heat removal, i.e., dehumidification, necessitates additional equations for moisture transport (Yang et al. 2018) that further complicate derivation of daily correlations of the type of Eq. (5).

Although extension of our method to a cooling season is not straightforward, the way we estimate the wind-driven air leakage—i.e., the coefficient C_1 in Eq. (9)—remains valid in the cooling season[12] because it compares system runtimes over similar time windows with high versus low wind, and a proper matching of the time window pairs cancels the unknown latent heat and also solar heat gains from the difference. If interval electricity consumption data for home's AC system were available, we could then estimate the value of coefficient C_1 and compare this estimate to its heating-season counterpart. The equality of the two estimates would imply potential applicability of our method to the cooling season data.

Because electric interval data are not available to us in this project, we cannot directly compare the cooling- and heating-based C_1 estimates. Nonetheless, we can calculate the ratio between the estimates of Q_{heat} in the heating season to the average value of Q_{cool} in the cooling season by using Eq. (9). Assuming that the average value of Q_{cool} is of the order of a nominal (i.e., nameplate) cooling capacity of the home AC system, we estimate that the calculated Q_{heat}/Q_{cool} ratio in a Massachusetts climate should be in the range of 1 to 3.[13]

The CT data we obtained from vendor #3 included data from a central AC system for some homes. In all, we identified 71 homes with good-quality CT data for both heating and cooling seasons, of which 52 had cooling data from a single zone (i.e., a single CT in the cooling season, whose ID may or may not be the same as those in the heating season). We modified our air-leakage algorithms for the cooling season and then applied them, together with their heating-season counterparts, to the data from these 71 homes.

Figure 19 shows the results. As expected, the ratio is always greater than unity, but many ratios are appreciably larger than anticipated. To a significant extent, this may reflect the relatively short cooling season duration and modest summertime outdoor temperatures in Massachusetts. This results in fewer days with appreciable cooling, making it challenging to find well-matched periods of time and decreasing the number of data points in regressions (compromising regression quality).

More work and data, specifically interval electricity consumption data, are needed to fully explore the algorithmic capabilities during the cooling season.

[12] Assuming that wind predominantly comes from similar directions during the cooling and heating seasons.
[13] This is based on a 2,000 ft^2 home with a 50 to 150 kBtu space-heating system and assuming 500 ft^2/ton of cooling (= 4 tons * 12,000 Btu/ton = 48 kBtu).

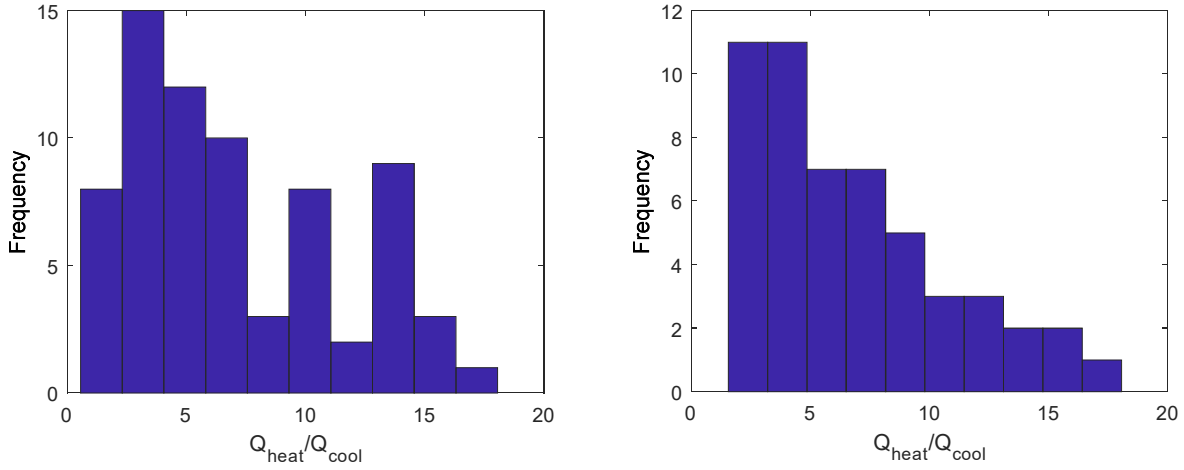

Figure 19. Estimated heating to cooling capacity ratio for homes with both heating and cooling data available

Left – all homes with heating and cooling data (71), right – only homes with one cooling zone and heating data (52)

7.2 EM&V for Retrofits

It is well known that conventional methodology for evaluation, measurement, and verification (EM&V) of post-retrofit savings requires several years of post-retrofit fuel bills. In addition, significant uncertainty remains whether household maintain the same level of thermal comfort post-retrofit—in other words, increasing thermostat setpoint reduces the potential savings.

Our approach may provide a quicker and more reliable opportunity for EM&V if sufficient CT data are available pre- and post- ECM implementation. This, in turn, depends on the data scatter in the daily correlation plots (see, for example, Figure 2[14]). As a rough estimate,[15] we can use the results from linear regression theory for the standard deviation of the slope and then compare the difference in slope pre/post-retrofit with this standard deviation. The standard deviation of the slope is inversely proportional to the square root of data point number (i.e., the number of days with CT runtime data in our case) so that the minimum number of days with CT data required to statistically discern pre/post data and estimate savings can be calculated.

In addition, because ECM implementation should not change the thermal mass of a home, the conventional first-order lumped resistance-capacitance "cooling curve" approach (see Figure 10 and corresponding text) can be used to check if the decay rate has changed post-ECM implementation. However, because of the coarseness of the temperature data from vendors #2 and #3 (see Table 1; temperatures reported in 1°F increments), the calculated decay rates can be too coarse for a meaningful comparison.

[14] Note that the scatter in Figure 2 can be attributed in part to wind-driven air leakage.
[15] Because the infiltration model in Eq. (4) incorporates the nonlinear stack effect, the overall dependence of the runtime on temperature difference is slightly nonlinear.

In our data sets, we identified 17 homes with ECMs implemented during the data collection process; of those, six had data at least two weeks both before and after ECM implementation. Two examples of these homes are detailed below.

7.2.1 Example 1

The first example is a home heated by an atmospheric (i.e., non-condensing) furnace. The home was built in 1989, and the HEA values for wall R-value is 8.1 and for attic R-value is 12. The following ECMs were installed:

- Air sealing, door stripping
- No wall insulation
- Attic: (1) attic floor open blow cellulose 5", (2) Propavent 2' or 4', (3) attic stair cover thermal barrier with carpentry.

Figure 20 and Figure 21 show the computed correlations for the CT data obtained by the home's two thermostats. Whereas Figure 20 suggests no noticeable pre-post difference for the correlations computed by the "downstairs" CT, dramatic difference is evident for both types of correlations in Figure 21. This is in line with the performed retrofits that are mainly upgrading the attic.

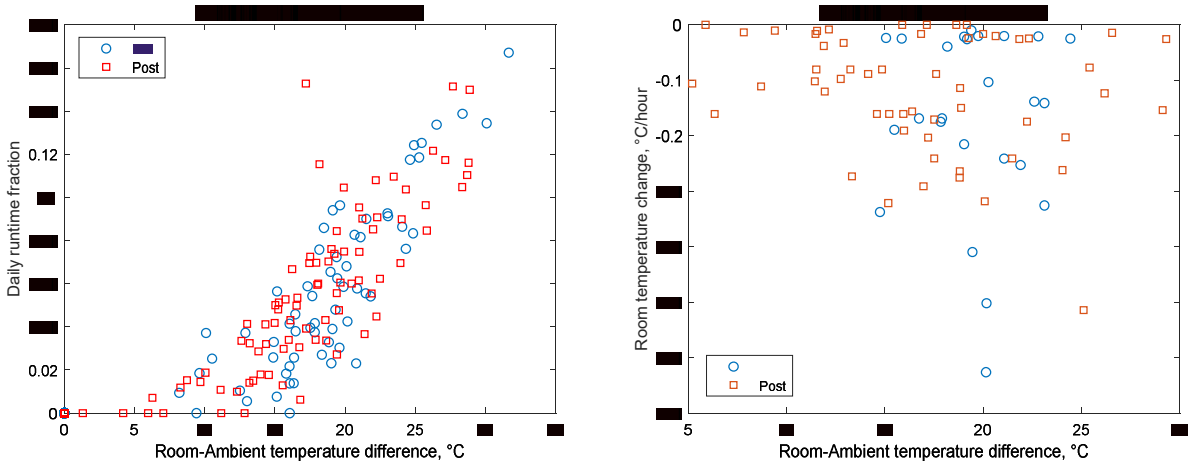

Figure 20. Correlations for pre/post ECM implementation for a home from vendor #3, downstairs CT

Left – runtime correlations, Eq. (5), right – cooling decay rates over thermostat setbacks

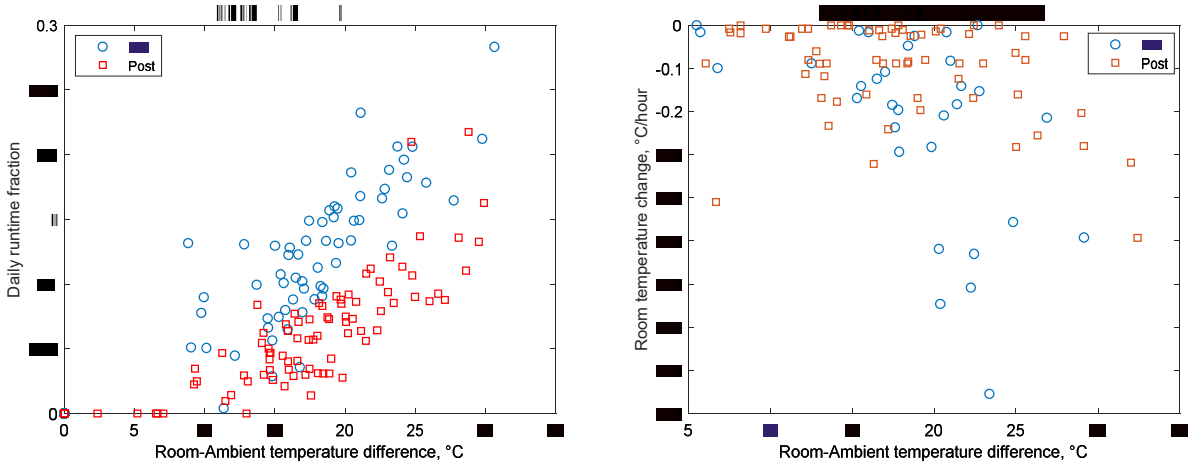

Figure 21. Correlations for pre/post ECM implementation for a home from vendor #3, upstairs CT

Left – runtime correlations, Eq. (5), right – cooling decay rates over thermostat setbacks

Quantitatively, we calculated the pre-post slopes of the regression lines using the whole-home equivalent approach (see Section 4), as well as their standard deviations. The pre-retrofit slope value is 0.0151 with a standard deviation of 0.0012. The post-retrofit value is 0.0120 with a standard deviation of 0.0008. The savings can be roughly estimated as (0.0151 – 0.0120)/.0151 = 20.5%. The difference in slopes is 0.0031 and its standard deviation is 0.0014, i.e., the slope difference is about 2.2σ. This difference is statistically significant at a significance level of 0.014 for a normal distribution. Encouragingly, this change in the second-floor thermal performance is consistent with the major upgrade in attic insulation, i.e., adding ~R-18 of insulation to an R-12 attic.

7.2.2 Example 2

The second example is a home heated by a condensing furnace. The home was built in 2000, and the HEA values for wall R-value is 11.4 and for attic R-value is 17.8. The following ECMs were installed on or about February 21, 2017:

- Air sealing, door stripping
- No wall insulation
- Attic: (1) Propavent 2’ or 4’, (2) attic stair cover thermal barrier with carpentry, (3) attic floor open blow cellulose 4”.

Figure 22 shows the computed correlations for the CT data obtained by the home’s thermostat. Although the installation date is more than a month later than that in the first example, there are relatively few pre-ECM data points because the CT was connected on February 2, 2017—less than three weeks prior to ECM implementation. Yet, some pre/post difference is visible in the runtime correlations. The calculated decay rates are not very meaningful, mainly because the temperature data from vendor #2 are too coarse (see Table 1).

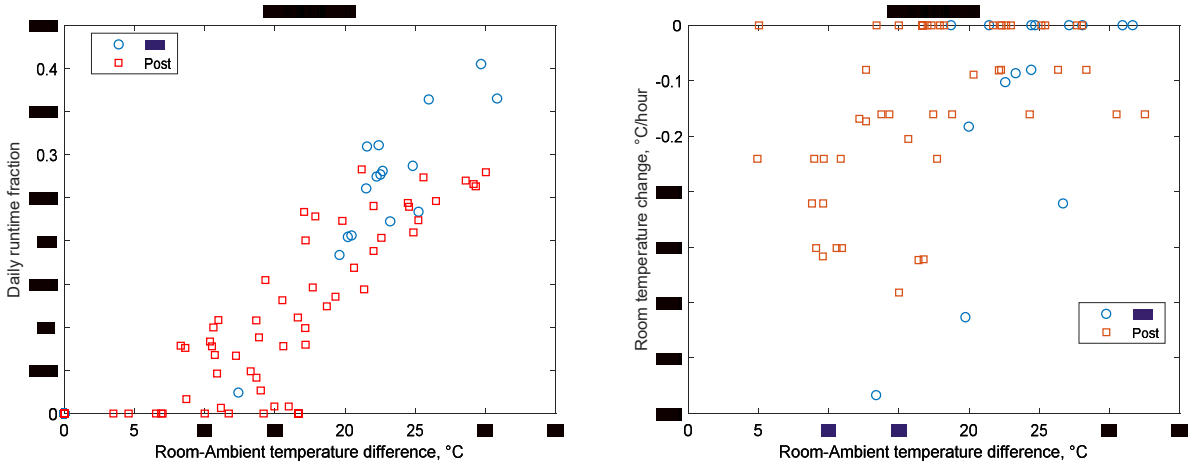

Figure 22. Correlations for pre/post ECM implementation for a home from vendor #2

Left – runtime correlations from Eq. (5), right – cooling decay rates over thermostat setbacks

Quantitatively, the calculated pre-retrofit slope value is 0.0118 with a standard deviation of 0.0025. The post-retrofit value is 0.0089 with a standard deviation of 0.0011. The savings can be roughly estimated as (0.0118 – 0.0089)/.0118 = 24.6%. The difference in slopes is 0.0029 and its standard deviation is 0.0027, so the slope difference is about 1.1σ. This difference, though close to the slope difference in the previous example (0.0031) is not statistically significant because the standard deviation of the pre-retrofit slope is too large due to relatively few data points. Our calculations suggest that increasing the number of good-quality pre-retrofit data points by a factor of about two would make the slope difference statistically significant at significance level of 0.08.

These examples are promising, yet large pre-post data samples and additional work are required to establish an EM&V method on the basis of the proposed approach.

7.3 Cases Without Fuel Bills

The requirement to have gas bills available can limit the applicability of the project's results because (1) a CT reward program may not be controlled by a gas utility, and (2) many homes, especially in the Northeast, are heated by delivered fuels, i.e., heating oil or propane.

Gas bills enable us to estimate the heat supply Q in Eqs. (2)-(5); this means that without bills, we need an alternate approach to estimate Q. All other things being equal, Q should scale approximately linearly with the home's conditioned area. Indeed, the estimated Q values reported in the HEAs show this dependence. Such a correlation implies a relatively narrow distribution of heat supply for a given conditioned area in a home which could ultimately result in a more definitive answer to the question of whether a given home with no gas bills is a good candidate for insulation and/or air sealing retrofit.

Suppose the probability density function of Q for a given home is known. Our approach so far uses a single value or a *point estimate* of Q to estimate the most important physical parameters

characterizing retrofit opportunities, the overall home R-value and ACH_{50}. Both parameters have a functional relationship with Q: the R-value is inversely proportional to Q (see Eq. (5)) and ACH_{50} is linearly proportional to Q (see Eqs. (9)-(12)). Thus, it is possible to derive a confidence interval for the estimated R-value or estimate the probability that the R-value is less than the threshold value (assumed to be 8 in our work). In this way, we could potentially overcome the need for fuel bills. Additional sources of uncertainty (e.g., those related to home measurements and model simplifications) can be accounted for as well, although our previous analysis related to the modeling of a home with two CTs shows very small variability related to these factors (see Section 4).

The statistical distribution of Q for a home can be characterized using a sample of similar homes with known Q values. As an example, we calculated values of Q for 628 homes from vendor #1, part of the RCT data set (see Section 6.1). With the limited information provided to us, we can only characterize "similarity" of homes by their floor space and the number of floors. Figure 23 shows the calculated values of Q for two-story homes.

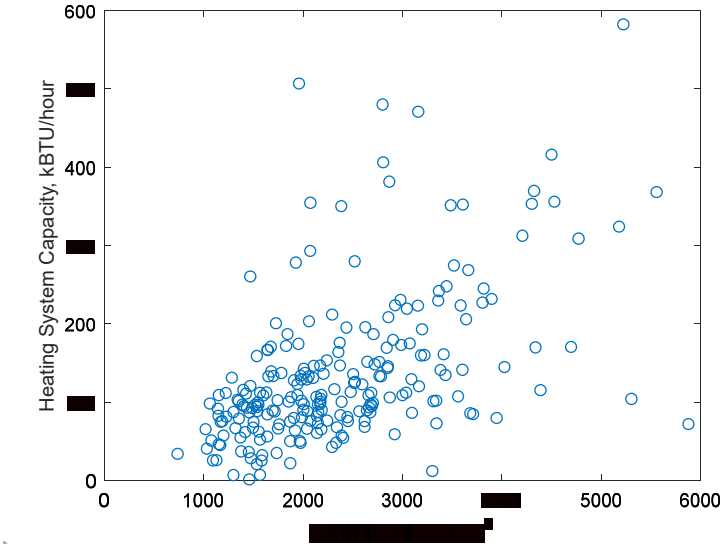

Figure 23. Calculated heat supply, Q, for a sample of two-story homes

The RCT data set from vendor #1 was used. Results for 230 homes are shown.

Suppose now we have a two-story home with an unknown heat supply from the same population, and its conditioned area is 2,000 ft^2. Ideally, to characterize the statistical distribution of this home's heat supply, we would down-select homes with the same conditioned area from our sample. Because the sample is not large, we allow an arbitrary tolerance of ±300 ft^2 for a home to be "similar." A histogram of the calculated heat supply for such homes is shown in Figure 24.

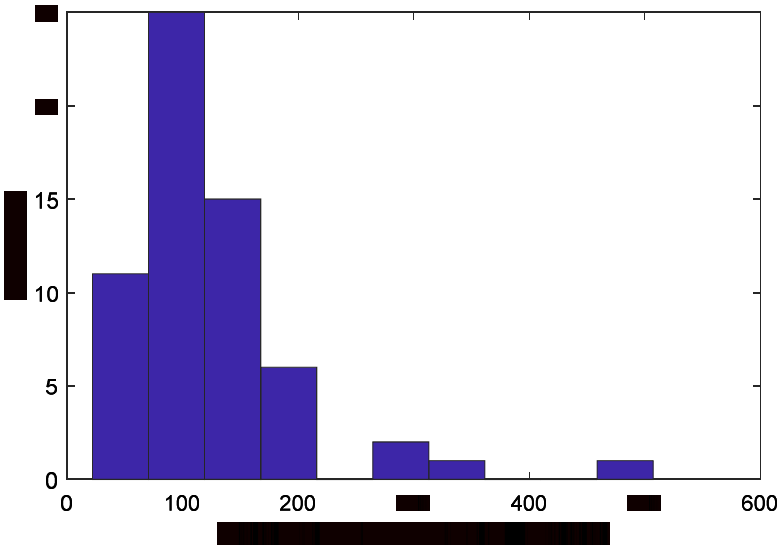

Figure 24. Distribution of calculated heat supply for a sample of two-story homes with conditioned area of about 2,000 ft²

The histogram is built from homes used in Figure 23 that have conditioned area of 2,000±300 ft² (101 in total).

It is seen in the figure that the distribution mode (i.e., the most frequent value) is about 100 kBTU/h. Suppose we used this value as a point estimate of Q and derived, using our approach and available CT data, the corresponding estimates for the overall R-value of 7.5 and ACH_{50} of 17.2 for this home. Using the empirical distribution (plotted in Figure 24), we can obtain the distribution of R-values and ACH_{50} corresponding to these two estimates.

These distributions (which can be considered to be statistical distributions of actually estimated values with the value of Q being known), are shown in Figure 25. As expected, Figure 25 suggests that there is a wide range of possible "actual" parameter estimates. Quantitatively, we can estimate the probability that this home is a candidate for insulation retrofit, meaning that the "actual" R-value <8. This probability equals the number of R-values <8 (58) divided by the total sample size (101), i.e., 57.4%. Analogously, the probability that the home is an air sealing candidate equals the number of ACH_{50} >10 (85) divided by the total sample size (101), i.e., 84.1%. Whether these probabilities are high enough to justify approaching this home with a retrofit offer would depend on the program design and could be decided, e.g., by calculating the chance that the expected savings exceed a threshold or by comparing the saving expectation with HEA cost. Such calculations could be performed along the lines of the proposed approach.

In practice, actual (i.e., nameplate) installed furnace and boiler capacities come in discrete capacities and depend on several factors, foremost contractor system sizing decisions. Consequently, using distributions of *actual* Q values derived from field data (e.g., from

completed retrofit projects) as a function of relevant variables (e.g., vintage)[16] would likely yield more accurate distributions.

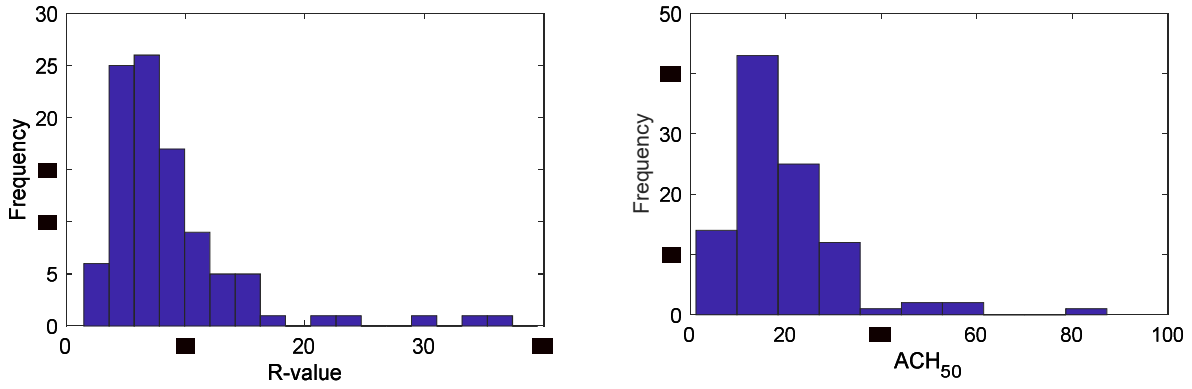

Figure 25. Distributions of potential values of R-values and ACH_{50} for a home with unknown heat supply

The distributions are obtained for a two-story home with 2,000 conditioned area and point estimates of 7.5 for R-value and 17.2 for ACH_{50}.

8 Conclusions

In this work, we successfully developed and validated a set of scalable algorithms that automatically analyze CT data to accurately characterize home retrofit opportunities and predict expected energy savings for homes with one or two CTs. Unlike other research groups, we use home energy assessment reports as the source of comparison and verification and a "static" (i.e., averaged over time) approach for estimation of physical home parameters that characterize the retrofit opportunity. Importantly, our physics-based approach is based on first principles, so it does not require "training" that would limit its applicability.

The results for the building envelope R-values and for air leakage characteristic ACH_{50} suggest that we can accurately identify low-performing homes that are prime candidates for insulation and air sealing retrofits, achieving classification accuracies of 89% and 96% for insulation and air sealing retrofits, respectively. Among scalable and computationally efficient approaches, ours can uniquely separate conduction heat losses from infiltration losses. Initially, we developed our algorithms for homes with a single CT and then subsequently successfully extended the approach to homes with two CTs. Thus, our method is generally applicable to homes using gas-fired furnaces and boilers controlled by CTs located in heating-dominant climates. We have not, however, evaluated the effectiveness of the algorithms for combi-systems, i.e., boilers that meet both space and water heating loads. Neither we were able to quantitatively characterize the heating system efficiency of a home.

That said, we found several challenges related to the quality of CT data and precision of on-site home energy assessments. In particular, obtaining high-quality CT data proved to be a major

[16] We presume that the Q estimates reported from the HEAs reflect algorithms based on these kinds of underlying data.

challenge, with two of the three vendors providing lower resolution data that negatively impacted algorithm effectiveness, particularly for EM&V. In addition, missing CT runtime data also appears to be a significant challenge. Consequently, utilities and energy efficiency programs need to carefully consider what data fields and resolution/precision they expect CT providers to report when developing CT procurement specifications to take advantage of emerging use cases such as remote home energy assessments.

Ultimately, the targeted and customized outreach enabled by the algorithms has the potential to increase the energy savings and cost-effectiveness of energy efficiency programs in multiple ways. First, we expect this will appreciably increase the number of HEAs requested by homes with significant (10%+ savings) insulation and air sealing opportunities, as well as the number of insulation and air sealing projects completed. Not only does this increase the quantify of savings realized, it also increases the cost-effectiveness of the HEAs because the savings per HEA increase. Second, the information provided by the algorithms can be used to focus the HEAs and, therefore, decrease their cost. Third, the algorithms can also be used by the programs to perform automated home-specific post-retrofit EM&V of retrofits performed by comparing actual post-retrofit thermal performance with expected performance. This will reduce program quality control and evaluation costs and should increase the ratio of actual to expected energy savings.

To better understand how targeted, customized outreach affects HEA and ECM uptake, we conducted an RCT field pilot with our utility partners. The RCT results suggest that personal messaging leads to a significant increase in the HEA uptake rate. Because of a time lag of up to a year between an HEA and ECM implementation in a home, it is too early to conclude whether the ECM realization rate also increased for such homes. Potential future work could involve gathering precise pre-retrofit field data measurements (insulation levels, air infiltration, and other characteristics) to further refine/validate the algorithm.

References

Afram, A. and F. Janabi-Sharifi. 2015. "Black-box modeling of residential HVAC system and comparison of gray-box and black-box modeling methods." *Energy and Buildings* 94, pp. 121-149.

Bacher, P. and H. Madsen. 2011. "Identifying suitable models for the heat dynamics of buildings," *Energy and Buildings* 43, pp. 1511–1522.

Barbour, N. 2021. "Smart thermostats gain traction in US, point to modest electricity savings." S&P Global Intelligence. July 21, 2021.

Berthou, T. 2013. "Développement de modèles de bâtiment pour la prévision de charge de climatisation et l'élaboration de stratégies d'optimisation énergétique et d'effacement" (in French). Ph.D. thesis.

Berthou, T. et al. 2014. "Development and validation of a gray box model to predict thermal behavior of occupied office buildings." *Energy and Buildings* 74, pp. 91-100.

Blasnik, M. 2018. Senior Building Scientist, Google Nest. Private communication.

BPI. 2012. "Technical Standards for the Building Analyst Professional." Building Performance Institute Standard.

Brand L., and W. Rose. 2012. "Measure Guideline: High Efficiency Natural Gas Furnaces." U.S. Department of Energy Office of Energy Efficiency and Renewable Energy, Building America Program. https://www.nrel.gov/docs/fy13osti/55493.pdf.

Chong, H. and A. George. 2018. "Buildings and Behavior: Rebounds Effects and other Inconvenient Findings in Residential Energy Consumption and Building Shell Performance." Working Paper. Dated 8/29/2016.

Cutler, D., et al. 2013. "Improved Modeling of Residential Air Conditioners and Heat Pumps for Energy Calculations." Golden, CO: National Renewable Energy Laboratory. https://www.nrel.gov/docs/fy13osti/56354.pdf.

DOE/EIA. 2009. "Annual Energy Outlook 2009." U.S. Department of Energy and the U.S. Energy Information Administration. Washington, D.C. www.eia.gov/outlooks/archive/aeo09/.

DOE/EIA. 2018. "Annual Energy Outlook 2018: Residential Sector Key Indicators and Consumption". U.S. Department of Energy and the U.S. Energy Information Administration. Washington, D.C. https://www.eia.gov/outlooks/aeo/.

Energy Assessment Standards. 2012. "Technical Standards for the Building Analyst Professional." Building Performance Institute Standard.

Fels, M. 1986. "PRISM: An introduction." *Energy and Buildings* 9, pp. 5–18.

Fisette, P. 2005. "Cellulose Insulation – a Smart Choice," https://bct.eco.umass.edu/publications/articles/cellulose-insulation-a-smart-choice/.

Gaasch, W.H., et al. 2014. "A Comparison of Methods for Early-Stage Retrofit Analyses," Proceedings of ACEEE Summer Study on Energy Efficiency in Buildings, pp. 11.149-161.

Goldman, E., et al. 2014. "Are Thermostats the New Energy Audits?" Efficiency Vermont. https://www.efficiencyvermont.com/news-blog/whitepapers/are-thermostats-the-new-energy-audits.

Goldman, E., et al. 2018. "Free X-Ray Specs—Just Send in Your Smart Thermostat Data for an Automated Peek Inside Your Home's Envelope!" In: Proceedings of ACEEE Summer Study on Energy Efficiency in Buildings Pacific Grove: ACEEE, pp. 12.1-12.12.

Hales, D. 2014. "Residential Ventilation in the Pacific Northwest." Washington State University Extension Energy Program. https://www.proctoreng.com/dnld/Hales-Ventilation.pdf.

Hallinan, K.P. et al. 2011. "Targeting Residential Energy Reduction for City Utilities Using Historical Electrical Utility Data and Readily Available Building Data." ASHRAE Transactions 117, p. 577.

Harish, V.S.K.V and A. Kumar. 2016. "Reduced order modeling and parameter identification of a building energy system model through an optimization routine." *Applied Energy*, p. 162.

Harley, B. 2011. "EnergyMeasure™ – Building Model Algorithms," Conservation Services Group (nonpublic).

Home Energy Services. 2018. "No-Cost Energy Assessments." Massachusetts. https://www.masssave.com/en/saving/energy-assessments.

Huang, J., J. Hanford, and F. Yang. 1999. "Residential Heating and Cooling Loads Component Analysis." Berkeley, CA: Lawrence Berkeley National Laboratory. Report #44636. https://simulationresearch.lbl.gov/dirpubs/44636.pdf.

Klint, P. 2018. Program Manager, Eversource. Private communication regarding Eversource Residential Program.

Krigger, J. and C. Dorsi. 2004. Residential Energy Cost Savings & Comfort for Existing Buildings. *Pearson*, 4th Edition, p. 284.

Lazrak, A. and M. Zeifman. 2017. "Estimation of Physical Buildings Parameters Using Interval Thermostat Data." Proceedings of 4th ACM International Conference on Systems for Energy-Efficient Built Environments (BuildSys 2017), Article No. 22.

Lee, S. H. and T. Hong. 2017. "Leveraging Zone Air Temperature Data to Improve Physics-Based Energy Simulation of Existing Buildings." Lawrence Berkeley National Laboratory, published in the 15th IBPSA Conference Proceedings. San Francisco, CA. https://eta-publications.lbl.gov/sites/default/files/bs2017_137.pdf.

Lin, Y., T. Middelkoop, and P. Barooah. 2012. "Issues in identification of control-oriented thermal models of zones in multi-zone buildings." 51st IEEE Conference on Decision and Control. Maui, Hawaii. https://ieeexplore.ieee.org/document/6425958.

Ljung, L. 2017. "System Identification Toolbox™ User's Guide." Natick: The MathWorks, Inc.

Lstiburek, J. 2010. "BSI-030: Advanced Framing." Building Science Corporation. Building Science Insight. https://www.buildingscience.com/documents/insights/bsi-030-advanced-framing.

Mark, N. et al. 2016. "Bridging the Gap Between Direct Install and Whole House Programs: Minneapolis Home Energy Squad Residential Engagement Pilot" Proceedings of ACEEE Summer Study on Energy Efficiency in Buildings Pacific Grove: ACEEE, pp. 2.1-2.12.

Massachusetts DPU. 2016. "2016-2018 Energy Efficiency Plan Order." Massachusetts Department of Public Utilities. https://www.mass.gov/doc/2016-2018-three-year-plans-order/download.

Mejri et al. 2011. "Energy performance assessment of occupied buildings using model identification techniques." *Energy and Buildings*, 43, pp. 285-299.

New, J.R., et al. 2012. "Autotune E+ Building Energy Models," Proc. 5th National SimBuild of IBPSA-USA.

Newsham, G. et al. 2017. "Remote energy auditing: Energy efficiency through smart thermostat data and control." In ECEEE Summer Study proceedings. https://www.eceee.org/library/conference_proceedings/eceee_Summer_Studies/2017/5-buildings-and-construction-technologies-and-systems/remote-energy-auditing-energy-efficiency-through-smart-thermostat-data-and-control/.

Opinion Dynamics Corporation. 2009. "Massachusetts Residential Appliance Saturation Survey (RASS): Volume 1: Summary Results and Analysis." Final Report. https://ma-eeac.org/wp-content/uploads/11_MA-Residential-Appliance-Saturation-Survey_Vol_1.pdf.

Pathak, N., et al. 2019. "A Bayesian Data Analytics Approach to Buildings' Thermal Parameter Estimation," Proceedings of the Tenth ACM International Conference on Future Energy Systems, pp. 89-99.

Peeters, L. et al. 2008. "Control of heating systems in residential buildings: Current practice." *Energy and Buildings* 40, p. 1446-1455.

Roth, K. and M. Zeifman. 2020. "Using Communicating Thermostats to Identify Home Insulation and 1 Air Sealing Retrofit Opportunities: Best Practices Guide for Scale Up." Prepared for the U.S. Department of Energy Building America Program, Office of Energy Efficiency and Renewable Energy. January 2020.

S3C. 2019. "Energy audits for households." Guidelines by S3C. https://www.smartgrid-engagement-toolkit.eu/learning/list-of-guidelines-and-tools/.

Siemann, M.J. 2013. "Performance and Applications of Residential Building Energy Grey-Box Models." Ph.D. Thesis, University of Maryland. College Park, Maryland.

Walker, I.S. and D.J. Wilson. 1990. "The Alberta air infiltration model (AIM-2)." Alberta University. https://www.aivc.org/sites/default/files/airbase_3705.pdf.

Walker, I. 2017. Leader, Residential Building Systems Group at Lawrence Berkeley National Laboratory. Private communication.

Yang, S., et al. 2018. "A state-space thermal model incorporating humidity and thermal comfort for model predictive control in buildings," *Energy & Buildings* 170, pp. 25-39.

Zeifman, M. and K. Roth. 2016. "Residential Remote Audits: Estimation of Building Thermal Parameters Using Interval Energy Consumption Data." Proceedings of ACEEE Summer Study on Energy Efficiency in Buildings. Pacific Grove: ACEEE, pp. 12.1-12.12.

Zeifman, M., K. Roth, and B. Urban. 2017. "Communicating Thermostats as a Tool for Home Energy Performance Assessment." Proceedings of 35th IEEE International Conference on Consumer Electronics. Las Vegas: IEEE, pp. 191-192.

Zeifman, M., A. Lazrak, and K. Roth. 2018. "Residential retrofits at scale: opportunity identification, saving estimation and personalized messaging based on communicating thermostat data." Proceedings of ACEEE Summer Study on Energy Efficiency in Buildings Pacific Grove: ACEEE, pp. 12.1-12.13.

Zeifman, M., A. Lazrak, and K. Roth. 2020a. "Residential retrofits at scale: opportunity identification, saving estimation and personalized messaging based on communicating thermostat data." *Energy Efficiency* 13, pp. 393-405.

Zeifman, M., A. Lazrak, and K. Roth. 2020b. "When Two Isn't Better Than One: Algorithms to Remotely Identify Insulation and Air Sealing Retrofit Opportunities in Homes with Two Communicating Thermostats." Proceedings of ACEEE Summer Study on Energy Efficiency in Buildings Pacific Grove: ACEEE, pp. 12.467-12.478.

U.S. DEPARTMENT OF ENERGY

Office of
ENERGY EFFICIENCY & RENEWABLE ENERGY

For more information, visit:
energy.gov/eere/buildings/building-america

DOE/GO-102022-5750 ▪ July 2022