Energy and Buildings 80 (2014) 556–561

Contents lists available at ScienceDirect

Energy and Buildings

journal homepage: www.elsevier.com/locate/enbuild

Developing a pre-retrofit energy consumption metric to model post-retrofit energy savings: Phase one of a three-phase research initiative

Kate Goldstein [a], Michael Blasnik [b], Michael Heaney [c,*], Ben Polly [c], Craig Christensen [c], Les Norford [a]

[a] Massachusetts Institute of Technology, United States
[b] Michael Blasnik and Associates, United States
[c] National Renewable Energy Laboratory, United States

ARTICLE INFO

Article history:
Received 13 November 2013
Accepted 31 March 2014
Available online 6 May 2014

Keywords:
Single family residential
Energy efficiency
Data
Retrofit
Utility programs

ABSTRACT

This paper details the process and results from the first step of a three-step research process. This first step looks to identify the most predictive pre-retrofit metric of energy consumption to utilize in a model to predict the energy savings post retrofit. The ultimate goal of this research is to predict candidacy for retrofit using only a combination of demographic and home-characteristics data that is available for the entirety of the U.S. residential housing stock. This is important, as utility data is almost always protected for privacy and thus unavailable to assist in targeting where energy efficiency retrofits will be successful.

It is found that the best metric is the simplest, total energy consumption divided by total floor area. In addition to evaluating which pre-use metric is most indicative of post retrofit savings, the paper evaluates the endogenous component of pre-use to post use and a potential method to alleviate this endogeneity. The research finds that by removing the year that is used to calculate the savings as the baseline pre-use year removes a portion of the endogeneity. It is also found that one year before the savings base year is the best year to utilize as the base.

1. Introduction

This paper details the process and results from the first step of a three-step research process. This first step looks to identify the most predictive pre-retrofit metric of energy consumption to utilize in a model to predict the energy savings that will occur post retrofit. The ultimate goal of this research is to predict a home's candidacy for retrofit using only a combination of demographic and home-characteristics data that is available for the entirety of the U.S. residential housing stock. This is important, as utility data is almost always protected for privacy and thus unavailable to assist in targeting where energy efficiency retrofits will be successful. However, in order to build and verify this model, it is necessary to have actual data on both what homes consume, and how much homes that have undergone retrofits have saved as a result of the work. As a very small fraction (less than one percent) of homes in the US have actually completed retrofits, the best way to predict for retrofit potential is to use a sample of homes where energy data is available to build a model for pre-retrofit consumption or energy intensity. Further, in order to verify that this energy intensity is indeed indicative of retrofit savings potential, another (albeit much smaller) dataset is necessary for which both pre-retrofit and post-retrofit data is available.

Available to the research, as diagrammed in what follows, is 6.5 years of gas billing data (monthly bills from 2007 to 2013) for about 1 million customers in Ohio. This data essentially lists the number of therms consumed per month, per customer. Approximately 13,000 of these homes underwent an energy efficiency retrofit in 2011. The work completed, cost of work completed, projected savings, change in characteristics as a result of the retrofit are all also available for these homes. For each of the following research steps, "pre-use" and "post-use" refers to a weather normalized annual gas consumption metric which is calculated through a billing analysis of monthly utility bills in combination with recent and long-term weather data. A Pre-use and post-use metric refers to a particular formula used to characterize the consumption before and after retrofit. For example, weather normalized annual consumption could include just the component of the gas usage attributable to heating

* Corresponding author at: National Renewable Energy Laboratory, 15013 Denver West Parkway, Golden, CO 80401, USA. Tel.: +1 401 345 6543.
E-mail address: Michael.Heaney@nrel.gov (M. Heaney).

http://dx.doi.org/10.1016/j.enbuild.2014.03.068
0378-7788/

demand, or it could also include baseload. Furthermore, each of these values could be further normalized by the area, volume, or surface area of the home to yield "energy intensity", or, average therms consumed per year per square foot of home. Further detail on the billing analysis methodology, variants in the classic methodology utilized in this research, and the various different ways to quantify energy intensity pre-retrofit are discussed in Section 4.

The research workflow takes the following form:

(1) Find what pre retrofit-use metric is most indicative of retrofit savings potential – how does pre-retrofit usage predict post-retrofit usage?

a. Use a sub-set of homes for which both pre and post retrofit consumption is available (~7000 homes)
b. Test various pre retrofit-use metrics and their correlation with retrofit savings (utility data post-retrofit)
c. Choose metric with best fit and explanatory power

(2) Build a model to predict this chosen pre-use metric using home characteristics and demographics – how do home characteristics and demographics predict home energy consumption (pre-usage)?

a. Use a large dataset of representative homes where energy data is available (~1,000,000 homes)
b. Test multiple data mining and modeling techniques for predictive power (utility data pre-retrofit)
c. Choose model(s) and validate candidacy with sub-sample of homes that have both pre-use and post-use and underwent retrofit

(3) Scale this model up to homes for which no energy data is available

a. Use chosen modeling technique to predict chosen pre-use metric
b. Validate model using homes left out of model development as test sample
c. Quantify scalability of model to other regions/demographics/home-stocks (independent sample of homes in different region with pre-use/post-use/retrofit data available).

This particular paper focuses on step 1 of this process (the blue section in Fig. 1), using the sub-set of data with both pre and post retrofit consumption to validate the pre-use metric most predictive of retrofit savings potential. Completing this first phase is essential to rest of the project; and underlying assumption of the research flow is that pre-retrofit consumption is an effective proxy for retrofit potential. Further, many of the analysis techniques utilized during this phase of the research will be extended to the subsequent phases.

2. Motivation

The market potential for residential energy efficiency retrofits is extremely high. Buildings represent 40% of the United States' primary energy consumption [1]. According to the Rockefeller Foundation [2], building energy retrofits represent a $279 billion dollar investment opportunity with potential savings over 10 years of $1 trillion dollars. The avoided carbon amounts to 600 million metric tons of CO_2 per year or, 10% of the total U.S. emissions in 2010. In addition, of the 3000 trillion Btu of energy savings potential, the residential sector alone represents almost 2000 trillion Btu of this possibility, and almost $200 billion of the $300 billion investment opportunity.

The residential sector is also the most fragmented opportunity, with small diffuse opportunities that are difficult to aggregate to a substantial scale. While single-family homes are 88% of the total residential housing stock and use 80% of the total energy, they are also the smallest consumers per square foot of real estate; the usage is extremely diffuse. In addition [3], notes that 71% of end-use potential in the residential sector are tied up in improving the building shell and upgrading the heating and cooling equipment in existing homes.

In order to solve energy efficiency, energy efficiency in buildings must be solved. To solve energy efficiency in buildings, energy efficiency in single-family homes must be solved. In order to succeed, this must be accomplished at great scale, but the solutions are well known and do not vary substantially across the sector. Fix the shell, and upgrade HVAC systems. Historically, the pure lack of data on the United States' housing stock has been one of the primary barriers to market penetration of residential energy efficiency retrofits. Without knowledge of the homes and customers to reach, outreach has been untargeted and inefficient. As such, what is needed is a model to help those on the ground perform better, more targeted outreach for retrofit measures, leading to higher audit to retrofit conversion rates, and to help inform the very initial steps of the energy audit process. This research does not pretend to approach the equally critical factor of home-owner engagement. Figuring out whether or not a home has the potential for energy savings is just the first step to completing the retrofit. The home owner has to both agree to, and then pay for the retrofit (or some of the retrofit) once their home has been identified.

3. Literature review

In order to take raw energy consumption data and turn this into useful information regarding how a home uses energy on an annualized and weather normalized basis, a billing analysis must be completed. This analysis combines long term historical weather data, recent weather data corresponding to the usage period, and the monthly bills. With this information, an estimate of the gas used for baseload (cooking/water-heating), an estimate of the gas used for home-heating, and an indoor balance point temperature (related both to the indoor set-point and internal loads) are extracted. Furthermore, a metric for the amount of therms used per heating degree day (an aggregated measure indicating indoor–outdoor temperature differences) can be scaled utilizing the long-term weather data to estimate the energy consumed for heating energy in the typical year (normalized annual heating consumption). An estimate of total gas consumption in a normalized year can also be estimated by adding the yearly baseload estimate to the normalized annual heating consumption to get a normalized annual consumption. The historic development of this type of analysis and the ways in which the analysis used in this research differs from the original is discussed in what follows.

The PRISM method, developed in the late 1970s and greatly expanded upon by Margaret Fels in 1986, was originally designed to evaluate the effectiveness of home energy retrofit programs. The method utilizes a set of customer utility bills in addition to daily averaged outdoor temperatures both before and after retrofit to generate a weather-adjusted indicator of a home's performance called the Normalized Annual Consumption [4]. This statistical method is also applicable when simply evaluating a home's candidacy for retrofit and results in three critical numbers describing the home's energy usage: the reference temperature, the baseload, and an indicator of the home's heating system and envelope performance, the heating slope. The model plots the average daily outdoor temperature against the rate of gas consumption, *f*. The amount of