DATA-DRIVEN RANKING OF BUILDING ENERGY EFFICIENCY UTILIZING STOCHASTIC ENERGY EFFICIENCY FRONTIERS (SEEF)

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF
CIVIL AND ENVINRONMENTAL ENGINEERING
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF
PHILOSOPHY

Amir Ehsan Kavousian

June 2014

Abstract

Quantifying building energy efficiency is essential to prioritizing investments and monitoring energy efficiency programs for utilities and large-scale facility managers. Two inherent challenges of quantifying energy efficiency are (a) energy consumption is influenced by many different factors, some of which are outside the scope of influence or interest of users to change; and (b) energy consumption is highly variable, hence energy efficiency can also considerably change over time. The contribution of my research is a computational method that utilizes smart meter data to rank buildings' energy efficiency while (a) removing the effect of factors that are outside the scope of influence and interest of users to change, and (b) accounting for the fluctuation in energy consumption, by specifying a confidence interval for energy efficiency estimates. Particularly, I extend frontier methods for energy efficiency ranking by developing Stochastic Energy Efficiency Frontiers (SEEF), a method to estimate the uncertainty in efficiency rankings and validate the functional form of the frontier. I validate the power and generality of my method using three data sets in residential and commercial settings. I demonstrate how SEEF helps identify factors that drive energy efficiency, and use its result to identify interventions that contribute to improvement to energy efficiency. For instance, in residential buildings, SEEF results validate that certain efficient features such as double-pane windows are more correlated with higher efficiency ranks, while other factors such as programmable thermostats have a smaller impact on rankings. The results of my research will enable practitioners to compare energy efficiency of buildings more accurately, and to compare them over time.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank my family, parents Hamid and Maryam and brothers Ali and Pouya for their unconditional love and support throughout my studies. Higher education is one of our family's most respected values, and I am happy to share this moment with my two brothers, Pouya who received his doctorate in Electrical Engineering from Stanford University, and Ali who is working toward his MBA degree at McGill University, Canada. My girlfriend, Mitra, was a big source of joy, motivation, and emotional and intellectual support during my doctoral studies. We also met at Stanford when she was working on her Masters in Management Science and Engineering. I am grateful to her and would like to share this moment with her.

My advisors at Stanford, Professors Martin Fischer and Ram Rajagopal have been instrumental in my research, as well as personal, intellectual, and professional growth. In particular, I would like to thank Professor Fischer for his support when I decided to change my focus to energy four years ago. His style of advising—giving his students autonomy and responsibility to devise and manage their own research program while maintaining very high academic standards—considerably helped me honing in my research skills while maintaining my motivation. I would also like to thank Professor Rajagopal for consistently raising the bar for me and mentoring me to push myself beyond what I could imagine I would be able to do. For that, I will always be grateful to him.

I would also like to thank my committee members, Professor James Sweeney, Doctor June Flora, and Professor John Weyant. Professors James Sweeney has been a great mentor and advisor since the beginning of my PhD research. His accomplishments and impact on the industry was an important source of inspiration when I decided to focus on the energy industry. Doctor June Flora brought a unique behavioral perspective to my personal framework of energy efficiency. Her knowledge of human behavior and the impact of attitudes toward energy saving on people's behavior and habits were valuable resources throughout my doctoral studies. Professor John Weyant has been another source of inspiration for me to focus on the energy industry. His class, MS&E 295 was instrumental for me in gaining an appreciation of how small changes in energy consumption patterns can create significant effects on the industry as a whole. I highly recommend MS&E 295 to anyone who wants to get an appreciation of the interplay of efficiency, clean tech, and the environmental impacts.

My colleagues and friends at the Center for Integrated Facility Engineering (CIFE), Stanford Sustainable Systems Lab (S3L), and Precourt Energy Efficiency Center have made my PhD journey much more joyful and manageable. To name a few, I would like to thank Adrian Albert, Alissa Cooperman, Forest Flager, Nelly Garcia Lopez, Robert Graebert, Taewan Kim, Jungsuk Kwac, Roshan Mehdizadeh, Rene Morkos, Forest Peterson, Pat Shiel, and Jasmine Wei. I apologize in advance for inadvertently leaving out any names; I am certainly grateful for their support.

Last but not least, I would like to thank the organizations that supported my work financially and institutionally. My doctoral work at different points in time has been supported by the Center for Integrated Facility Engineering (CIFE), Stanford Sustainable Systems Lab (S3L), and Precourt Energy Efficiency Center.

Chapter 1

Data-Driven Energy Efficiency in Buildings – An Introduction to This Thesis

Residential and commercial buildings account for about 41% of total primary energy consumption in the US. More than 75% of this energy is procured using fossil fuels (US DOE, 2014) (Figure 1). It is economically and technically feasible to reduce energy consumption of buildings by 30%, using existing technologies; this figure is expected to reach 80% by 2020 (Farese et al., 2012).

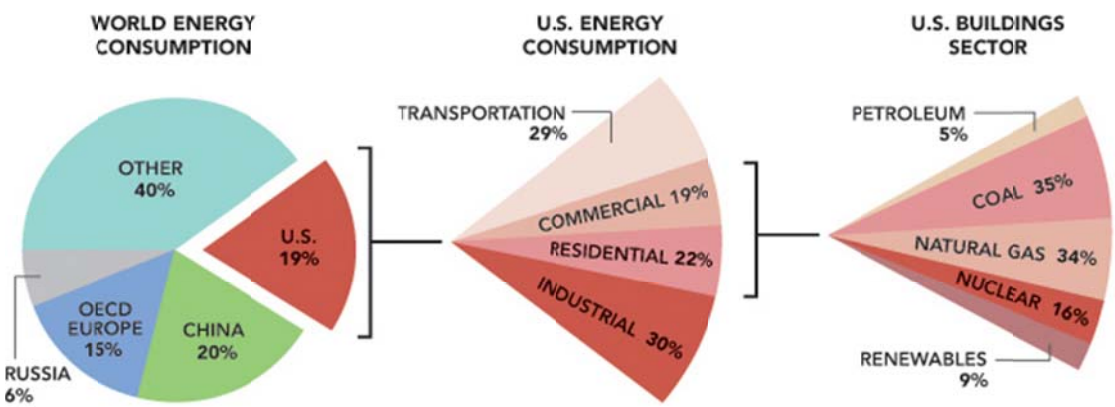


Figure 1 - A Breakdown of US energy consumption, with an emphasis on the buildings sector [figure source: US Department of Energy].

Building energy retrofits in the US offer a $279 billion dollar investment opportunity, which can lead to creating more than 3.3 million new cumulative job years, and mitigating more than 600 million metric tons of CO2 per year equal to about 10% of US emissions in 2010 (McKinsey, 2009; CAP, 2009; EIA CBECS, 2003; EIA RECS, 2009; in Rockefeller Foundation, 2012). However, research shows that most energy efficiency projects do not achieve their energy reduction goals. Some efficiency projects even result in increased consumption (Deutsche Bank & Living Cities, 2011;

Shapiro, 2011). Ranking energy efficiency is essential to the improvement of energy performance and prioritization of investments, and is a crucial step in planning, monitoring, and verifying energy efficiency programs.

1 Building Energy Efficiency Ranking: Definition, Goals, and Requirements

Building energy efficiency ranking is an ongoing review of energy performance to determine if a building is getting better or worse compared to itself or other buildings in the portfolio[1].

The problem of ranking buildings' energy efficiency follows two goals: (a) identify efficient and inefficient buildings, and (b) understand what drives efficiency or inefficiency. The former guides the selection of buildings to retrofit, and the latter informs the focus of energy efficiency policies and the choice of retrofit projects to implement on selected buildings.

Furthermore, energy efficiency ranking methods need to be *actionable* and *unbiased*. To be actionable, ranking methods need to not only rank buildings based on their energy performance, but also to help practitioners make decisions and take specific actions to improve the energy efficiency of buildings. The actions inferred from the ranking methods need to be feasible and within the control and interest of the users to change. For example, as Chapter 3 of the thesis explains, my research shows that families with pets consistently use more energy than similar families without a pet.

[1] In this thesis, the terms "building energy efficiency ranking", "quantifying energy efficiency", and "benchmarking" are used interchangeably, and are sometimes abbreviated as "ranking methods".

However, the suggestion to encourage families with pets to switch to not having a pet is not practical. On the other hand, the conclusion that buildings with programmable thermostats only use slightly less energy than similar houses without programmable thermostats can be a valuable insight in prioritizing energy efficiency investments and resource allocation. Therefore, an actionable ranking method should remove the effect of immutable factors before ranking buildings, and find the correlation of energy efficiency rankings with other factors after ranking. Chapter 4 of the thesis explains this process in detail.

To be unbiased, ranking methods need to remove the effect of random fluctuations of energy consumption or inherent advantages of users (such as enjoying favorable weather or a particular lifestyle) from energy efficiency ranks. The effect of inherent advantages on energy consumption can be removed together with removing the effect of immutable factors, as explained above. The effect of random fluctuations can be accounted for by providing a measure of the uncertainty of the ranks. This process is explained in Chapter 3 of the thesis, and illustrated using an example in Chapter 4.

In short, an unbiased and actionable ranking method should remove the effect of certain inherent and immutable variables and account for the effect of random fluctuations in energy consumption in its estimates of energy efficiency.

Prior to the introduction of Advanced Metering Infrastructure (AMI), monthly bills used to be the only source of information on energy consumption of residential and small to medium businesses. Recently, the deployment of smart meters has enabled utilities and building owners to have access to a wealth of consumption data to better account for the effects mentioned above. However, computational methods to process smart meter data and generate energy efficiency rankings that are actionable and unbiased are missing. The contribution of my doctoral research is a computational method to utilize smart meter data to improve building energy efficiency ranking, called the Stochastic Energy Efficiency Frontiers (SEEF) method.

In the next sections of this chapter, I will explain the existing methods for building energy efficiency ranking and identify their shortcomings which led to my research questions. I then describe the research tasks to address those research questions and explain my claimed contributions.

2 Background on Residential Building Energy Consumption

A building can be modeled as a dynamic system: it receives energy in the form of electricity, gas, etc. and provides a range of services such as air conditioning, refrigeration, and lighting to its occupants. The input-output relationship between energy consumption and services is influenced by several internal and external factors such as climate, building size and type, appliances, and behavior. At another level of influence, the socioeconomic and demographic status of the household can affect the building size and type, appliance selection, and behavior towards energy consumption and efficiency (Figure 2).

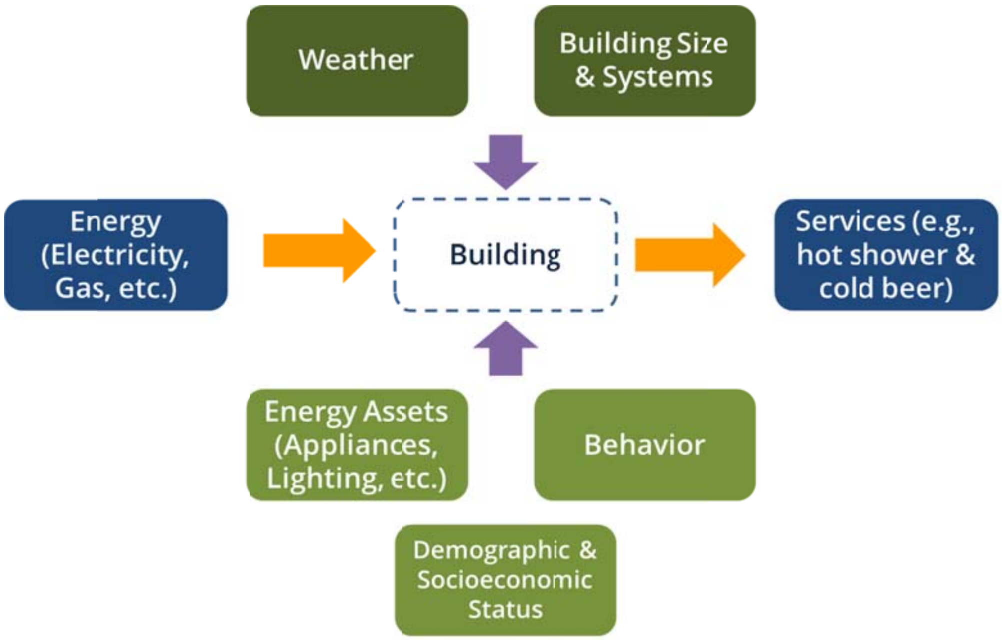


Figure 2 – A schematic of residential buildings' energy consumption. Blue boxes show inputs and outputs of a house as an energy system; dark green boxes show the external factors that impact consumption; light green boxes show internal factors that influence consumption.

Service is a central concept in my thesis; thus, I offer a more detailed description of how I quantify service. I measure Level of Service (LOS) in residential buildings as a weighted sum of the appliances in the building, with the weight of each appliance being proportional to the energy intensity of that appliance. The energy intensity values are estimated using industry averages of annual energy consumption of each appliance, and consider both the power and expected usage hours of the appliance. The weighted average explained above is called Appliance Ownership Index (AOI) (Capasso et al., 1994). Furthermore, I normalize AOI by dividing each user's AOI by the maximum AOI value in the population to be able to cross-compare users across different populations.

$$AOI_n = \sum_{i=1}^{k} A_{i,n} E_i$$

AOI_n: Appliance Ownership Index of user *n*

$A_{i,n}$: the count of appliance *i* that the household *n* owns

E_i: expected (industry average) energy consumption of appliance *i*

$AOI *_n = AOI_n \ / \ \max_n AOI_n$ (Normalized to range [0,1])

Having established a simple model of energy consumption in buildings, I will provide a review of the most important energy efficiency benchmarking methods.

3 A Review of Benchmarking Methods

Building energy efficiency benchmarking methods can be categorized and compared based on different metrics and characteristics. For example, at a very high level, benchmarking methods are divided into longitudinal and cross-sectional methods (Granderson, 2011). Longitudinal methods compare a building with its own performance over time, while cross-sectional methods compare a group of buildings, normally all within the same time interval. Longitudinal benchmarking methods are primarily used in measurement and verification (M&V) of savings after energy efficiency programs. Cross-sectional methods, on the other hand, compare the energy efficiency of buildings at a particular point in time. In the rest of this thesis, whenever I use the term benchmarking, I refer to cross-sectional benchmarking.

Another common division of benchmarking methods is theoretical versus empirical. Theoretical methods compare buildings' performance with a theoretical model of the building, while empirical models compare observed values of energy consumption.

Theoretical models can be further categorized based on the type of the computations involved; e.g., thermodynamic, and statistical models.

Table 1 summarizes the building ranking categories described above. A more detailed comparison of theoretical and empirical benchmarking methods is offered in the following section.

Table 1 - A classification of benchmarking methods.

Benchmarking Basis	Type of Benchmark	Method	Example
Longitudinal	Theoretical	Thermodynamic	EnergyPlus
		Statistical	IPMVP
Cross-Sectional	Theoretical	Thermodynamic	EnergyPlus
		Statistical	Energy Star
	Empirical	Index Model	Energy Intensity Method
		Frontier Model	Data Envelopment Analysis

3.1 Theoretical Benchmarking Models

Theoretical benchmarks are hypothetical cases that are calculated using engineering, thermodynamic, or statistical methods using a theoretical understanding of the relationship between energy consumption and its determining factors (Clevenger et al. 2012). These models utilize buildings' design information and industry averages of the fixtures, as well climate and assumptions about occupant behavior to predict energy usage (Xydis et al., 2008). For example, Griffith & Crawley (2006) explain a method for estimating the technical potential for energy efficiency in buildings using detailed energy modeling. Thermodynamic models create a detailed thermodynamic (bottom-up) model of the building, while statistical models create a top-down statistical model of the building based on a few variables such as weather and occupancy.

Statistical models use historical data to create a statistical model that predicts energy consumption of the *average* building based on building characteristics and climate. Then, the measured energy consumption of the building of interest is compared with the consumption of the average building. Any deviation from the average building is assumed to be due to the efficiency or inefficiency of the building of interest. For instance, Energy Star is a regression-based method for quantifying energy efficiency (EPA, 2012). It compares the energy consumption of a building to the average building of similar characteristics, calculated using a regression model. For office buildings, the regression model has the following seven variables: square foot, number of occupants, operating hours, number of workers per square foot, location and the amount of conditioned and heated space. The majority of statistical methods depend on regression models, and hence are limited by the shortcomings of ordinary regression (e.g. Atkinson, 1982; Riani & Atkinson, 2000; Savage, 2009). For example, single-building consumption models have reported errors of up to 41% (NREL, 2011).

Depending on the level of detail of the model, availability of reliable data, quality of assumptions, and the accuracy of the simulation engine, theoretical models can be very accurate in identifying sources of inefficiency in a building (Griffith & Crawley, 2006; Pisello et al. 2012; Waltner et al., 2012; Goldstein, 2012). However, such detailed models are time and skill intensive to build and maintain, hence cost prohibitive in most instances (Patterson 1995; Yezioro et al. 2008; Maile, 2009).

3.2 Empirical Benchmarking Models

Empirical models utilize measured data (usually from a population of buildings) to compare and rank building energy efficiency. These models can be based on comparing a few metrics across buildings (e.g., energy per square foot); or, they can be based on the identification of the most efficient buildings and comparing all other buildings with them (e.g., frontier models). While empirical models use a realistic benchmark against which to compare any building, they can be limited by data availability. With the advent of Advanced Metering Infrastructure (AMI) and development of statistical methods specialized for analyzing AMI data, empirical models are becoming more practical for benchmarking purposes. Two important empirical models that will be explained in more detail in the following sections are Index Models and Frontier Models.

3.2.1 Index Models

The most common form of empirical models are Index Models. Index Models compare a common metric such as energy per square foot across different buildings. Some other examples of metrics that have been used in Index Models are aggregate energy consumption as reported by the meter, normalized energy consumption (i.e., the effect of weather and seasonal effects are removed), and energy intensity metrics (energy consumption divided by floor area, number of occupants, heating/cooling degree days, etc.). Index Models have also been used for macro-level analysis purposes (for a review, see Phylipsen et al., 1997). These macro-level indices decompose the energy consumption of a unit (be it a building, a city, an industry, or a country) and then quantify its energy efficiency based on that decomposition, hence the title "Index Decomposition Analysis". Some examples of Index Decomposition Analysis metrics are: Divisia index, Laspeyres index, Malmquist Factor index. For more detailed analysis of the pros and cons of each index, see Boyd et al. (1988) and Greening et al. (1997).

In the absence of detailed data, Index Models need to make strong assumptions about the dynamics of energy consumption without having enough data to validate such assumptions. For example, a common assumption behind Index Models is the assumption of linearity of relationship between consumption and metrics such as floor area (e.g., when using energy per square foot as a comparison metric). Such assumptions often ignore the complexities of energy consumption processes in buildings. For example, Schipper (1983) questioned the value of using energy intensity (energy per square foot) to compare energy efficiency and calls for more comprehensive energy efficiency indices. Battles (1997) summarized the challenges in defining indices for measuring energy efficiency. Particularly, Battles highlighted data limitations and diversity of possible energy-efficiency indicators (i.e., diversity of metrics by which owners perceive the value of their buildings) as major challenges in defining a uniform energy efficiency index. Moreover, Amory Lovins (2004) summarized the complexity of energy efficiency measurement by highlighting how different users get different utilities from using energy fixtures. Huntington (1995) described the challenges in quantifying energy efficiency and the shortcomings of

Index Models. He suggested the use of non-parametric models (such as frontier models) to deal with those complexities and shortcomings. Previous research has also shown that buildings are not linear systems, and as buildings grow in size, the processes with which they consume energy changes (Kavousian et al. 2012). For instance, the volume of the house may increase with a different rate than the square footage; or, the penetration of certain equipment such as pools will change as the size changes. Another example of how and Index Model such as energy per square foot can be a misleading metric is the comparison of two coffee shops, one being frequently busy and filled by people who spend longer time there, while the other one being mostly quiet and with fewer customers. The first coffee shop in this case is expected to use more energy and should not be penalized for using more energy serving more customers. Frontier methods solve this problem by comparing the energy consumption of buildings of the same service level.

Another class of empirical benchmarking models, called Frontier Methods, quantify the energy efficiency of buildings by forming an efficient frontier (best-practice technology) and comparing all buildings against that frontier. Frontier methods offer several advantages over other empirical models, and will be discussed in more detail in the following section.

3.2.2 Frontier Methods

Frontier methods have been used in finance and economics to specify a tradeoff between two or more metrics, such as returns versus risk or output versus input (Pareto, 1906; Farrell, 1957). As described in previous sections, in a buildings context, Level of Service (LOS) of a building can be used as a proxy for the output of the building, while energy consumption can be used as a proxy for the input to the building. Frontier methods for energy efficiency benchmarking are based on the tradeoff between energy and LOS.

As Figure 3 illustrates, the plot of LOS versus normalized energy consumption (the effect of weather and seasonal effects are removed from consumption). LOS values are calculated as explained above: a building with LOS of 1 has the highest appliance ownership index in the set (estimated based on energy intensity of appliances); therefore, it is expected to consume the most amount of energy on its appliances across the entire data set. A building with LOS of 0.5 is expected to consume half as much appliance energy as the building with LOS of 1, and so on. The figure shows a general increase in energy consumption for increase in LOS. However, for any given LOS, there is a minimum energy consumption level achievable. This minimum energy level for a given LOS can be represented by a boundary, which I call energy efficiency frontier.

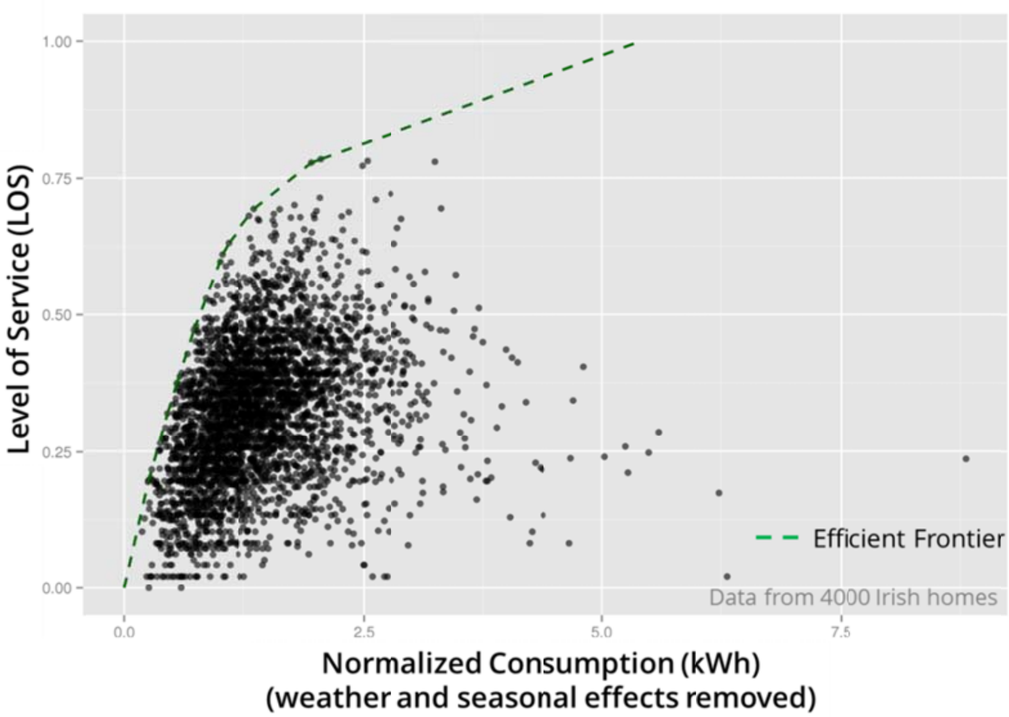


Figure 3 – While energy consumption tends to increase with increase in LOS, for each LOS there is a minimum energy consumption that is achievable.

Once the efficient frontier is established, the energy efficiency of each user can be estimated by comparing its consumption against the energy consumption of the corresponding user on the frontier. In 1957, Farrell developed mathematical relationships for calculating efficiency, which were all based on this simple equation (see Figure 4 for a description of the terms used in the equation):

Efficiency of user $\Theta_n := e^* / e_n$

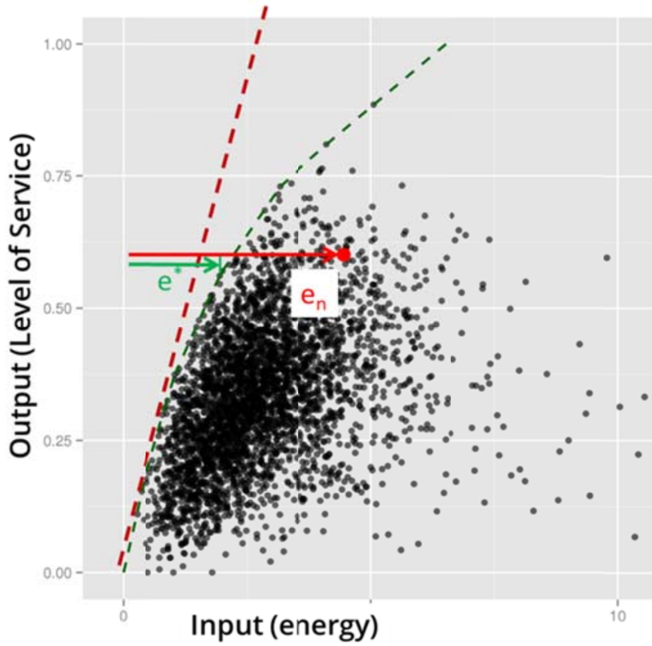


Figure 4 – Schematic of efficiency calculations using efficient frontiers. The green frontier is directly fitted to the data, while the red frontier assumes a linear form for the frontier.

Frontier methods can be both parametric and non-parametric. Parametric frontier methods, such as Stochastic Frontier Analysis (SFA), assume that the efficient frontier can be presented using a mathematical formula plus a stochastic error term called random component (Aigner et al., 1977). Non-parametric frontier methods make no assumptions about the mathematical formula of the frontier, and instead fit the frontier to the measured data. In this thesis, when I use the term “frontier methods”, I refer to non-parametric frontier methods.

In 1976, Charnes, Cooper, and Rhodes analyzed different functional forms of the frontier (e.g., linear versus curved as illustrated in Figure 4), and extended the work of Farrell. They coined the term Data Envelopment Analysis (DEA). Simar and Wilson (2000) estimated the uncertainty in efficiency estimations based on the uncertainty in input measurements. However, Simar and Wilson's method did not consider the case of repeated measurements.

DEA has been applied to energy efficiency ranking too. For example, Huntington used DEA to assess the productivity of US firms using energy and capital as inputs and products, goods, and services as outputs (Huntington, 1994). Carnes used DEA to estimate energy efficiency of Texas State universities based on course credits and research expenditure in each building (Carnes et al., 1998). Onut and Soner (2000) estimated the efficiency of five-star hotels in Antalia, Turkey, using occupancy, revenue, and number of guests as outputs and energy, water, and number of employees as inputs. DEA has also been applied to administrative buildings in Taiwan (Lee & Lee, 2008) and residential buildings in the US (Grosche, 2009).

3.2.3 A Comparison of Frontier Methods with Other Empirical Benchmarking Methods

Frontier methods are based on the basic definition of energy efficiency, which is "doing more using less" (Gillingham et al., 2009). Normally, energy managers and energy efficiency planners are not interested in changing the services in buildings (e.g., removing a dish washer without replacing it with a new one). Instead, they are interested in keeping the LOS almost the same, while minimizing the amount of energy. Therefore, we are interested in managing the tradeoff between LOS and energy consumption. Frontier methods explicitly illustrate this tradeoff.

Frontier methods have several advantages over existing energy efficiency benchmarking methods including:

- Frontier models compare a building to the best practice rather than average practice. It is often more actionable to learn from the best rather than to compare a building against mediocre performance.
- Frontier methods use an input-output paradigm in suggesting actions to improve efficiency. For example, the efficiency of a building can be improved both by reducing energy consumption for a given LOS and increasing LOS (e.g., accommodating more occupants) for the same amount of consumption.
- Since frontier methods are non-parametric, they relax some of the strong assumptions that Index Models make about the relationship of factors such as floor area with energy consumption.

Benchmarking methods are commonly compared based on two metrics: complexity, which measures the data and expertise needed to run the benchmarking process; and functionality. Complexity is a measure of scalability of the method. For example, collecting information for modeling individual buildings in great detail is a highly time and skill intensive task; hence, simulation models are not very scalable beyond comparing a few to a dozen buildings. On the other hand, Index Models only require aggregate consumption and a few important features such as floor area and number of occupants to create a metric to compare buildings. Since collecting such information about thousands of buildings is not highly time or skill intensive, Index Models are less complex and more scalable compared to simulation models. However, the ability of Index Models to make inferences about the causes of inefficiency or strategies to improve the efficiency of buildings is limited. Functionality is a measure of the power of the method to make inferences and actionable insights for future energy efficiency planning. For instance, simulation models can be used to generate different scenarios to examine hypothetical interventions on buildings, and therefore are more functional compared to Index Models. In the middle of the spectrum of complexity-functionality lie the engineering models. Engineering models are often parametric statistical models that model the energy efficiency of buildings using external factors (e.g., climate) and internal factors (e.g., occupancy patterns). Energy Star is an example of a statistical model. Since these models are based on parameters and statistical models, they still require making assumptions about the building. Frontier models often use similar parameters as parametric statistical models; however, they compare buildings non-parametrically. Thus, frontier models are relatively simpler compared to statistical parametric models. However, since frontier methods illustrate the tradeoff between energy and services, they can be used to infer ways to save energy by leveraging this tradeoff (Figure 5).

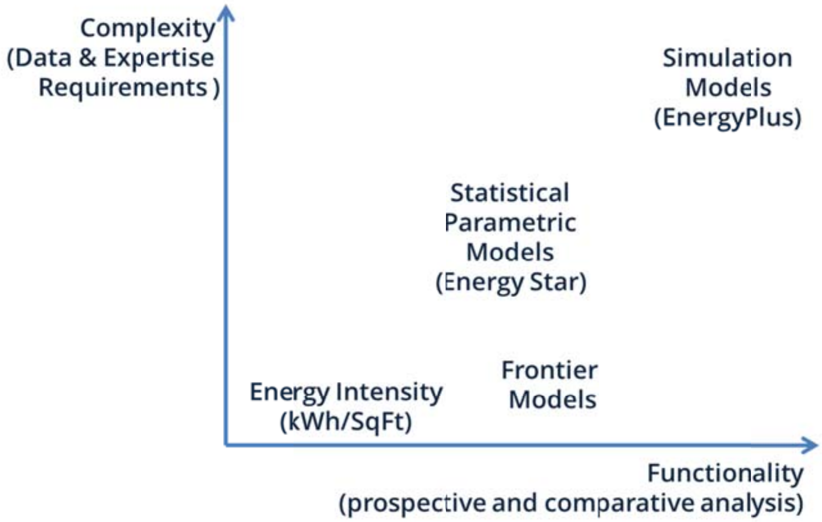


Figure 5 - A comparison of the most common benchmarking methods along complexity and functionality metrics. Each class of models is represented based on their simplest form (e.g., simplest frontier models can be as simple as using a single metric to quantify the output, but they can also be very complex as more variables are added to the model).

4 Contribution: Stochastic Energy Efficiency Frontier (SEEF)

Previous implementations of frontier methods in energy efficiency either treat efficiency scores as deterministic values, or estimate their uncertainty by resampling from one set of measurements. Availability of smart meter data (repeated measurements of energy consumption of buildings) enables utilizing measured data to estimate the uncertainty of efficiency scores. Furthermore, existing applications assume a linear form for the efficient frontier; i.e., they assume that the best-practice technology scales up and down proportionally with building characteristics. However, previous research shows that buildings are non-linear systems (for examples from the literature, see Kavousian et al., 2013). Chapter 3 explains these shortcomings in more detail.

I propose a statistical method called Stochastic Energy Efficiency Frontier (SEEF) to estimate bias-corrected efficiency scores and their confidence intervals from measured data. SEEF offers several advantages over existing applications of frontier methods in the energy efficiency context. Most importantly, SEEF treats energy efficiency as a random variable and finds bias-corrected efficiency estimates and their respective confidence intervals based on repeated measurements. Confidence intervals of efficiency ranks can be used to compare the efficiency of different buildings.

5 Structure and the Audience of the Thesis

Chapter 2 of the thesis examines driving factors of energy consumption in buildings, based on the results of my study of smart meter data for 1600+ homes in the U.S. in 2010. I examine the role of climate, building systems, appliances, and occupant behavior in driving energy demand using real-world data, and compare my findings with results from previous studies. By the end of Chapter 2, readers will gain an appreciation of the power and limitations of data-driven energy management methods, and the importance of accounting for the effect of internal and external factors when ranking building energy efficiency. Furthermore, Chapter 2 establishes data requirements, analysis options, and potential insights from statistical analyses of building consumption data. This work has been peer-reviewed and published in the Energy journal in 2013.

Chapter 3 incorporates the core contribution of my thesis, which is the algorithm behind the Stochastic Energy Efficiency Frontier (SEEF). I start the chapter with a critical review of existing benchmarking methods and in particular the application of frontier methods in building energy efficiency. Next, I explain the SEEF algorithm to calculate the confidence interval for efficiency estimates. Finally, I present the results of applying SEEF to a real-world case study. Through the case study, I validate the power of SEEF, by demonstrating that SEEF is able to better profile buildings based on the adoption of energy efficient features. While SEEF is introduced in the context of residential appliance usage efficiency, it can also analyze other building types (e.g., commercial buildings) and other aspects of energy efficiency (e.g., total building energy efficiency, heating and cooling efficiency, etc.). This chapter was published in January 2014 issue of ASCE Journal of Computing in Civil Engineering.

Chapters 4 and 5 serve to validate the power and generality of SEEF, by applying it to two different data sets of building energy consumption, one in the residential sector and the other in a commercial setting. The first case study (Chapter 4) presents the results of applying SEEF on a data set of 3600 residential customers' electricity consumption (30-min interval) from July 20th, 2009 through December 26th, 2010 for a total of 525 days. In addition to consumption data, the data set includes a detailed list of dwelling characteristics, demographic, and socioeconomic status (SES) data about households. The household data describe (a) building structural variables, such as building type, insulation level, vintage, (b) buildings systems variables including type of water heating, cooking fuel, percent efficient light bulbs, (c) appliances ownership, for instance number of TV sets, refrigerators, (d) demographic and socioeconomic status (SES), such as number of people, age, and (e) behavioral features, including motivation for energy saving and energy efficient habits (or lack of them). Due to the diversity of behavioral and demographic factors among the households of the data set, I was able to examine specific socioeconomic factors that contribute to energy efficiency. For example, I demonstrated that inherent demographic factors such as employment and home ownership have considerably larger influence on energy efficiency compared to self-assessed behavioral variables such as the willingness to make lifestyle changes to save energy. In Chapter 4, I also explain removing the effects of external and internal factors that influence energy consumption (e.g., weather and lifestyle). Therefore, this chapter validates that SEEF is both more unbiased and more actionable than other ranking methods.

The second case study (Chapter 5) analyzes energy consumption and operational data for 120+ buildings on the campus of a technology firm in Washington, US, spanning from July 2008 to July 2012. The data set includes monthly energy consumption (electricity and gas) of individual buildings, as well as their headcount, number of computers, data center load, HVAC system type, and a breakdown of floor area (office space, data center space, café space, and warehouse). The data set also includes energy management data including dates and descriptions of Energy Conservation Measures (ECMs), retrofits, and Retro-commissioning (RCx) projects. In Chapter 5, I explain how SEEF results can provide actionable feedback to building energy managers, by examining the effects of several different energy efficiency interventions on ranking scores. For instance, I demonstrate that low-cost interventions such as changing HVAC setpoint can be more effective in long-term improvement of efficiency score compared to lighting projects.

Finally, Chapter 6 summarizes the research questions and contributions of this thesis, and suggests future research that can build on and extend its findings and offerings.

The audience of this thesis include energy managers, utility energy efficiency planners, energy users, and researchers with an interest in data-driven energy efficiency.

6 References

Battles, S. (1997) Developing Energy Indicators in the U.S.: The EIA Experience. Energy Information Administration, In Proceedings of 18th North American Conference of US Association for Energy Economics., United States Association for Energy Economics, Cleveland, Ohio, USA.

Carnes, D., Hunn, B., Jones, J. (1998) The Use of Data Envelopment Analysis for Evaluating Building Energy Consumption in Terms of Productivity. Proceedings of 1998 Summer Study in Building Energy Efficiency, Pacific Grove, CA. American Council for an Energy Efficiency Economy, Washington, DC, USA.

Charnes, A., Cooper, W., Rhodes, E. (1976) Measuring the Efficiency of Decision Making Units. European Journal of Operations Research. 2 (1978) 429-444.

Deutsche Bank and Living Cities. 2011. Recognizing the Benefits of Energy Efficiency in Multi-Family Underwriting. New York.

EPA (2012) Energy Star Guidelines for Energy Management Overview. http://www.energystar.gov/index.cfm?c=guidelines.guidelines_index

EPA (Environmental Protection Agency). (2011). ENERGY STAR Performance Ratings Technical Methodology for Office, Bank/Financial Institution, and Courthouse. Environmental Protection Agency. Retrieved from www.energystar.gov/ia/business/evaluate_performance/office_tech_desc.pdf

EPA (Environmental Protection Agency). (2011). ENERGY STAR Performance Ratings Technical Methodology, General Overview. Environmental Protection Agency. Retrieved from http://www.energystar.gov/ia/business/evaluate_performance/General_Overview_tech_methodology.pdf

Farese, P., Gelman, R., Hendron, R. (2012) A Tool to Prioritize Energy Efficiency Investments. National Renewable Energy Lab (NREL) Technical Report NREL/TP-6A20-54799, August 2012, Golden, CO, USA.

Farrell, M. J. (1957). The Measurement of Productive Efficiency. Journal of the Royal Statistical Society, 120(3), 253-290.

Gillingham, K., Newell, R., Palmer, K. (2009) Energy Efficiency Economics and Policy. Annual Review of Resource Economics, Annual Reviews, vol. 1(1), pages 597-620, 09.

Goldstein, David (2012) Personal communication.

Granderson, J., M.A. Piette, B. Rosenblum, L. Hu (2011). "Energy Information Handbook: Applications for Energy-Efficient Building Operations". Lawrence Berkeley National Laboratory, Berkeley, Californian. LBNL-5272E

Grosche, P. (2009) Measuring Residential Energy Efficiency Improvements with DEA. Productivity Analysis, 31, 87-94.

Hutington, H. (1994) Been Top Down So Long It Looks Like Bottom Up to Me. Energy Policy, 22 (10), 833-839.

Kavousian, A., Rajagopal, R., Fischer, M. (2013) Determinants of residential electricity consumption: Using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants' behavior. Energy, Volume 55, 15 June 2013, Pages 184-194, ISSN 0360-5442, http://dx.doi.org/10.1016/j.energy.2013.03.086.

Kavousian, A. and Rajagopal, R. (2014). "Data-Driven Benchmarking of Building Energy Efficiency Utilizing Statistical Frontier Models." J. Comput. Civ. Eng., 10.1061/(ASCE)CP.1943-5487.0000327.

Lovins, A. (2004) "Energy Efficiency, Taxonomic Overview," Encyclopedia of Energy 2:383-40, 6 vols.

Lutzenhiser, L., Moezzi, M., Hungerford, D., Friedmann, R. (2010). Sticky points in modeling household energy consumption. ACEEE Summer Study on Energy Efficiency in Buildings, Pacific Grove, CA. American Council for an Energy Efficiency Economy, Washington, DC, USA.

Patterson, M. (1996) What is Energy Efficiency: Concepts, Indicators, and Methodological Issues. Energy Policy, 24 (5), 377-390.

Peterman, A., Kourula, A., Levitt, R. (2012) A roadmap for navigating voluntary and mandated programs for building energy efficiency. Energy Policy, 43 (1) 415-426

Schipper, L. (1983) Energy Efficiency and Conservation in Buildings: The Use of Indicators. Energy 8(1), 7-14.

Shapiro, I. (2011) Ten Most Common Problems with Energy Audits. ASHRAE Journal, Debruary 2011, 26-3.

Chapter 2

Determinants of residential electricity consumption: Using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants' behavior[2]

1 Abstract

We propose a method to examine structural and behavioral determinants of residential electricity consumption, by developing separate models for daily maximum (peak) and minimum (idle) consumption. We apply our method on a data set of 1628 households' electricity consumption. The results show that weather, location and floor area are among the most important determinants of residential electricity consumption. In addition to these variables, number of refrigerators and entertainment devices (e.g., VCRs) are among the most important determinants of daily minimum consumption, while number of occupants and high-consumption appliances such as electric water heaters are the most significant determinants of daily maximum consumption. Installing double-pane windows and energy-efficient lights helped to reduce consumption, as did the energy-conscious use of electric heater. Acknowledging climate change as a motivation to save energy showed correlation with lower electricity consumption. Households with individuals over 55 or between 19 and 35 years old recorded lower electricity consumption, while pet owners showed higher

[2] This chapter was first published in Energy Journal as:
Kavousian, A., Rajagopal, R., Fischer, M. (2013) Determinants of residential electricity consumption: Using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants' behavior. Energy, Volume 55, 15 June 2013, Pages 184-194, ISSN 0360-5442, http://dx.doi.org/10.1016/j.energy.2013.03.086.

consumption. Contrary to some previous studies, we observed no significant correlation between electricity consumption and income level, home ownership, or building age. Some otherwise energy-efficient features such as energy-efficient appliances, programmable thermostats, and insulation were correlated with slight increase in electricity consumption.

2 Introduction

Residential buildings consume 39% of the total electricity in the US, more than any other sector or building type [1]. Their electricity consumption has increased by 27% from 1990 to 2008 and—provided the expected efficiency gains are realized—is projected to increase by 18% from 2009 to 2035. To meet this demand, 223 GW of new generating capacity will be needed between 2010 and 2035, 75% of which is projected to be provided by fossil fuels [2]. Detailed planning and execution of demand-side energy efficiency programs is needed to reduce or stabilize residential electricity consumption, and to prevent its harmful impact on the environment and on energy security [3] and [4].

To effectively plan and execute energy efficiency programs, a sound understanding of the determinants that drive household electricity consumption (such as floor area, average outside temperature, and number of occupants) is needed [5]. However, because of lack of easily-accessible, high-resolution consumption data, underlying determinants of energy use and energy-related behaviors have not been extensively examined before [6].

With growing deployment of smart meters and real-time home energy-monitoring services, data that allow studying such underlying determinants are becoming available (for examples of studies using high-resolution consumption data, see Refs. [7], [8] and [9]). However, the methodologies to analyze the data and infer the results that can be used to support decision making at the household level have not yet been formalized [6].

To address that gap, this paper proposes a methodology to analyze large data sets of residential electricity consumption to derive insights for policy making and energy efficiency programming. In particular, it offers a method to disaggregate the impact of structural determinants (e.g., insulation level of the residence) from behavioral determinants (e.g., occupant habits). As a case study, we use a large data set of 10-min interval smart meter data for 1628 households in the U.S. The data set is collected over 238 days in 2010, and is supported by an extensive 114-question survey of household data. The household survey covers information about the climate, location, physical characteristics of the dwelling, appliances, and occupants.

After a brief review of the gaps in existing literature, this paper explains our model, experimental data set, and the results of the our analyses.

3 Summary of Limitations of Existing Models

Several studies have analyzed residential electricity consumption and its determinants. For our purposes, we needed a bottom-up model that could make use of high-resolution electricity consumption data and a large set of information about the households. We found that most existing studies were limited in supporting our requirements, mostly because of:

(a) Use of aggregate (low-resolution) consumption data: Most studies in the past have used monthly billing data, mainly because the advanced metering technologies of today were not easily accessible [10], [11], [12], [13], [14], [15], [16], [17] and [18]. However, residential electricity consumption has strong temporal variation, which is not captured with low-resolution consumption data such as monthly bills [9].

(b) Partial set of explanatory variables: Most previous studies have analyzed only a partial set of residential electricity consumption determinants; e.g., only appliance stock, weather conditions, or behavioral factors [19] and [20]. However, interactions among different factors (e.g., the relationship between weather, appliance load, lighting load, and heating load) offer considerable potential for improving energy efficiency [6]. Another limitation of some previous studies is the use of “bundle” variables (such as Zip Code) that combine (hence obscure) the effect of several underlying determinants.

(c) No distinction between “idle” consumption of the house and peak consumption: Most studies in the past have either used peak consumption or the total electricity load. However, analyzing the lower limit of electricity consumption (i.e., idle load) gives valuable insights on the physical characteristics of the building. For example, a leaky building may have higher idle load because the heater or AC (air conditioner) needs to run constantly. Furthermore, as we will show in the paper, the distinction between idle and maximum consumption helps to disaggregate the effect of structural and behavioral factors on residential electricity consumption.

(d) Using energy intensity as the only indicator for analyzing electricity consumption: Most studies have used energy intensity (kWh per square foot) to compare residential electricity consumption [21], [22], [23], [24], [25], [26] and [27]. Scaling the electricity consumption by floor area implies that, for example, a refrigerator in a 2000 sq.ft house consumes twice as much as the same refrigerator in a 1000 sq.ft house. Instead, we scale only those factors whose consumption is dependent on floor area (e.g., lighting and heating loads), and use the unscaled kWh value for other factors.

4 Model Setup

Our proposed model addresses the limitations of existing models by: (a) classifying explanatory variables based on their interaction with each other and with electricity consumption; (b) selecting a subset of variables that best explain electricity consumption; (c) using different load features (e.g., daily peak and idle load) to distinguish the effect of different variables on electricity consumption; and (d) using a stepwise method to rank explanatory variables based on their correlation with electricity consumption.

4.1 Explanatory Variables

Through a review of residential electricity consumption models and building sciences literature [5], we identified four major categories of determinants:

1. Weather and location. Examples: daily outdoor temperature and climate zone; these determinants are normally outside the scope of influence of the household.

2. Physical characteristics of the building. Examples: level of insulation and fuel use for water heating; modifying these determinants is normally considered long-term investments [28]. In return, the effect of such modifications is likely to persist longer.

3. Appliance and electronics stock. Examples: the number of refrigerators or computers; modifying these determinants is normally considered medium to short-term investments.

4. Occupancy and occupants' behavior towards energy consumption: determinants in this category have different levels of effort and impact span. Some behavioral modification determinants are of short-term effort and impact (e.g., proper management of thermostat settings) [28]. Another group of determinants are associated with long-term effort and impact (e.g., purchasing energy-efficient appliances). Finally, some determinants in this category are outside the scope of interest to change (e.g., occupancy level during the day).

Behavioral factors show strong correlation with each other, making multicollinearity an issue for the power and stability of the regression model. We used FA (Factor Analysis) [29] to remove multicollinearity of the variables, and identified latent variables that are not captured by direct behavioral questions. Factor Analysis reduces the number of variables by replacing them by potentially lower number of linear combinations of original variables (called factors) while conserving as much information as possible. Because these factors are linear combinations of the original variables, they are easy to interpret and can be assigned physical significance. In short, Factor Analysis identifies the set of k latent factors (f1,f2,…,fk) that drive q observable variables (responses to survey questions) indexed as xt=(x1,x2,…,xq), where ([30]):

$$x = \Lambda f + u$$

where Λ is the q × k matrix of factor loadings, and f and u are q × l matrices of factors and variances.

While estimating Λ by the Maximum Likelihood method, Factor Analysis identifies a rotation and a scale of Λ that contracts as many coefficients to zero as possible. Having a sparse matrix Λ increases the interpretability of the factor model, since any factor will be created using only a few observable variables, hence allowing us to bundle several inter-related variables and label them as a factor.

4.2 Model Selection

When working with a large number of explanatory variables, even after Factor Analysis, the number of model variables may be too large to support a model that is easy to interpret and statistically stable. Furthermore, it is important to identify variables that contribute the most to the variation in consumption to inform future data collection efforts. Our preferred method for model selection is stepwise selection[31] because (a) it ranks the variables based on their importance; and (b) in sequentially adding variables to the model, it minimizes multicollinearity.

The forward stepwise algorithm starts with the mean value of the consumption (i.e., the intercept) and then sequentially adds to the model the variable that best improves the fit, as measured by the AIC (Akaike Information Criteria [32]), given by Equation (2):

$$AIC = -2.\log L + 2.edf$$

where L is the likelihood and edf is the equivalent degrees of freedom (i.e., the number of free parameters for usual parametric models) of fit. The backward stepwise algorithm starts with the full model, then sequentially drops the variables that least hurt the fit, as measured by AIC. Forward and backward stepwise procedures usually result in similar models [31].

4.3 Response Variables

We considered four different features of the hourly electricity consumption data: daily average, minimum, maximum, and maximum-minus-minimum (also called "range"). Daily minimum, maximum, and average values for each user were calculated for each day of the experiment (238 values for each user). Then, these 238 values for each user were averaged over the summer and winter (partial) periods, resulting in two daily minimum, two daily maximum, two daily average, and two daily maximum-minus-minimum consumption values for each user (eight values in total for each user). Then separate regression models were fitted to each load feature for each season (eight regression models). Each regression model consists of 952 response variables (one value for each user).

We used a weighted regression model to explain the variation in household electricity consumption. Those determinants whose contribution to electricity consumption has a linear relationship with floor area (e.g., insulation level) are multiplied by the floor area of the residence, while other variables (e.g., number of refrigerators) are included without a multiplier. The regression equation of our model is given by:

$$y_j \beta_{0j} = \sum_{i=1}^{M} \beta_{ij} X_{ij} + X_j \times \sum_{i=M+1}^{K} \beta_{ij} X_{ij} + \varepsilon_j$$

where y_j is the electricity consumption (kWh) of household j, X_{ij} is the value of the determinant i for household j, and β_{ij} is the regression coefficient for that determinant. M is the number of variables (household features) that do not depend on floor area, while K is the total number of variables, and ε is the error term.

To summarize, our method enables working with large data sets of electricity consumption data and extensive household surveys, by (a) using different load features that help to understand different aspects of consumption (e.g., long-term steady idle load versus short-term volatile peak load); (b) selecting a subset of variables that contribute the most to each load feature; (c) ranking the contribution of different variables to each load feature through a stepwise model; and, (d) properly considering the effect of floor area, by using a weighted regression model.

4.4 Data Summary and Preprocessing

We applied our model to a data set of 10-min interval smart meter data for 1628 households, collected over 238 days from February 28, 2010 to October 23, 2010. Detailed data about household characteristics were available via a 114-questions online survey, covering a wide range of characteristics including climate and location, building characteristics, appliances and electronics stock, demographics, and occupants' behavior.

Participants were provided with a device that recorded the electricity consumption of the household for every 10 min and sent the data to a central server to be stored. The device installation and server costs were covered by the experiment administrators (for more details of the experiment, refer to Ref. [33]).

Participants were employees of a technology company in Silicon Valley and voluntarily participated in the experiment. They all paid for their own electricity bill. More than 50% of the participants reported income levels higher than $150,000. However, it is worth mentioning that the mix of households in our study (i.e., well-educated, upper and middle class families who are also early adopters of new technologies such as home energy monitoring systems) is also more likely to respond to energy efficiency programs by investing in energy-efficient products [34]. Hence, the results of our analysis can be particularly helpful to energy efficiency program managers and policy makers to develop programs specifically targeted towards the households represented by our sample.

After collecting the data, 952 households for which reliable smart meter and survey data were available were selected for the analysis. Less than 3% of survey responses were inconsistent or missing, for which we imputed data using iterative model-based imputation techniques ([35] and [36]). After the initial screening, we did not observe any outliers in the data set. Selected households were located in 419 different Zip

Codes, 140 different counties, 26 different states, and were spread across all six climate zones defined by the Department of Energy [1]. California had the largest representation (53% of households) of all states in the data set. During the data collection process, the weather conditions in most areas where participant households resided were similar to the 30-year average climatic conditions; however, some areas, especially in the north east of the U.S., experienced slightly higher-than-normal temperatures [37]. Average electricity consumption in our sample lied between California and US averages. Some structural determinants such as household size, square footage of the house, and the proportion of single family detached units in our sample were close to US population averages [33].

We analyzed the consumption data at hourly level to ensure that the fluctuations in electricity consumption are considered, but not obscured by sudden spikes in the consumption. This also makes the results of our analysis comparable with those of previous studies on smart meter data and electricity market analysis [38].

We transformed some variables to better reflect the technical characteristics of buildings. For example, we transformed the construction year to a categorical variable that indicated the residential building code that was effective at the time of the construction (i.e., different revisions of ASHRAE 90.2 [1]). We also included a categorical variable for House Size to capture the effects of the floor area that are not completely explained by square footage. For example, when a building's floor area passes a certain threshold, the type of structural and architectural material that is used in the building often changes significantly. Since we do not have a separate variable for floor area and are not dividing the electricity consumption of the dwelling by its floor area, introducing the House Size variable does not create a multicollinearity problem (Table 1).

Table 2 - Summary statistics of the daily maximum, minimum, and average hourly consumption, averaged overall users in the case study. The variability in daily minimum hourly consumption is the lowest, and that of the daily maximum is the largest.

Variable	No. of days	Min	Median	Mean	Max	Standard Dev.
Daily minimum, kWha	238	0.2	0.4	0.4	0.5	0.039
Daily average, kWha	238	0.5	0.8	0.9	1.2	0.1
Daily maximum, kWha	238	1.7	2.5	2.5	3.5	0.256

The household survey included 40 questions on the attitudes of occupants towards energy consumption. Using Factor Analysis as was explained in previous sections, and informed by behavioral sciences research, we formed 22 major factors that collectively explain more than 80% of the information included in the original 40 questions; i.e., the threshold value in the scree test for choosing the number of factors was explaining 80% of the variance. The 22 variables explain the attitudes of households in three groups: (1) Energy Efficiency Actions, (2) Information Seeking Behaviors, and (3) Home Improvements Behaviors. After Factor Analysis, adding a number of transformations of the original variables, and removing some variables, the total number of variables was reduced from 114 to 97. We fit separate models to daily maximum, minimum, maximum minus minimum, and average consumption, both for summer and winter (for the period when the data were available), and ranked the variables by their importance through a forward stepwise model selection procedure (Fig. 1).

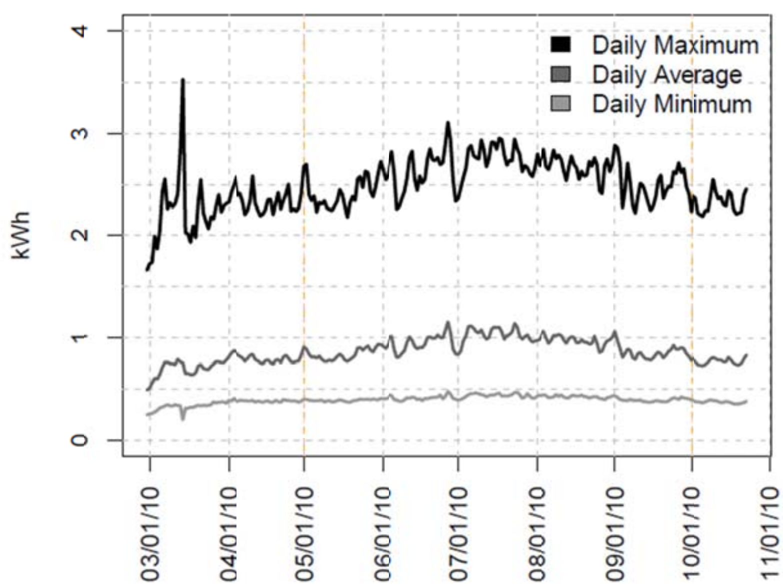


Figure 6 - Comparison of daily average, maximum, and minimum consumption, averaged over all users for each day of the experiment.

5 Results

A comparison of the results of our models shows that the daily minimum consumption is influenced the most by weather, location, and physical characteristics of the building. On the other hand, daily maximum consumption is influenced the most by end uses that are energy-intensive and do not run constantly (e.g., electric water heater). This group of end uses mostly depends on the occupancy levels and activities of occupants. These results are summarized in Tables 2 and 3 (forward stepwise results), and Table 4 (backward stepwise results).

Table 3 - Summary of the most important factors explaining different features of residential electricity consumption (forward stepwise regression). F: full model; P: partial model (excluding Zip Code and Floor Area). Only percentages greater than 2% are included in this table.

Variable	Min				Max				Max-Min				Average			
	Summer		Winter		Summer		Winter		Summer		Winter		Summer		Winter	
	F	P	F	P	F	P	F	P	F	P	F	P	F	P	F	P
Ave. of CDD		26%				31%				27%				38%		
Climate Zone						2%								3%		
Zip Code	12%		12%		39%		26%		37%		25%		46%		17%	
House Size	2%		21%		11%		2%		9%				12%		23%	
Type of bldg												12%		2%		
Ownership of elec. water heater						4%	4%	11%	2%	6%	5%	12%			2%	5%
Ownership of elec. clothes dryer						2%	2%	3%		3%	2%	4%				
Nb of spas/pools							2%				2%				2%	
Nb of freezers						3%								2%		
Nb of refrig's	7%	7%	7%	7%	3%	4%	4%	3%	2%	2%	3%	2%	3%	4%	6%	6%
Nb of entert't devices Except TV's		3%	2%	4%										2%		
Total nb of occup'ts							8%									
Total nb of occup'ts (sq. rt)					2%	2%	2%	4%	2%	3%	7%	4%			2%	2%
Pet ownership	2%	2%					2%	4%				3%			2%	3%
Purchasing E-Star Appl's	2%	2%	2%	3%												2%
Energy Conserv'n w.r.t. Elec. Heater Usage		2%	2%	2%										2%		
Turning lights off when not in use				19%				13%								20%
Motivated to reduce coumspt'n to address Global Warm.				2%											2%	3%

Summer model: total number of variables: 93; $R^2_{adj}=0.52$
Winter model: total number of variables: 82; $R^2_{adj}=0.48$

Table 4 - Model coefficients for statistically significant variables of forward stepwise regression models. The mean and standard deviation of each load feature are shown in parentheses at the table header. Shaded cells highlight negative values.

Mode	Min				Max				Max - Min				Average			
	Summer (0.404; 0.354)		Winter (0.358; 0.297)		Summer (2.616; 1.596)		Winter (2.307; 1.395)		Summer (2.212; 1.374)		Winter (1.948; 1.230)		Summer (0.925; 0.739)		Winter (0.769; 0.537)	
Variable	Full	Partial	Full	Partial	Full	Partial	Full	Partial	Full	Partial	Full	Partial	Full	Partial	Full	Partial
CDD		0.01	0.004	0.004	0.03	0.054	0.016	0.018	0.024	0.043		0.013	0.012	0.03	0.007	0.009
HDD			-0.006	-0.005											-0.006	
Climate Zone														-0.15 to 0.17 (ave=-0.02)		
Zip Code	-.43 to 1.56 (ave=.19)		-.29 to 1.08 (ave=.07)		-1.27 to 3.64 (ave=.34)		-2.84 to 2.03 (ave=-1.23)		-1.32 to 3.12 (ave=.12)		-2.59 to1.45 (ave=-1.17)		-.66 to 2.37 (ave=.30)		-.74 to 1.44 (ave=-.08)	
House Size	-0.57 to -0.24 (ave=-0.38)		-0.20 to 0.46 (ave=-0.06)		-0.62 to 1.40 (ave=0.15)		-1.65 to 0.62 (ave=-0.94)		-0.52 to 0.81 (ave=0.15)		-1.48 to -0.02 (ave=-0.98)		-0.43 to 0.97 (ave=-0.03)		-.62 to .50 (ave=-.38)	
Type of Bldg		-0.06 to 0.05 (ave=0.01)				-0.31 to 0.25 (ave=0.07)		-.21 to .39 (ave=.17)		-0.27 to 0.27 (ave=0.09)		-.17 to .39 (ave=.19)		-0.09 to 0.14 (ave=0.05)		-0.06 to 0.11 (ave=0.05)
Bldg Shell Insulation								0.074			0.069	0.072	0.04	0.033		
Energy-Efficient Lights	0.025		-0.016					-0.132			-0.082	-0.129				-0.04
Double-Pane Windows	-0.048	-0.052		0.036			-0.134						-0.093	-0.087		-0.034
Programmable Thermostat					0.142	0.14			0.124	0.123						
No. of Elec. Heaters						0.092	0.095	0.138		0.09	0.092	0.127			0.032	0.047
No. of AC	0.011	0.014														
No. of Elec. Water Heaters				0.054	0.534	0.0682	0.804	1	0.532	0.656	0.783	0.946	0.003	0.183	0.197	0.268
No. of Elec. Clothes Dryers					0.294	0.369	0.382	0.459	0.298	0.355	0.392	0.457				0.062
No. of Elec. Stoves					0.238	0.312		0.18	0.227	0.292		0.166		0.094		
No. of Spas/Pools	0.096		0.096	0.081	0.341	0.281	0.445	0.406	0.279	0.241	0.369	0.328	0.23	0.138	0.258	0.197
No. of Dishwashers					0.266	0.398			0.301	0.391						
No. of Refrigerators		0.121	0.104	0.105	0.356	0.348	0.331	0.335	0.242	0.218	0.237	0.222	0.188	0.195	0.177	0.174
No. of Freezers	0.016		0.047	0.057	0.23	0.28	0.249	0.303		0.2	0.222	0.246	0.149	0.16	0.131	0.138
No. of TV's	0.03	0.029	0.024	0.026	0.097	0.1	0.102	0.093	0.086		0.072		0.045	0.051	0.037	0.031
No. of Entertain't Devices (Except TV's)	0.011	0.022	0.016	0.02		0.059							0.031	0..042	0.019	0.023
No. of Occupants (sqrt)			0.066		0.959	0.771	0.902	0.772	0.883	0.57	0.867	0.758	0.31	0.137	0.273	0.198
Individuals under 5	0.012															
Individuals between 6 and 12		-0.013								0.049						
Individuals between 13 and 18		0.017												0.017		0.022
Individuals between 19 and 35					-0.096	-0.082	-0.047	-0.043	-0.084		-0.044	-0.036	-0.042	-0.015	-0.017	-0.011
Individuals between 36 and 54						-0.062			-0.061				-0.03			
Individuals between 55 and 65		-0.02	-0.021	-0.018	-0.15	-0.12	-0.098	-0.097	-0.128		-0.079	-0.082	-0.07	-0.045	-0.034	
Individuals Older than 65				-0.021	-0.195	-0.23	-0.156	-0.163	-0.178	-0.195	-0.14	-0.14	-0.071	-0.072	-0.058	-0.055
No. of Pets	0.037	0.039	0.023	0.027	0.126	0.128	0.157	0.166	0.085	0.073	0.135	0.139	0.06	0.059	0.055	0.057
Purchasing E-Star Appliances	0.011	0.014	0.015	0.016											0.019	0.021
Energy-Conscious Use of Elec. Heater (Thermostat Settings)	-0.023	-0.027	-0.017	-0.019	-0.082	-0.105		-0.042	-0.057	-0.08			-0.045	-0.064	-0.026	-0.027
Energy-Conscious Use of AC (Thermostat Settings)					0.036	0.035			0.032	0.032			0.01	0.013		
Turning Lights Off When Not in Use	0.064	0.022		0.022												0.046
Motivated to Reduce Consump'n to Address Global Warming	0.001		-0.012	-0.013		-0.028	-0.047	-0.049		-0.019	-0.034	-0.037	-0.015	-0.02	-0.025	-0.027

Table 5 - Model coefficients for statistically significant variables of backward stepwise regression models. Shaded cells highlight negative values.

Model	Min				Max				Max - Min				Average			
	Summer (0.404; 0.354)		Winter (0.358; 0.297)		Summer (2.616; 1.596)		Winter (2.307; 1.395)		Summer (2.212; 1.374)		Winter (1.948; 1.230)		Summer (0.925; 0.739)		Winter (0.769; 0.537)	
Variable	Full	Partial	Full	Partial	Full	Partial	Full	Partial	Full	Partial	Full	Partial	Full	Partial	Full	Partial
CDD	0.005	0.01	0.004	0.004		0.055	0.015	0.019	0.024	0.043		0.013	0.012	0.03	0.007	0.009
HDD			-0.006	-0.004											-0.006	
Climate Zone																
Zip Code	-0.41 to 1.57 (ave=0.19)		0.29 to 1.08 (ave=0.07)		-1.23 to 3.63 (ave=0.32)		-2.79 to 2.09 (ave=-1.17)		0.00 to 0.97 (ave=0.56)		-2.80 to 1.40 (ave=-1.22)		-0.65 to 2.35 (ave=0.27)		-0.75 to 1.42 (ave=-0.10)	
House Size	-0.35 to 0.03 (ave=-0.23)		-0.20 to 0.46 (ave=-0.06)		0.64 to 1.62 (ave=0.37)		-1.64 to 0.76 (ave=-0.91)		-0.53 to 0.79 (Ave=-0.16)				-0.43 to 1.01 (Ave=-0.01)		-0.62 to 0.52 (ave=-0.36)	
Type of Bldg						-0.31 to 0.28 (ave=0.08)		-0.22 to 0.41 (ave=0.18)		-0.23 to 0.27 (ave=0.09)		-0.17 to 0.39 (ave=0.19)		-0.09 to 0.13 (Ave=0.05)		-0.06 to 0.10 (ave=0.05)
Bldg Shell Insulation	0.025		-0.02		0.083		0.08				0.087	0.072	0.046	0.034	0.039	
Energy-Efficient Lights								-0.121			-0.1	-0.129				-0.041
Double-Pane Windows	-0.053	-0.059		-0.036			-0.128				-0.114		-0.108	-0.089		-0.036
Programmable Thermostat	0.024				0.105	0.141		0.095	0.127	0.134						
No. of Elec. Heaters						0.083		0.133	0.065	0.09	0.99	0.127		0.024		0.049
No. of AC	0.013	0.013														
No. of Elec. Water Heaters				0.054	0.523	0.677	0.802	0.992	0.523	0.662	0.783	0.946	0.13	0.186	0.194	0.269
No. of Elec. Clothes Dryers					0.301	0.371	0.388	0.465	0.306	0.367	0.409	0.457				0.061
No. of Elec. Stoves					0.22	0.315		0.184	0.223	0.286		0.166		0.093		
No. of Spas/Pools	0.092			0.081	0.335	0.291		0.416	0.28	0.221	0.342	0.328	0.237	0.135		0.196
No. of Dishwashers					0.304	0.396		0.232	0.341	0.475		0.186				
No. of Refrigerators	0.118	0.121	0.104	0.105	0.349	0.354	0.325	0.345	0.249	0.232	0.218	0.222	0.188	0.195	0.175	0.174
No. of Freezers	0.065	0.072	0.047	0.057		0.279	0.25	0.312		0.208	0.211	0.246	0.149	0.162	0.125	0.137
No. of TV's	0.029	0.035	0.024	0.026	0.094	0.097	0.096	0.09	0.085	0.087	0.074		0.045	0.052	0.036	
No. of Entertain't Devices (Except TV's)	0.017	0.022		0.02		0.059							0.03	0.043	0.019	0.024
No. of Occupants (sqrt)	0.103	0.081	0.066	0.035	0.963	0.784	0.919	0.789	0.885	-0.858	1.062	0.758	0.311	0.151	0.27	0.198
No. of Individuals under 5											-0.036					
Individuals between 6 and 12	-0.02	-0.022														
Individuals between 13 and 18				0.014												
Individuals between 19 and 35	-0.021	-0.02			-0.098	-0.084	-0.048	-0.04	-0.084	-0.08	-0.076	-0.036	-0.045	-0.016	-0.017	-0.012
Individuals between 36 and 54	-0.015	-0.019			-0.071	-0.062			-0.061	-0.071	-0.052		-0.033			
Individuals between 55 and 65	-0.03	-0.03	-0.021	-0.018	-0.155	-0.117	-0.096	-0.094	-0.129	-0.121	-0.112	-0.082	-0.072	-0.046	-0.035	
Individuals Older than 65	-0.025	-0.026		-0.021	-0.194	-0.229	-0.155	-0.159	-0.178	-0.223	-0.154	-0.14	-0.071	-0.072	-0.057	-0.054
No. of Pets	0.038	0.042	0.023	0.027	0.119	0.125	0.155	0.166	0.085	0.084	0.124	139	0.062	0.061	0.054	0.056
Purchasing E-Star Appliances		0.104	0.015	0.016		-0.075			-0.059						0.018	0.02
Energy-Conscious Use of Elec. Heater (Thermostat Settings)	-0.023	-0.025	-0.017	-0.019	-0.085	-0.104		-0.042		-0.076				-0.065	-0.026	-0.029
Energy-Conscious Use of AC (Thermostat Settings)					0.038	0.034			0.034	0.033			0.011	0.012		
Turning Lights Off When Not in Use		0.044											-0.045			0.041
Motivated to Reduce Consump'n to Address Global Warming	-0.009	-0.011	-0.012	-0.013	-0.024	-0.028	-0.047	-0.045		-0.021	-0.032	-0.037	-0.016	-0.019	-0.025	-0.027

Overall, locality (indicated by a proxy such as Zip Code) and House Size demonstrate considerable correlation with residential electricity consumption, most likely because they are correlated with several other variables that influence energy consumption. For example, Zip Code is often correlated with climate, building type, type of systems used in the building, building materials, and socioeconomic status of the household. On the other hand, House Size is often correlated with affluence, socioeconomic status, number of residents, and appliance stock.

5.1 The Effect of External Determinants on Residential Electricity Consumption

Zip Code explains up to 46% of the variability in consumption. However, once Zip Code is removed from the models, underlying drivers of electricity consumption such as Cooling Degree Days are highlighted.

CDD (Cooling Degree Day) is the dominant factor in the summer, explaining 38% of the variability in total electricity consumption. On the other hand, HDD (Heating Degree Day) is not a significant factor, even in the winter model. We offer an explanation for this observation in Section 5.2. Other studies have also reported similar results on the effect of Zip Code and cooling load on electricity consumption [39], [40], [41] and [42].

5.2 The Effect of Physical Characteristics of Building

House Size is the most important factor among building characteristics in our models, while house age, type of building, and ownership status do not show significant impact on electricity consumption in our sample. Other variables such as installing double-pane windows and energy-efficient lighting fixtures show correlation with reduced electricity use. The following sections explain these results in more detail.

5.2.1 Type of Building

Type of Building is a significant factor in the winter daily maximum model, where heating load dominates. In the winter, households who live in multifamily apartments have the lowest daily maximum consumption, followed by town houses; finally, detached (free-standing) houses have the highest daily maximum consumption in the winter. Similar results have been reported previously ([5] and [43]).

5.2.2 House Size

House Size has a dual effect on daily minimum and daily maximum consumption. Daily minimum consumption is much more correlated with House Size during the winter (House Size explains 21% of the variability in winter minimum consumption, while it explains only 2% of the variability in summer minimum). On the other hand, House Size has a much larger impact on daily maximum consumption during the summer compared with winter. Both observations can be explained by the relationship of House Size with heating and cooling loads, and the nature of these two loads. First, larger houses require more heating and cooling loads, both because they have more volume to condition, and because they have higher heat loss or gain with the outside. Also, heating load tends to have a more consistent nature, especially when systems such as radiant heating are used. Cooling load, on the other hand, tends to have a more intermittent nature, since the air conditioner compressor only needs to run intermittently to produce cool air.

5.2.3 House Age

We did not observe any significant correlation between electricity consumption and building age. While some previous studies have observed an increase in household electricity consumption of new houses due to more penetration of air conditioning and other high-consumption appliances [44], other studies have observed the reverse, reporting a decrease in household electricity consumption for newer houses, attributing that pattern to improved insulation and use of more efficient lighting and air conditioning stock [45] and [46]. Our hypothesis is that in our sample, these two forces have canceled out each other's effect, resulting in a uniform trend between household electricity consumption and the age of the house. Another possible explanation for the uniform trend in our data is that the physical conditions of buildings have been maintained through time, possibly due to the enforcement of building regulations.

When we grouped the households into different time periods based on the prevalence of different ASHRAE 90.2 residential building codes, we observed that the houses that were built before 1975 on average consumed less electricity than the houses that were built between 1993 and 2003 (p-values = 0.00266 and 0.00105, respectively). A potential explanation for this trend is the increased penetration of air conditioners and other high-consumption appliances in newer houses.

5.2.4 The Effect of Appliance Stock and Electronics

As Tables 3 and 4 show, number of refrigerators is a statistically significant factor in almost all models, but its effect is more highlighted in the daily minimum consumption. Former studies show that (a) refrigerators are the largest electricity consumer among household appliances (according to Ref. [44] in 2001, refrigerators consumed 14% of total electricity delivered to U.S. homes, only second to air conditioners who consumed 15%), and (b) the secondary refrigerators in the US households are on average considerably older (and less-efficient) than the primary refrigerators [44]. Therefore, the ownership of more than one refrigerator in a household implies a high probability of having an inefficient, high-consumption fixture, hence the significant contribution of the number of refrigerators to household electricity consumption.

Other than refrigerators that have a steady load, most high-consumption, intermittent appliances such as electric water heater, electric clothes dryer, and Spas/Pools primarily contribute to daily maximum consumption. These are the appliances that are not “always on” and their operating schedules are dependent on the activities and habits of the occupants. Therefore, they are indicators of (and are driven by) occupants' habits and activities rather than the location and physical characteristics of the dwelling, hence their correlation with daily maximum load. Our results are in line with former studies who have shown the large impact of high-consumption appliances on total electricity consumption. For example, according to the U.S. Energy Information Administration, air conditioners, electric water heaters, and laundry appliances consume 16.0%, 9 1%, and 6.7% of the total electricity consumption in US households, respectively [44]. Similarly in Europe, according to EuroAce, 57% of the energy consumed in buildings is used for space heating, 25% for hot water, 11% for lighting and electrical appliances, and 7% for cooking [47].

In some models, some building shell insulation variables such as basement insulation had positive coefficients, implying that controlled for other factors, insulated houses in our sample consumed more electricity on average. However, this effect was not physically substantial. Previous studies who have examined the effect of insulation on electricity consumption have also reported mixed results. While some have verified that insulation reduces electricity consumption [48], [49] and [50], some others have reported no significant correlation between insulation level and residential electricity consumption [39]. Still, another group of studies have reported that buildings with higher insulation ratings are likely to consume more electricity [51]. Several reasons may contribute to increased energy consumption in well-insulated homes. One potential reason for positive correlation of insulation and energy consumption is higher cooling load in well-insulated houses due to the heavier use of AC instead of natural ventilation [52]. This is relevant since our data set covered the entire summer but not the entire winter months, so the effect of cooling loads on our results is more significant. Another potential reason for this observation is the positive correlation of insulation with House Size, appliance stock, or income [53], all of which are correlated with higher electricity consumption. Finally, the relationship between insulation and electricity consumption may be in the opposite direction: larger houses with higher energy consumption are more motivated (and more plausible) to add extra insulation; i.e., insulation might have been a result of high consumption, not a cause for it. If it was not for the insulation, these houses would have consumed much higher energy than other houses, but insulation helps to mitigate that effect [51].

Use of energy-efficient lights was correlated with lower consumption, as did the use of double-pane windows. Finally, use of programmable thermostats was correlated with increased electricity consumption. One possible explanation for this trend is that programmable thermostats turn the air conditioner or heater on at certain times, even if the house is not occupied. Therefore they may actually increase heating or cooling load compared with manual settings, especially when the house is not occupied regularly. Finally, our data did not show a statistically significant difference in electricity consumption between rented and owned houses, contrary to several previous studies [43].

5.3 The Effect of Occupants

We analyzed the effect of occupants from three different perspectives: the effect of occupancy level, the effect of occupant behavior (long-term habits and preferences), and the effect of occupant socioeconomic status. The following sections summarize our results.

5.3.1 The Effect of Occupancy Level

Number of occupants is a significant variable in daily maximum models, while it is not a significant variable in daily minimum models. This observation supports the notion that the presence of occupants primarily impacts the consumption in excess of the daily minimum. Furthermore, the models suggest a non-linear relationship between household electricity consumption and the number of occupants, selecting the square root of number of occupants over the number of occupants. In other words, our model verifies that when the number of occupants double, electricity consumption increases at a slower rate (1.4 in our data), leading to the conclusion that larger households have higher aggregate electricity consumption but lower per capita consumption. A similar concave relationship between number of occupants and electricity consumption has been reported in former studies [54], [55] and [56].

Pet ownership is a statistically significant factor in the models, while the magnitude of its impact is the largest for the summer daily minimum, winter daily maximum, and winter daily maximum–minimum models. The effect is significant even after removing the effect of other significant variables. Our hypothesis for explaining the correlation between pet ownership and higher electricity consumption is two-fold: (a) it is likely that people who have pets spend more time at home, and (b) even when they leave home, they do not completely turn the AC off. Therefore, pet ownership may be a proxy for the percentage of time that the house is "active". Further research is needed to validate this hypothesis.

Age of occupants showed a significant correlation with electricity consumption. Specifically, households with occupants older than 55 and between 19 and 35 showed lower electricity consumption. The older household members tend to be more conscious about the way they use electricity, and also tend to use less electric gadgets. On the other hand, household members between 19 and 35 are more likely to have a full-time job and therefore are less at home. These observations suggest that both the occupancy rate (how much time one spends at home) and intensity of energy use of household members influence electricity consumption [57].

5.3.2 The Effect of Long-Term Habits and Preferences

Behavioral factors that have long-term impacts (such as Purchasing Energy-Star Appliances and Air Conditioners) or are considered long-term habits (such as adjusting thermostat settings moderately) are significant explanatory variables for daily minimum consumption.

As Table 3 shows, in the daily minimum model, the behavior of Purchasing Energy-Star Appliances and Air Conditioners has a positive coefficient. This suggests that, in our study sample, contrary to common belief, households that have expressed motivation to buy energy-efficient appliances and air conditioners have higher levels of daily minimum consumption, after adjusting for all other variables. Since most new higher-end appliances are energy-star, the tendency to purchase energy-star appliances may simply mean buying new, high-quality appliances, which is by itself a proxy for wealth [58], [59], [60] and [61]. Another explanation for the positive coefficient of energy-star appliances the "rebound effect", where an increase in the efficiency of appliances results in increased use of them, hence an increase in overall energy consumption [3] and [6].

Another long-term habit, Turning Off Lights When Not in Use, is significant in most winter models but not at all significant in summer models. A closer look at this variable reveals a significant geographical pattern: people in the two coasts in our sample declare much higher rates of turning off lights when not in use. Also, this variable becomes insignificant when Zip Code is included in the model, which is another evidence to its geographical trend. Last, in models where this variable is indeed significant, its coefficient is positive, implying that those in our sample who declare they turn lights off when leaving the room consume more energy on average. Our conclusion is that Turning Off Lights When Not in Use is capturing a latent factor in our sample, and further data are needed to quantify the effect of energy-conscious behavior of turning off unnecessary lights.

5.3.3 Effect of Income Level

We did not observe any statistically significant correlation between income level and electricity consumption. The relationship between household income and energy consumption has been the subject of extensive research. While a large number of studies have concluded that energy consumption increases monotonically with income [19], [28], [51], [62], [63], [64] and [65], a number of studies have reported observing an inverted U-path relating energy consumption and household income [59], [66] and [67]. At the same time, the effect of income on household electricity consumption has been shown to be mediated by ownership of appliances; i.e., income of the household impacts the consumption through affecting the stock (quantity and quality) of appliances [17], [27], [68] and [69]. This is especially true for populations like our sample where households are similar to each other in socioeconomic status and therefore the income effect is minimal.

6 Conclusions

Summarizing our findings, we showed that:

(a) Factors that influence residential electricity consumption can be categorized into four major groups: external conditions (e.g., location and weather), physical characteristics of dwelling, appliance and electronics stock, and occupants.

(b) Each of the four categories above, on average, has a different time span and effort level for modification; while location, weather, and occupancy are outside the scope of influence for modification, physical characteristics of the building, appliance stock, and occupant behavior factors can be modified in long-term, medium-term, and short-term investment spans, respectively. Accordingly, the persistence of the modification effect is generally proportional to the level of effort and investment that was allocated to it.

(c) Daily minimum and daily maximum consumption are explained by different sets of explanatory factors. Daily minimum has a lower variation level compared with daily maximum, and is best explained by factors that are steady through time, such as weather (Degree Days), location (Zip Code), House Size, and number of refrigerators. On the other hand, daily maximum is best explained by large and intermittent loads such as electric water heater and air conditioners.

(d) Using our model, we were able to explain 55–65% of the variability in electricity consumption, as measured by the R2 of the regression model. This is comparable with most studies in the past. Using variable transformations and other machine learning techniques [70], we were able to achieve R2 values of above 70%. However, we prefer the linear model due to its interpretability. Furthermore, we deliberately did not use variables such as households' estimate of their electricity bill (available from the survey) that improve the R2 of the fit, but add little explanatory significance to the results of the model.

(e) Overall, weather and physical characteristics of the building illustrate more influence on residential electricity consumption compared to other categories such as occupant behavior. These results are comparable with the work of [43] who showed that building characteristics determine 42% of the variability in residential electricity consumption, whereas occupant behavior explains 4.2%. Within the physical characteristics of the building, floor area, type of building, and use of electric water heater contributed the most to consumption, whereas within the appliance stock, number of refrigerators was the most important factor. Finally, pet ownership (which can be considered a proxy for the percent of time that the house is active) was a significant factor in explaining variation in electricity consumption.

7 Policy Implications

Based on the results of our models, we highly recommend policies and regulations aimed at improving the thermal performance of buildings, including both improvements to the insulation level of the dwelling and improving the efficiency of the stock of air conditioning and electric heaters. Since certain end uses such as space heating are more prone to rebound effects [19], we strongly recommend provisions for regular home energy audits in codes and regulations [49]. Furthermore, we recommend policies and regulations aimed at improving the efficiency of the appliance stock. Certain end uses such as refrigerators illustrate great potential for consumption reduction. Policies and programs that encourage the purchase of energy-efficient refrigerators and other appliances must also devise provisions for buying back the old refrigerators or make the financial incentive contingent on households returning the old refrigerators.

Most high-consumption, intermittent appliances such as electric water heater, electric clothes dryer, and Spas/Pools demand high volume and intermittent electric loads, hence are attractive targets for load shifting programs. Since these end uses are primarily driven by occupants' habits and activities, respective energy efficiency programs should be focused on behavioral modification. For example, educational campaigns encouraging households to use larger loads of laundry, to lower the temperature of their electric water heaters, or to shift their laundry time to a more appropriate time in the day can be effective in this regard.

In terms of the impact of behavioral factors, our results agree with some previous works that residential electricity consumption is primarily determined through the way households use electricity, and less by the way that they value energy efficiency [28] and [40]. However, energy-saving values usually impact efficiency through “habits” such as Purchasing Energy-Star Appliances. A dual approach as suggested by Kelly [51], where less efficient homes are encouraged to perform structural changes (such as adding insulation) while more efficient homes are motivated to reduce consumption through behavioral interventions and economic incentives is supported by our analyses [40]. Energy saving values and habits are shown to respond to changes in price of electricity [63] and [71]; therefore, behavior modification programs can be more effective when supported by monetary and regulatory policies.

8 Future Work

More data are needed to validate some of the findings of this paper. Specifically, household data from a more heterogeneous sample over a larger period of time are needed for validating the generality of these results. The use of self-reports to measure behavior may have introduced some bias in the data, called "social desirability" bias [72]. However, since the purpose of this study was to explain the variability in electricity consumption, and the households in our study were all from middle and upper social class, we assume that the bias in responses was uniform over the respondents and therefore the results of the model explain the variability in electricity consumption with a reasonable accuracy.

In this paper, we examined energy consumption and its features such as daily maximum and minimum consumption, and explained their variability using household data. A potential follow-up to this study is to develop a metric for quantifying energy efficiency of the households, and compare households using that metric instead of their consumption data. Such metric needs to be defined in a way that recognizes the inherent differences among different groups of households and at the same time enables comparison across those different groups.

7 Acknowledgments

The majority of this work was funded through Advanced Research Projects Agency-Energy (ARPA-E) grant to Stanford University. Ram Rajagopal was supported by the Powell Foundation Fellowship. The authors would like to thank Dr. June Flora at H-STAR Institute at Stanford University for helpful inputs and comments during this study. The authors are also grateful to anonymous referees for their valuable comments.

8 References

[1] US Department of Energy. Building energy data book 2010 (2011). Tech. Rep.. US Department of Energy, Office of Energy Efficiency and Renewable Energy. URL http://buildingsdatabook.eren.doe.gov/DataBooks.aspx; 2011.

[2] US Energy Information Administration. Annual energy outlook 2011 (2011). Tech. Rep. US Energy Information Administration. http://dx.doi.org/DOE/EIA-0383(2011). URL, www.eia.gov/forecasts/aeo/; 2011 April.

[3] Beerepoot M. (2007) Energy policy instruments and technical change in the residential building sector. Alblasserdam, The Netherlands: Haveka; 2007. ISBN 9781586038113.

[4] Pina A, Silva C, Ferrão P. (2012) The impact of demand side management strategies in the penetration of renewable electricity. Energy 2012; 41(1); 128:137. http://dx.doi.org/10.1016/j.energy.2011.06.013. URL, http://linkinghub.elsevier.com/retrieve/pii/S0360544211003902.

[5] Haas R. (1997) Energy efficiency indicators in the residential sector: what do we know and what has to be ensured? Distribution 1997; 25; 789-802. PII:S0301-4215(97)00069-4.

[6] Abrahamse W, Steg L, Vlek C, Rothengatter T. (2005) A review of intervention studies aimed at household energy conservation. Journal of Environmental Psychology 2005; 25(3); 273-291. http://dx.doi.org/10.1016/j.jenvp.2005.08.002. URL http://linkinghub.elsevier.com/retrieve/pii/S027249440500054X.

[7] Bartels R, Fiebig D. (1998) Metering and modelling residential end use electricity load curves. Forecasting 1998; 15(6); 415-426. http://dx.doi.org/10.1002/(SICI)1099-131X(199611)15:6<415::AID-FOR636>3.0.CO;2-J.

[8] Tonn B. (1990) Residential wood combustion for space heating in the Pacific northwest difficulties for electric utilities? Energy Policy 1990; 18(3); 283-292. http://dx.doi.org/10.1016/0301-4215(90)90219-T. URL, http://linkinghub.elsevier.com/retrieve/pii/030142159090219T.

[9] Masiello JA, Parker DS. (1992) Factors influencing water heating energy use and peak demand in a large scale residential monitoring study. In: Residential buildings: technologies, design, performance analysis, and building industry trends 1992. p.157-170.

[10] Aigner D, Sorooshian C, Kerwin P. (1984) Conditional demand analysis for estimating residential end-use load profiles. The Energy Journal 1984; 5(4); 81-97.

[11] Aydinalp M, Ugursal V, Fung AS. (2003) Modelling of residential energy consumption at the national level. International Journal of Energy Research 2003; 27(4); 441-453. http://dx.doi.org/10.1002/er.887. URL, http://doi.wiley.com/10.1002/er.887.

[12] Caves DW, Herriges JA, Train KE, Windle RJ. (1987) A Bayesian approach to combining conditional demand and engineering models of electricity usage. The Review of Economics and Statistics 1987; 69(3); 438-448.

[13] Hsiao C, Mountain C, Illman KH. (1995) A Bayesian integration of end-use metering and conditional-demand analysis. Journal of Business and Economic Statistics 1995; 13(3); 315-326. http://dx.doi.org/10.2307/1392191. http://dx.doi.org/10.1080/07350015.1995.10524605,

[14] Goldfarb DL, Huss R. (1988) Building utility scenarios for an electric. Long Range Planning 1988; 21(2); 78-85.

[15] LaFrance G, Perron D. (1994) Evolution of residential electricity demand by end-use in Quebec 1979-1989: a conditional demand analysis evolution of residential electricity demand by end-use in Quebec. Energy Studies Review 1994; 6(2): 164-173.

[16] Lins M, DaSilva A, Rosa L. (2003) Regional variations in energy consumption of appliances: conditional demand analysis applied to Brazilian households. Annals of Operations Research 2003; 235-246.

[17] Parti M, Parti C. (1980) The total and appliance-specific conditional demand for electricity in the household sector. The Bell Journal of Economics 1980; 11(1); 309-321. http://dx.doi.org/10.2307/3003415.

[18] Swan LG, Ugursal VI. (2009) Modeling of end-use energy consumption in the residential sector: a review of modeling techniques. Renewable and Sustainable

Energy Reviews 2009; 13(8); 1819-1835.
http://dx.doi.org/10.1016/j.rser.2008.09.033.URL
http://linkinghub.elsevier.com/retrieve/pii/S1364032108001949.

[19] Cayla JM, Maizi N, Marchand C. (2011) The role of income in energy consumption behaviour: evidence from French households data. Energy Policy 2011; 39(12); 7874-7883. http://dx.doi.org/10.1016/j.enpol.2011.09.036. URL http://linkinghub.elsevier.com/retrieve/pii/S0301421511007257.

[20] Sütterlin B, Brunner TA, Siegrist M. (2011) Who puts the most energy into energy conservation? A segmentation of energy consumers based on energy-related behavioral characteristics. Energy Policy 2011; 39(12); 8137-8152. http://dx.doi.org/10.1016/j.enpol.2011.10.008. URL http://linkinghub.elsevier.com/retrieve/pii/S0301421511007920.

[21] Baltagi B. (2002) Comparison of forecast performance for homogeneous, heterogeneous and shrinkage estimators: some empirical evidence from US electricity and natural-gas consumption. Economics Letters 2002; 76(3); 375-382. http://dx.doi.org/10.1016/S0165-1765(02)00065-4. URL http://linkinghub.elsevier.com/retrieve/pii/S0165176502000654.

[22] Haas R, Schipper L. (1998) Residential energy demand in OECD-countries and the role of irreversible efficiency improvements. Energy Economics 1998; 20(4); 421-442. http://dx.doi.org/10.1016/S0140-6701(99)96911-4. URL, http://linkinghub.elsevier.com/retrieve/pii/S0140670199969114.

[23] Halvorsen R. (1975) Residential demand for electric energy. The Review of Economics and Statistics 1975; 57(1) 12-18.

[24] Hirst E. (1978) A model of residential energy use. Simulation 1978; 30(3); 69-74. http://dx.doi.org/10.1177/003754977803000301. URL, http://sim.sagepub.com/cgi/doi/10.1177/003754977803000301.

[25] Houthakker H. (1980) Residential electricity revisited. Energy 1980; 1(1); 29-41.

[26] Kamerschen DR, Porter DV. (2004) The demand for residential, industrial and total electricity, 1973-1998. Energy Economics 2004; 26(1); 87-100 PII: S0140-9883Z03.00033-1. URL, http://linkinghub.elsevier.com/retrieve/pii/S0140988303000331.

[27] Dubin JA, Mcfadden D. (1984) An econometric analysis of residential electric appliance holdings and consumption. Society 1984; 52(2); 345-362. URL, http://www.jstor.org/stable/1911493.

[28] Cramer JC, Miller N, Craig P, Hackett BM. (1985) Social and engineering determinants and their equity implications in residential electricity use. Energy 1985; 10(12); 1283-1291. http://dx.doi.org/0360-5442/85. URL http://linkinghub.elsevier.com/retrieve/pii/0360544285901392.

[29] Thorndike R. (2005) History of factor analysis: a statistical perspective. Encyclopedia of Statistics in Behavioral Studies. URL, http://www.wiley.com//legacy/wileychi/eosbs/; 2005.

[30] Everitt B, Hothorn T. (2011) Exploratory factor analysis. In: An introduction to applied multivariate analysis with R; chap. 05. NY: Springer. p. 135-161. ISBN 978-1-4419-9649-7. URL, http://www.springerlink.com/index/10.1007/978-1-4419-9650-3; 2011.

[31] Hastie T, Tibshirani R. (2011) The elements of statistical learning. Springer; 2011.

[32] R Development Core Team (2008) R language and environment for statistical computing. ISBN 3-900051-07-0. URL, http://www.R-project.org; 2008.

[33] Houde S, Todd A, Sudarshan A, Flora JA, Armel CK. (2012) Real-time feedback and electricity consumption: a field experiment assessing the potential for savings and persistence. Energy 2012; 1-22.

[34] Ehrhardt-Martinez K, Donnelly KA. (2010) Advanced metering initiatives and residential feedback programs: a meta-review for household electricity-saving opportunities. Report Number E105; American Council for Energy Efficient Economy, Washington, DC, USA.

[35] Courrieu P, Rey A. (2011) Missing data imputation and corrected statistics for largescale behavioral databases. Behavior Research Methods 2011; 43(2); 310-330. http://dx.doi.org/10.3758/s13428-011-0071-2. URL, http://www.ncbi.nlm.nih.gov/pubmed/21424187.

[36] Gelman A, Hill J. (2007) Missing data imputation. Behavior Research Methods 2007; 43(2); 310-330. http://dx.doi.org/10.1186/1471-2288-11-105.

[37] NOAA. (2010) Divisional temperature maps of the US Jan-Nov 2010.

[38] Lijesen M. (2006) The real-time price elasticity of electricity. Energy Economics 2007; 29(2); 249-258. http://dx.doi.org/10.1016/j.eneco.2006.08.008. URL, http://linkinghub.elsevier.com/retrieve/pii/S0140988306001010.

[39] Cramer JC, Hackett B, Craig PP, Vine E, Levine M, Dietz TM, et al. (1984) Structural and behavioral determinants of residential energy use: summer electricity use in Davis. Energy 1984; 9(3); 207-216. http://dx.doi.org/0360-5442/84. URL, http://linkinghub.elsevier.com/retrieve/pii/0360544284901087.

[40] Gouveia JAP, Fortes P, Seixas J. (2012) Projections of energy services demand for residential buildings: insights from a bottom-up methodology. Energy 2012; 47(1); 430-442. http://dx.doi.org/10.1016/j.energy.2012.09.042. URL, http://linkinghub.elsevier.com/retrieve/pii/S0360544212007207.

[41] Howard B, Parshall L, Thompson J, Hammer S, Dickinson J, Modi V. (2011) Spatial distribution of urban building energy consumption by end use. Energy and Buildings 2012; 45; 141-451. http://dx.doi.org/10.1016/j.enbuild.2011.10.061. URL, http://linkinghub.elsevier.com/retrieve/pii/S037877881100524X.

[42] Pelenur MJ, Cruickshank HJ. (2012) Closing the Energy Efficiency Gap: a study linking demographics with barriers to adopting energy efficiency measures in the home. Energy 2012; 47(1); 348-357. http://dx.doi.org/10.1016/j.energy.2012.09.058

[43] Guerra Santin O, Itard L, Visscher H. (2009) The effect of occupancy and building characteristics on energy use for space and water heating in Dutch residential stock. Energy and Buildings 2009; 41(11); 1223-1232. http://dx.doi.org/10.1016/j.enbuild.2009.07.002. URL, http://linkinghub.elsevier.com/retrieve/pii/S0378778809001388.

[44] US Energy Information Administration. (2005) US household electricity report 2005. Tech. Rep. Washington, DC: US EIA; 2005. URL http://www.eia.gov/emeu/reps/enduse/er_contents.html.

[45] Liao HC, Chang TF. (2002) Space-heating and water-heating energy demands of the aged in the US. Energy Economics 2002; 24; 267-284.

[46] Linden AL, Carlssonkanyama A, Eriksson B. (2006) Efficient and inefficient aspects of residential energybehaviour:what are thepolicy instruments for change? Energy Policy 2006; 34(14); 1918-1927. http://dx.doi.org/10.1016/j.enpol.2005.01.015. URL, http://linkinghub.elsevier.com/retrieve/pii/S0301421505000583.

[47] Janssen R. (2004) Towards energy efficient buildings in Europe. Tech. Rep.. London, UK: EuroAce, The European Alliance of Companies for Energy Efficiency in Buildings; 2004. URL http://www.euroace.org/MediaPublications/PublicationsReports.aspx

[48] Assimakopoulos V. (2000) Residential energy demand modelling in developing regions: the use of multivariate statistical techniques. Energy Economics 02/1992; 14(1); 57-63. DOI:10.1016/0140-9883(92)90025-9.

[49] Hirst E, Goeltz R. (1985) Comparison of actual energy savings with audit predictions for homes in the north central region of the U. S. A. Building and Environment 1985; 20(1):1-6.

[50] Sonderegger RC. (1977) Movers and Stayers: the Resident's contribution to variation across houses in energy consumption for space heating. Energy and Buildings 1977; 1:313-324.

[51] Kelly S. (2011) Do homes that are more energy efficient consume less energy?: a structural equation model of the English residential sector. Energy 2011; 36(9):5610-5620. http://dx.doi.org/10.1016/j.energy.2011.07.009. URL, http://linkinghub.elsevier.com/retrieve/pii/S0360544211004579.

[52] Li DH, Yang L, Lam JC. (2012) Impact of climate change on energy use in the built environment in different climate zones e a review. Energy 2012; 42(1):103-112. http://dx.doi.org/10.1016/j.energy.2012.03.044. URL, http://linkinghub.elsevier.com/retrieve/pii/S0360544212002459.

[53] Smiley R. (1979) The levels and determinants of household insulation activity. Energy 1979;4:1087-1093.

[54] Barnes DF, Krutilla K,HydeW. (2004) The urban household energy transition. Tech. Rep.. Resources for the Future; 2004 March. URL

http://siteresources.worldbank.org/INTGENENERGY/Resources/UrbanEnergyTransitionV51.pdf.

[55] Heltberg R. (2005) Factors determining household fuel choice in Guatemala. Environment and Development Economics 2005; 10(3):337-361. http://dx.doi.org/10.1017/S1355770X04001858. URL, http://www.journals.cambridge.org/abstract_S1355770X04001858.

[56] Xiaohua W, Zhenmin F. (2003) Common factors and major characteristics of household energy consumption in comparatively well-off rural China. Renewable and Sustainable Energy Reviews 2003; 7(6):545-552. http://dx.doi.org/10.1016/S1364-0321(03)00080-7. URL, http://linkinghub.elsevier.com/retrieve/pii/S1364032103000807.

[57] O'Neill B, Chen BS. (2002) Demographic determinants of household energy use in the United States. Population and Development Review 2002; 28:53-88.

[58] Elias RJ, Victor DG. (2005) Energy transitions in developing countries e a review of concepts and literature. Working Paper 2005;40. URL, http://pesd.stanford.edu.

[59] Foster V, Tre JP, Wodon Q. (2000) Energy consumption and income: an inverted-U at the household level? World Bank, the Office of the Chief Economist for Latin America (Guillermo Perry) and by the ESMAP program. Latin America Region.

[60] Kowsari R, Zerriffi H. (2011) Three dimensional energy profile. Energy Policy 2011;39(12):7505e17. http://dx.doi.org/10.1016/j.enpol.2011.06.030. URL, http://linkinghub.elsevier.com/retrieve/pii/S0301421511004861.

[61] Leach G. The energy transition. (1992) Energy Policy 1992;20(2):116e23. http://dx.doi.org/0301-4215/92/020116-0. URL, http://linkinghub.elsevier.com/retrieve/pii/030142159290105B.

[62] Biesiot W, Noorman KJ. (1999) Energy requirements of household consumption: a case study of The Netherlands. Ecological Economics 1999; 28:367-383.

[63] Filippini M. (2011) Short- and long-run time-of-use price elasticities in Swiss residential electricity demand. Energy Policy 2011;39(10):5811-5817. http://dx.doi.org/10.1016/j.enpol.2011.06.002. URL, http://linkinghub.elsevier.com/retrieve/pii/S0301421511004563.

[64] Vringer K, Aalbers T, Blok K. (2007) Household energy requirement and value patterns. Energy Policy 2007; 35(1):553-566. http://dx.doi.org/10.1016/j.enpol.2005.12.025

[65] Daioglou V, van Ruijven BJ, van Vuuren DP. (2012) Model projections for household energy use in developing countries. Energy 2012;37(1):601-615. http://dx.doi.org/10.1016/j.energy.2011.10.044. URL, http://linkinghub.elsevier.com/retrieve/pii/S0360544211007110.

[66] Choi E, Heshmati A, Cho Y. (2010) An empirical study of the relationships between an empirical study of the relationships between CO2 emissions,

economic growth and openness. Tech. Rep. 5304. Bonn, Germany: The Institute for the Study of Labor; 2010. URL http://econpapers.repec.org/paper/izaizadps/dp5304.htm.

[67] Boopen S, Vinesh S. (2011) On the relationship between CO2 emissions and economic growth: the Mauritian experience. In: CSAE 25th anniversary conference 2011: economic development in Africa 2011. p. 1-25.

[68] Sudarshan A. (2011) Deconstructing the Rosenfeld curve: structural determinants of energy consumption. Phd thesis. Stanford University, Stanford, California, USA; 2011.

[69] Reiss PC, White MW. (2005) Household electricity demand, revisited. Review of Economic Studies 2005; 72(3):853-883. http://dx.doi.org/10.1111/0034-6527.00354. URL, http://restud.oxfordjournals.org/lookup/doi/10.1111/0034-6527.00354.

[70] Kolter JZ, Ferreira J. (2009) A large-scale study on predicting and contextualizing building energy usage. Journal of Machine Learning Research 2009.

[71] He Y, Wang B, Wang J, Xiong W, Xia T. (2012) Residential demand response behavior analysis based on Monte Carlo simulation: the case of Yinchuan in China. Energy 2012;47(1):230e6. http://dx.doi.org/10.1016/j.energy.2012.08.046. URL, http://linkinghub.elsevier.com/retrieve/pii/S0360544212006779.

[72] Randall DM, Fernandes MF. (1991) The social desirability response bias in ethics research express. Journal of Business Ethics 1991; 10(11):805-817. http://dx.doi.org/10.1007/BF00383696. URL: http://link.springer.com/article/10.1007/BF00383696.

Chapter 3

Data-Driven Benchmarking of Building Energy Efficiency Utilizing Statistical Frontier Models[3]

Amir Kavousian, Ram Rajagopal

1 Abstract

Frontier methods quantify the energy efficiency of buildings by forming an efficient frontier (best-practice technology) and comparing all buildings against that frontier. Because energy consumption fluctuates over time, the efficiency scores are stochastic random variables. Existing applications of frontier methods in energy efficiency either treat efficiency scores as deterministic values, or estimate their uncertainty by resampling from one set of measurements. Availability of smart meter data (repeated measurements of energy consumption of buildings) enables using actual data to estimate the uncertainty in efficiency scores. Also, existing applications assume a linear form for efficient frontier; i.e., assume that the best-practice technology scales up and down proportionally with building characteristics. However, previous research shows that buildings are non-linear systems.

This paper proposes a statistical method called Stochastic Energy Efficiency Frontier (SEEF) to estimate bias-corrected efficiency score and its confidence intervals from measured data. The paper proposes an algorithm to specify the functional form of the

---

[3] This chapter was first published in the ASCE Journal of Computing in Civil Engineering as: Kavousian, A. and Rajagopal, R. (2014). "Data-Driven Benchmarking of Building Energy Efficiency Utilizing Statistical Frontier Models." J. Computing in Civil Engineering, 10.1061/(ASCE)CP.1943-5487.0000327 (January, 2014).

frontier, identify the probability distribution of efficiency score of each building using measured data, and rank buildings based on their energy efficiency.

To illustrate the power of SEEF, results from applying SEEF on a smart meter data set of 307 residential buildings in the US are presented. SEEF efficiency scores are used to rank individual buildings based on energy efficiency, to compare sub-populations of buildings, and to identify irregular behavior of buildings across different time-of-use periods.

SEEF is an improvement over energy intensity method (comparing kWh/Sq.Ft): While SEEF identified efficient buildings across the entire spectrum of building sizes, the energy intensity method showed bias towards smaller buildings. The results of this research can help researchers and practitioners compare and rank (i.e., benchmark) buildings more robustly and over a wider range of building types and sizes. Eventually, this results in improved resource allocation in energy efficiency programs.

2 Introduction

Quantifying building energy efficiency is essential to developing and monitoring energy efficiency programs (Ahn et al., 2010). Two categories of methods for quantifying building energy efficiency are engineering and empirical methods (Lutzenhiser et al., 2010).

Engineering methods simulate building performance by creating thermodynamic models of building operations (Clevenger et al. 2012). These models utilize buildings' design information and industry averages of the fixtures to predict energy usage

(Xydis et al., 2008). Depending on the level of detail of the model, availability of reliable data, quality of assumptions, and the accuracy of the simulation engine, engineering models can be very accurate in identifying sources of inefficiency in a building (Griffith & Crawley, 2006; Pisello et al. 2012; Waltner et al., 2012; Goldstein, 2012). However, such detailed models are time and skill intensive to build and maintain, hence cost prohibitive in most instances (Patterson 1995; Yezioro et al. 2008; Maile, 2009).

Empirical methods utilize measured energy consumption data to create models using variables such as climate, building characteristics, energy fixtures (e.g., appliances and equipment), and occupancy. Most empirical methods (such as Energy Star®) depend on regression models, and hence are limited by the shortcomings of ordinary regression (e.g. Atkinson, 1982; Riani & Atkinson, 2000; Savage, 2009). Regression methods identify the average trend over the entire population, and then compare each building with that overall trend. Thus, if a few badly-performing buildings skew the regression line toward less-efficient buildings, a building with modest performance may appear to be highly efficient. In contrast, non-parametric empirical benchmarking models such as Data Envelopment Analysis (DEA—Farrell, 1957) compare a building with the best-performing building in its own class. Comparing each building with the best-practice technology in its own class (or of similar functionality) enables estimating total potential for energy efficiency. The goal is to estimate how much reduction in energy consumption is possible without compromising the functionality of the building.

In DEA model, the class of each building is quantified by a measure of its output capacity, which is called Level of Service (LOS) in this paper. LOS is an indicator of the utility (service, output) that a building provides to its occupants. The LOS can be defined in many ways, such as floor area, number of occupants, a weighted combination of energy fixtures, or the physical output of a building (e.g., number of products produced in a manufacturing plant). Once the LOS is defined, the efficient frontier in DEA model specifies the lowest amount of energy for a given LOS (Figure 7).

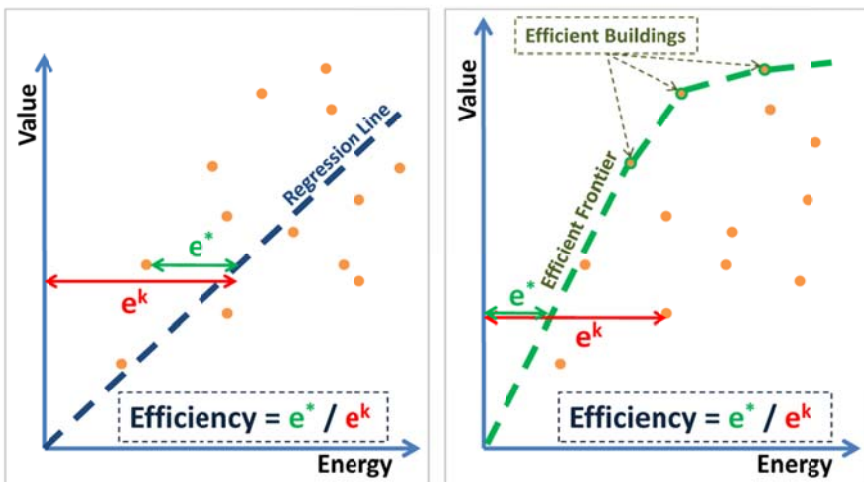


Figure 7 - (a) Left: Schematic of a regression model. (b) Right: Schematic of a DEA model.

Frontier methods such as DEA quantify energy efficiency by examining how much energy a building consumes in relation to the service that it provides. Frontier methods divide into two categories: parametric and non-parametric. Parametric frontier methods, also called Stochastic Frontier Analysis (SFA), assume that the efficient frontier can be presented using a mathematical formula plus a stochastic error term called random component (Aigner et al., 1977). Non-parametric frontier methods such as Data Envelopment Analysis (DEA) make no assumptions about the mathematical formula of the frontier, and instead form the frontier using measured data. Our model (Stochastic Energy Efficiency Frontier—SEEF) is based on the non-parametric assumption, and therefore is an extension of the DEA model.

In the rest of this paper, we first explain DEA model and its applications in energy efficiency field, and then introduce SEEF model and explain its innovations over existing methods. To illustrate the power of SEEF, we apply it to a smart meter data set and describe the insights that can be obtained using SEEF.

2.1 DEA Model

Consider k buildings, each consuming a vector of n energy sources $x_i \in \mathcal{R}_+^n$ to provide a vector of m services $s_i \in \mathcal{R}_+^m$. We denote the matrix of observed services by $S_{(k,m)}$ and the matrix of observed inputs by $X_{(k,n)}$.

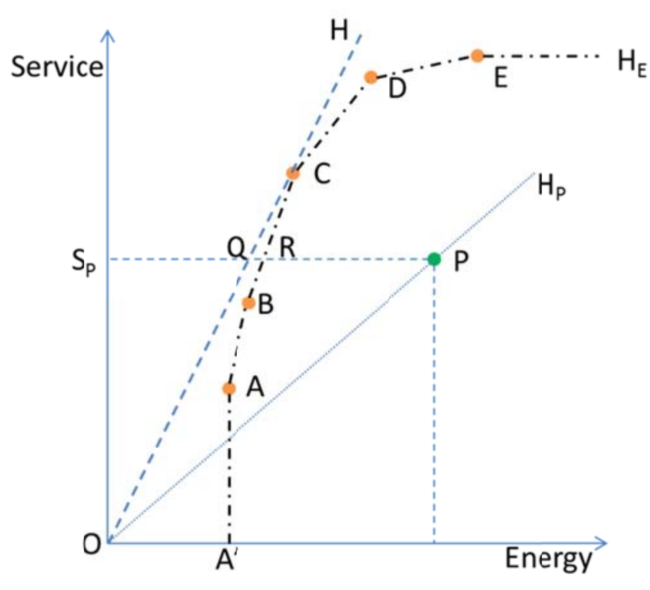


Figure 8 - Different functional forms for the efficient frontier.

A Linear Programming Problem (LPP) such as LPP 1 calculates the efficiency values of buildings given input matrix X and service matrix S (Farrell, 1957). For illustration purposes, assume the LPP corresponding to linear frontier (ray OH in Figure 8):

$\theta_i(s,x) = \min \lambda$
Subject to:

$$zS \geq s \qquad\qquad \text{(LPP1 = LPP(S,X))}$$
$$zX \leq \lambda x$$
$$z \in \mathcal{R}_+^k$$

The efficiency of observation P is given by the solution to the LPP (1); i.e., $\theta_i(s,x)$, which is also equal to the ratio S_PQ / S_PP. In other words, λ is the ratio by which it is possible to reduce the energy consumption of a building without reducing its level of service. By minimizing λ we find the maximum achievable energy reduction without reducing level of service.

z is the intensity vector and serves to construct the smallest cone which covers the data in the sample. Adding constraints to the values of z changes the shape of the frontier, as well as the efficiency values. These constraints depend on the assumptions about the relationship between energy consumption and LOS, and are called "returns to scale technologies". The term *returns to scale* is frequently used in economics to describe the relationship between changes in inputs versus outputs. For instance, a Decreasing Returns to Scale (DRS) implies that the output (in our case LOS) grows with a rate slower than that of input (in our case energy consumption). A Constant Returns to Scale (CRS) implies that the LOS and energy consumption can increase or decrease with the same rate. Finally, a Variable Returns to Scale (VRS) means the rate of change of LOS can be higher or lower than that of energy consumption depending on the size of energy consumption. Table 1 lists different assumptions for returns to scale and their corresponding LPP constraints.

Table 6 - Different returns to scale technologies and their respective LPP constraints and frontier shapes.

	Constant Returns to Scale	**Decreasing Returns to Scale**	**Variable Returns to Scale**
Constraint for $\sum_{i=1}^{k} z_i$ term	None	$\sum_{i=1}^{k} z_i \leq 1.$	$\sum_{i=1}^{k} z_i = 1.$
Frontier Shape	Ray OH	points O, C, D, E and the ray EH_E and the x axis	A', A, B, C, D, E plus the ray EH_E and the x axis starting at A'.

2.2 Related Work

Previous studies have applied DEA to estimate the energy efficiency of buildings. For example, Carnes et al. (1998) applied DEA to examine energy efficiency of Texas State universities. They used electricity, fuel, and thermal energy as energy inputs, and used course credits taught and research expenditures as the outputs of the energy consumption for Texas universities. They discussed several variations of DEA, including the use of multiple outputs and multiple inputs (e.g., non-energy inputs such as capital investment). Onut & Soner (2006) used DEA to analyze the energy efficiency of five-star hotels in Antalya, Turkey and made recommendations for increasing the energy efficiency. Grosche's (2008) applied DEA to the US Residential Energy Consumption Survey (RECS) data from 1997 and 2001 and compared the energy efficiency of US residential building stock in that period. Grosche concluded that the energy efficiency in US households improved from 1997 and 2001, but most of that improvement came from households using natural gas and fuel oil for their heating purposes (Grosche, 2008).

One limitation of studies applying DEA to energy efficiency is that they all have used a linear parametric frontier, without a data driven validation of the frontier shape. We show later in this paper how the shape of the frontier can be validated using the data.

Another limitation in existing studies of DEA in energy efficiency is the assumption of deterministic energy efficiency scores. Before widespread use of high-frequency meters, studies that analyzed energy consumption used a limited data set of energy consumption, usually available for only a short period of time. Due to lack of repeated measurements, these studies assumed a deterministic consumption level for each building. However, energy consumption has a significant stochastic component, impacted by factors such as weather and occupancy patterns (Azar & Menassa, 2011; Berges et al. 2011; Chen et al. 2012). Since DEA is highly adaptive to data, efficiency estimates based on single measurements can be very biased and unreliable if reported without estimating their error distributions.

In limited data settings, Simar & Wilson (1998) proposed a bootstrapping method that estimates a confidence band for the frontier by reassigning efficiency scores across buildings. This resampling across buildings attempts to simulate input stochasticity by reutilizing the same fixed measurements. The implied assumption of this method is that efficiencies are not a function of individual and potentially unobserved building characteristics. Thus, when multiple measurements per building are available, the bootstrapping method is not the optimal method to estimate uncertainty in energy efficiency scores; instead, the actual repeated measurements of energy consumption should be used.

This paper addresses these various challenges and limitations by introducing a new methodology, Stochastic Energy Efficiency Frontier (SEEF), that combines: (a) a hypothesis-testing method to choose efficient frontier shape; (b) a methodology to estimate the uncertainties in the energy efficiency scores using available repeated measurements; (c) a method to use the estimated uncertainties to create a bias-corrected efficient frontier and a confidence band for the frontier; and (d) a statistical ranking mechanism to compare efficiency scores across the building population.

3 SEEF Model and Algorithm

Stochastic Energy Efficiency Frontier (SEEF) is a stochastic performance-based method to rank buildings' energy efficiency. By introducing SEEF, this paper demonstrates the value of using frontier methods to rank buildings' energy efficiency and improves the way the energy efficient frontier is defined. The proposed SEEF methodology extends standard DEA analysis in energy efficiency field, by incorporating uncertainty to the estimation of the frontier as well as introducing a data-driven method to support the choice of frontier shape. SEEF consists of three stages: (1) set up the model by choosing the frontier shape and forming the LPP; (2) run Algorithm 1 to estimate efficiency scores and their confidence intervals; and (3) benchmark buildings by comparing their ranks. We describe each step in the rest of this section.

3.1 Choice of Frontier Shape

To validate the choice of frontier shape, we use the following hypothesis setup:

> Null hypothesis H_0: frontier is a straight line (representing a Constant Returns to Scale—CRS)
>
> Alternative hypothesis H_1: frontier is piecewise linear convex (representing a Variable Returns to Scale—VRS)

We use the following test statistic to reject or accept the null hypothesis H_0:

$$R = \frac{\sum_{k=1}^{K} \theta_{CRS}^{k}}{\sum_{k=1}^{K} \theta_{VRS}^{k}}$$

Where θ_{CRS}^{k} is the efficiency estimate of building k as calculated with respect to a CRS frontier, and similarly for VRS. If H_0 is true, then R will be close to 1, otherwise R < 1. Since we do not know the distribution of R a priori, we find its empirical distribution from the data set (see Algorithm 1 for details). Based on the empirical distribution of R, we define a critical value for the null hypothesis and check the statistic R against it.

3.2 Stochastic Energy Efficiency Frontier (SEEF) Algorithm

Once the frontier shape is known, efficiency scores can be calculated by solving LPP 1 (S,X). Before explaining SEEF method, we establish some assumptions:

(a) Energy consumption is stochastic. It is best explained by a random variable whose variance can be (partially) explained using variables such as climate zone and type of building (Kavousian et al. 2012).

(b) Service level values are deterministic. For example, the floor area of a house and the size of the family are not stochastic.

(c) The building sample adequately represents the population of interest; i.e., the sample selection error is not significant.

Assumptions a, b, and c above imply that the energy efficiency scores are also random variables, not deterministic values. Furthermore, the only source of uncertainty in efficiency score is the stochasticity of energy consumption measurements. As a result, the energy efficiency of two or more buildings cannot be compared using just a single value of efficiency. Algorithm 1 explains how uncertainty of energy consumption can be incorporated into the efficiency estimates.

Algorithm 1 – Stochastic Energy Efficiency Frontier (SEEF): Estimating the bias-corrected and confidence interval for efficiency scores based on repeated measurements.

Repeat for b=1,2,…,B (recommended B > 1000)

For each building, calculate x_k^b for k=1,2,…,K as follows:

Draw a random value from the Empirical Cumulative Distribution Function (ECDF) of energy consumption measurements (Figure 9).

a. Form the ECDF
b. Draw a random number from U[0,1]
c. Interpolate between the two measurements that correspond to the ECDF value closest to the random number.

$$if\ \frac{i}{n} \le u < \frac{i+1}{n} \qquad \text{then} \qquad x_k^b = n\,[x_i\left(u - \frac{i}{n}\right) + x_{i+1}\left(\frac{i+1}{n} - u\right)]$$

Where n is the number of measurements. For definition of other variables, see Figure 9.

Use *LPP1(S,X)* to calculate efficiency scores for each building. For the case of one input (e.g., electricity consumption) and one service level measure with finite number of values (*j=1,2,…,J)*, the solution is:

$$\theta_i^b = \frac{x_j^{b*}}{x_i^b} \qquad \text{where} \qquad x_j^{b*} = \min_k\{x_k^b : S(x_k^b) = S_j\}$$

S_j is the j-th level of service observed in the population, and $S(x_k^b)$ is the service level of the building k.

Using the efficiency values of each iteration ($\theta_i^1, \theta_i^2, \ldots, \theta_i^b$), calculate:

Bias-corrected efficiency score:

$\tilde{\theta}_i = E[\theta_i^b]$, *i=1,2,…,K; b=1,2,…,B*

Confidence interval for efficiency scores:

$\tilde{\theta}_i \pm sd(\tilde{\theta}_i)$ *i=1,2,…,K; b=1,2,…,B*

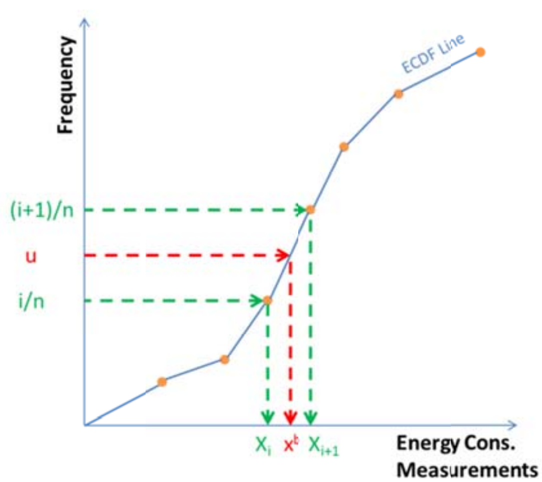


Figure 9 - Random selection of a measurement.

To illustrate the algorithm with an example, consider two buildings A and B, each with 90 days of energy consumption data. These 90 values can be presented with a random variable distribution; i.e., each is a realization of the energy consumption random variable. We start with randomly selecting a consumption value for each user. But instead of directly selecting one of the 90 values, we form an ECDF of observed consumption values for each user, then randomly select a probability on the ECDF, and finally find corresponding consumption value by interpolating between observe values (see Figure 9). Once the energy consumption values of the two buildings are selected, the energy efficient frontier can be formed. In this case, the building with a lower energy consumption over LOS ratio defines the frontier. Having specified the efficient frontier, we can find the efficiency values of each building using the formula presented in Figure 4. This process should be repeated several times (for the purpose of this example assume 1000 iterations), and the efficiency values at each iteration must be recorded. At the end of the iterations, the efficiency scores of each building form a distribution (in this case we have two distributions, each formed using 1000 values—one for each iteration). The mean of each distribution is the bias-corrected efficiency score, and the 95% confidence interval is the 95% confidence interval of the efficiency score.

3.3 Stochastic Benchmarking of Buildings Using SEEF

SEEF returns an efficient frontier and an estimate of the uncertainty in the frontier. It forms the distribution of efficiency score of each building, from which the bias-corrected score and its confidence interval are estimated. The efficiency score distribution can also be used to find a stochastic ranking of buildings: In each iteration b of Algorithm 1, record rank $q(b,k)$ of building k among all buildings. Then form ECDF of $q(b,k)$ to provide a distribution of ranks for building k. For example, the ECDF can be used to determine a 95% confidence interval for the rank of building k. Including reliability in rankings allows the analyst to make more informed comparisons.

4 Experimental Data and Model Setup

To illustrate the application of SEEF, we applied the method to a 10-min interval smart meter data set of 307 households in the US, collected between June and September 2010. A smart meter is a device that records electricity consumption of a building in regular intervals (every 10 minutes in this case) and communicates that information to a server where it is saved for later retrieval. The consumption data is accompanied by a survey of household information, including demographic features, building characteristics, and appliance ownership. The participants are spread across 164 zip codes and cover six of eight climate zones in the US. Climate zones in the data set include Cold, Hot-Dry, Hot-Humid, Marine, Mixed-Dry, and Mixed-Humid. The two excluded data sets are Very Cold and Sub-Arctic climate zones, the former only occurs in a few places within the US and the latter only occurs in Alaska. Climate zones are defined with reference to Building America definitions, and are based on annual precipitation and temperature records (PNNL, 2010). Building types of 307 subjects in our data set include apartment units, townhouses, and detached buildings. Participants were employees of a technology company in Silicon Valley and voluntarily participated in the experiment. They all paid for their own electricity bill.

More details on the data collection process can be found in Kavousian et al. (2012) and Houde et al. (2012).

The next section describes how we defined a measure of level of service for each house based on the information available in the survey.

4.1 Choice of Level of Service

Quantifying the utility that users receive from consuming energy is easier in some contexts compared to others. For example, for a manufacturing plant, the utility can be defined as the number of products that the plant produces. For a retail store, the utility can be defined as the number of customers it serves (Phylipsen et al, 1997). However, in a household model, quantifying the utility that a household gets from using energy is still an open question among economists, behavioral economists, sociologists, and psychologists (Lutzenhiser et al, 2010). This paper uses a technical approach (Lutzenhiser et al, 2010) to calculate a measure that is correlated with the service that users receive from electricity consumption. This measure is defined as the aggregate expected consumption of the stock of appliances, and is called Appliance Ownership Index (AOI) (Capasso et al, 1994). A household's AOI is defined as a linear combination of the number of most important appliances, with coefficients of the linear combination being the industry average (or model-based estimates) of the electricity consumption of each piece of appliance. The AOI is normalized (setting the maximum value to one) to form LOS. In other words, the LOS of user u, denoted by LOS_u is calculated using the following formula:

$$AOI_u = \sum_{i=1}^{k} A_k \times N_{k,u}$$

$$LOS_u = \frac{AOI_u}{Max(AOI)}$$

Where A_k is the average electricity consumption for appliance k, and $N_{k,u}$ is the count of appliance k that user u owns.

Choice of LOS is context-specific and depends on the purpose of energy efficiency studies. LOS is further discussed later in this paper and a sensitivity analysis is performed to examine the effect of the choice of level of service metric on energy efficiency estimates. But first in the following sections we demonstrate the power of SEEF model in informing energy efficiency programs by using Appliance Ownership Index as the LOS definition.

4.2 Weather Normalization

Some of the uncertainty in energy consumption of a building is due to local weather variations. If the weather data is available, it is important to normalize consumption values for weather impact. An examination of the relationship between energy consumption and outside temperature in our sample shows significant correlation for temperature lags of up to 24 hours. To remove this effect, we fit a stepwise regression model on outside temperature including lags of up to 24 hours.

$$y_i = a_i + \sum_j b_{ij} T_{ij} + \varepsilon_i \qquad i=1,2,\ldots,K; \quad j=0,1,2,\ldots,24$$

Where T_{ij}'s are the outside temperature readings for building i recorded j samples ago.

The estimated coefficients $\hat{b}_{ij}'s$ are then used to calculate normalized energy consumption, by assigning a reference temperature (65'F) to all households, and calculating their hourly consumption using the regression model.

$$\tilde{y}_i = a_i + \sum_j \hat{b}_{ij}\tilde{T}_{ij} + \varepsilon_i \quad i=1,2,\ldots,K; \quad j=1,2,\ldots,24; \quad \tilde{T}_{ij} = 65F \quad \forall i,j$$

4.3 SEEF Frontier Shape Choice

The hypothesis test for the shape of the frontier for this data set suggests a non-linear frontier ($\hat{R}$: 0.64; R_{crit}=0.91). Figure 10 illustrates two innovations of SEEF: Figure 10a compares linear frontier with curved frontier. Existing applications of DEA in energy efficiency use the linear frontier (dashed red line); but as the figure shows, using a linear frontier results in under-estimation of the energy efficiency of high energy users. SEEF offers a method to choose the functional form of the frontier based on measured data. Figure 10b shows the bias-corrected frontier and its confidence intervals. Existing applications of DEA in energy efficiency treat energy consumption as a deterministic value, or resample values from a fixed set of measurements to simulate fluctuation in daily energy consumption. SEEF offers a method to use repeated measurements of energy consumption to estimate bias-corrected and confidence intervals for the efficient frontier.

Note that bias-corrected frontier does not completely envelop data points as a deterministic frontier would do. This is due to the stochasticity assumption in estimating bias-corrected efficient frontier. However, the efficiency values are still strictly between zero and one, as explained here: at any given iteration, a random consumption value for each user is selected, and the efficient frontier is formed using the collection of those random consumption values. Therefore, at any given iteration, the frontier envelops all points, and thus the efficiency values calculated at any given iteration are between zero and one. Subsequently, the bias-corrected efficiency values that are averages of efficiency values of all iterations are still between zero and one. On the other hand, each data point on the plot represents the average consumption value which may or may not be enveloped completely by the bias-corrected frontier. In short, the seeming discrepancy of efficiency values higher than one is only due to the way the points and the frontiers are illustrated here and the actual efficiency values are still between zero and one.

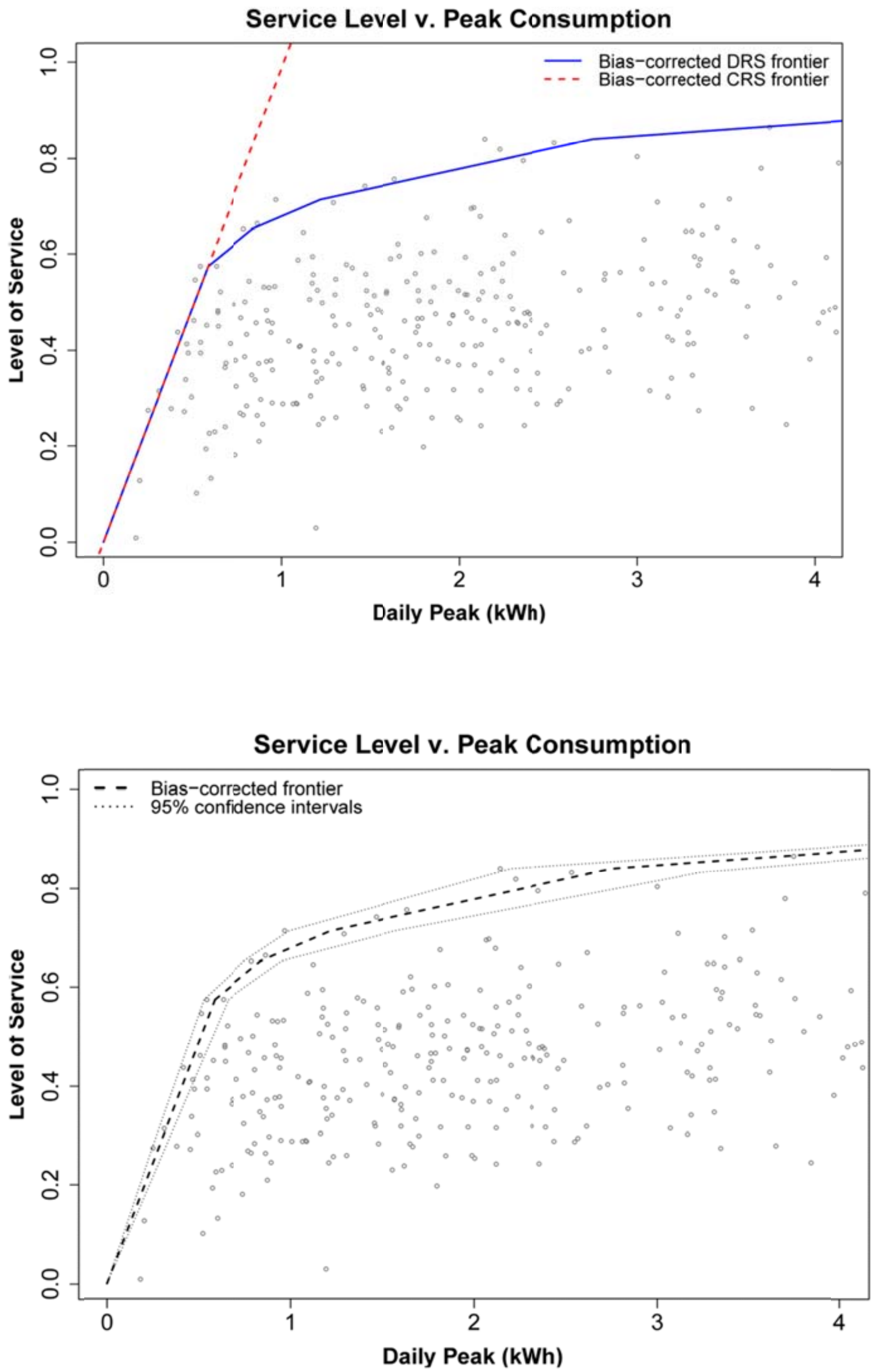


Figure 10 - Two innovations of SEEF: (a) Top: choice of frontier form; (b) Bottom: bias-corrected stochastic frontier and its confidence intervals. The points on the plots represent daily peak consumption of users (averaged over 92 days of the experiment period) versus Level of Service.

5 Results and Discussion

SEEF results can be used in a variety of applications, including ranking buildings based on their energy efficiency, comparing sub-populations of buildings in terms of their energy efficiency, and identifying irregularity in energy consumption patterns of buildings compared to the overall population. The following sections give examples of each analysis.

5.1 Rank Comparison

Figure 11 offers a visual aid to compare the efficiency scores of any two buildings. The less the two confidence intervals of the buildings overlap, the higher is the degree of confidence that one building is more efficient than the other one. As explained in previous sections, ECDF of the efficiency scores can be used to find the confidence level of this comparison.

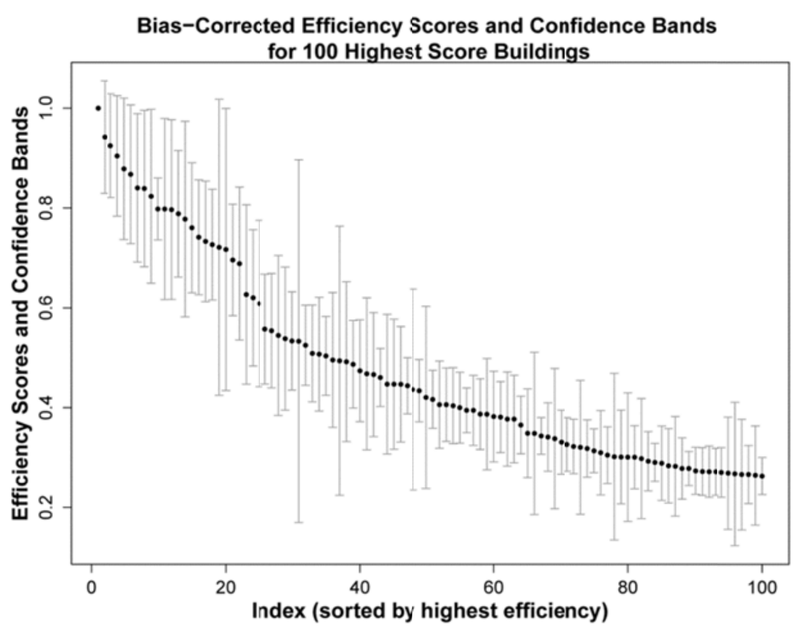


Figure 11 - Bias-corrected and confidence bands for efficiency scores of 100 most efficient buildings, as calculated by SEEF.

For example, in Figure 11, after 15th building, the efficiency scores drop sharply. For reference, we label the first 15 buildings as the *efficient buildings subset* (or *efficient homes*) and similarly call the rest of the population *typical buildings*. The choice of efficient set depends on the goals of the energy efficiency analysis; for instance one can label the first 7 buildings as efficient, because the confidence intervals of their efficiency values are all higher than 0.6 on efficiency scale. Once the efficient set is identified, the characteristics of its buildings can be compared with the rest of the population to infer insights.

For example, in our data set, whereas 65% of typical buildings resided in a Marine climate zone, this ratio was 73% within the efficient buildings subset. Since we have normalized the consumption data for outside temperature, the higher ratio of marine-climate residences in the efficient homes group suggests the effect of other factors that correlate with mild climate, such as lower penetration of air conditioners or design based on natural ventilation.

Newer buildings are more strongly represented in efficient buildings compared with the rest of the population: 40% of efficient homes are built after 2004, whereas only 18% of the typical buildings are built after 2004. On the other hand, while 45% of typical buildings are built before 1975, this ratio is 33% in efficient homes.

Efficient houses in our data set also illustrated higher penetration for some energy-efficient features. For example, 80% of efficient homes had installed double-pane windows, compared with 68% in the rest of the population. Similarly, 60% of efficient homes declared using energy-start air conditioners, while this ratio was 43% in typical homes.

Table 2 compares appliance ownership for a select group of appliances among efficient homes with that of the overall population. Some items such as electric water heaters, electric central heaters, and electric oven have higher penetration in efficient homes than the rest of the population. On the other hand, some items such as dishwashers, electric heaters, second refrigerators, and second or third AC's are less common in efficient houses. One possible reason for this observation is that having high-consumption items such as water heaters and central heaters will give some buildings extra "allowance" for energy consumption, because of the way we defined efficiency. In other words, having high-consumption items increases the service level of the building, which enables the building to use more energy while be considered as efficient. This behavior of the frontier method has two implications for practitioners. First, it suggests that efficiency is a relative term that is defined in the context of the service that a building offers. Frontier methods such as SEEF enable a more balanced view of efficiency by comparing buildings of the similar levels of service with each other rather than comparing the entire stock of buildings at once. Second, the effect of level of service measure on energy efficiency scores emphasizes the importance of choosing an appropriate level of service measure and performing sensitivity analysis before making judgments about energy efficiency of a building. We will discuss this insight later in the paper.

Table 7 - Appliance ownership in efficient homes versus the overall population.

Appliance Ownership	Penetration in Homes Labeled Efficient	Penetration in the Rest of the Population
Electric Water Heaters	27%	13%
Electric Central Heater	40%	20%
Dish Washer	66%	92%
Electric Heater	8%	19%
Electric Oven	87%	55%
Second Refrigerator	13%	20%
More than one AC	6%	14%

Another important application of the efficiency ranks is to prioritize and plan for energy efficiency programs (i.e., benchmarking). Buildings with lower efficiency scores should be given priority in efficiency programs.

5.2 Comparing Energy Efficiency of Buildings in Different Sub-Populations

The efficient frontier specifies the best-practice technology that a population of buildings can achieve. Therefore, by comparing efficient frontiers of sub-populations of buildings, one can identify which sub-population can achieve lower levels of energy consumption for the same level of service. Such insight can be used in several applications, including comparing the potential for energy efficiency in different populations of buildings and optimizing design practices to design a building with more energy efficient features.

Figure 12 shows an example of such comparison, where efficient frontiers for buildings in different climate zones are specified. The order of frontiers from left to right suggests that in our data set buildings with marine and hot-humid climate zones showed the highest efficiency values, followed by mixed-dry and mixed-humid climate zones.

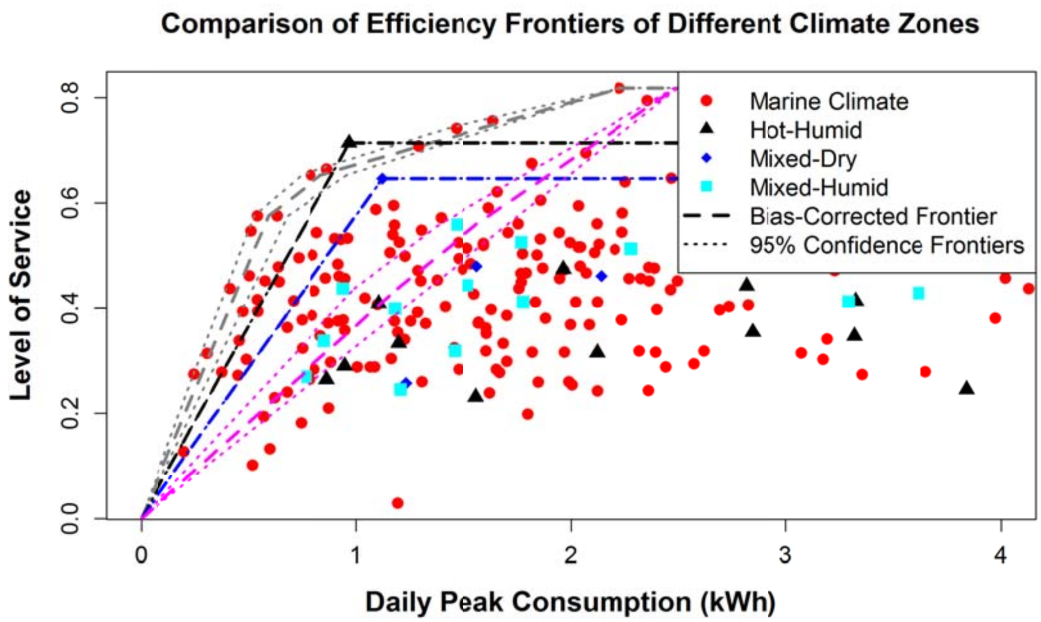


Figure 12 - Comparison of efficiency frontiers of sub-populations. Frontiers to the left imply higher efficiency for best-practice technology of their respective sub-population.

5.3 Time-of-Use Analysis

Comparing energy efficiency scores based on whole-day consumption may obscure hourly patterns of consumption in buildings. Time-of-Use (TOU) analysis can reveal buildings that are efficient during certain hours; e.g., peak hours. Energy efficiency planners or anyone also need this distinction because a building that is inefficient during afternoon hours may benefit from a more efficient air conditioner unit, whereas a building that is highly inefficient during evening hours may need a new, more efficient TV or water heater.

Figure 13 below compares the energy efficiency scores of buildings when estimated based on the consumption during peak hours versus off-peak hours. We define the peak versus off-peak schedule as follows: weekdays from 1pm to 7pm are considered peak hours; weekdays 10am-1pm and weekends 5pm-8pm are considered part-peak hours; and the rest of the week hours are considered off-peak (Pacific Gas & Electric, 2012). Most buildings stay close to the unit slope line, which implies that buildings that are efficient when measured based on their off-peak-hour consumption are also efficient when measured by their peak-hour consumption.

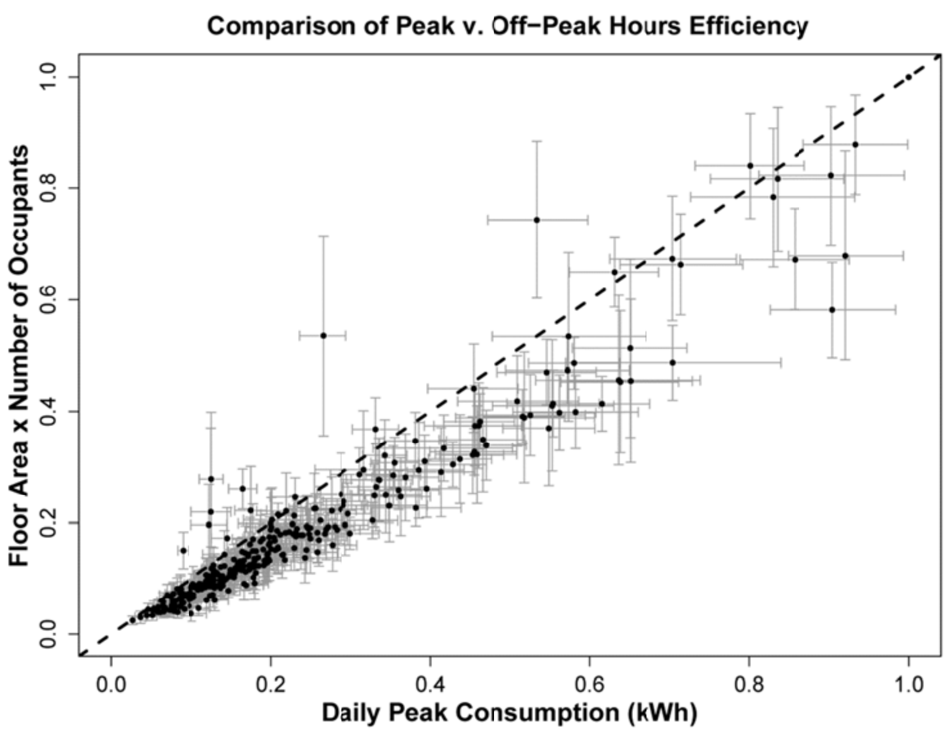


Figure 13 - Comparison of efficiency values when estimated based on consumption during peak hours versus off-peak hours.

5.4 Comparing SEEF with Other Benchmarking Methods

A common method for measuring energy efficiency of buildings is comparing energy intensity (energy per square foot) (Schipper, 1983). The energy intensity method is a fast and simple way for preliminary comparison among buildings of similar size and characteristics. However, it is too simplistic because it assumes the entire population of buildings (with their different types, sizes, and occupancy level, etc.) can be compared universally and with only one metric. In frontier methods terms, energy intensity method is comparable to a frontier model with square footage as the LOS and a linear frontier. However, research has shown that buildings are not linear systems, and as buildings grow in size, the processes with which they consume energy changes (Kavousian et al. 2012). For instance, the volume of the house may increase with a different rate than the square footage; or, the penetration of certain equipment such as pools will change as the size changes.

SEEF solves this problem by comparing the energy consumption of buildings of the same service level, without making any assumptions about linearity of buildings' energy consumption. Thus, when the efficient frontier is non-linear, the results of intensity method will be different than those of DEA (Figure 14).

To make the case more clear, we compared the efficient buildings as measured by the intensity method with the 15 most efficient homes as measured by SEEF. As Figure 14 shows, the energy intensity method has a bias towards buildings with lower floor area. On the other hand, SEEF identifies efficient buildings across the entire range of building sizes and energy consumption values. Here, the discrepancy between SEEF and energy intensity method is due to the non-linearity of the efficient frontier.

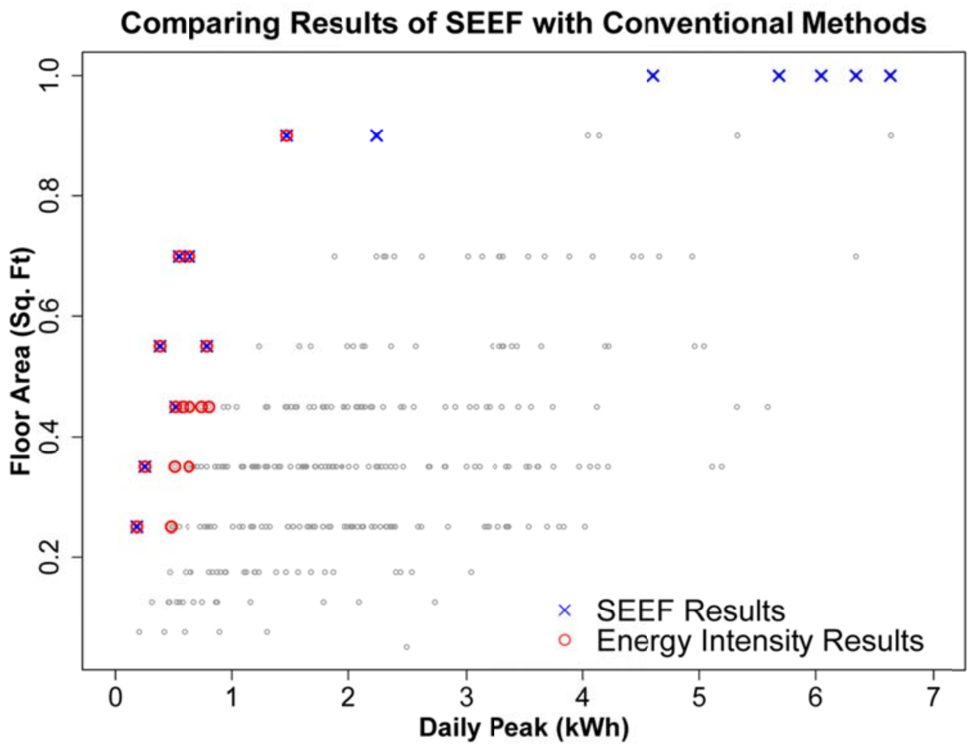


Figure 14 - Compared with energy intensity, SEEF identifies efficient buildings over a larger scope of sizes. In the plot above, energy intensity has a bias towards smaller buildings. However, since SEEF compares buildings within their own class of floor area, it is still able to find the most efficient buildings that are larger in size.

Some of the houses that were not identified as efficient by the energy intensity method indeed featured several energy efficiency attributes. For example, one of such houses has full insulation including walls, doors, ceiling, basement and window insulation, along with double-pane windows and a programmable thermostat, and consumes the lowest amount of energy among its peer houses. Still, that house was designated as inefficient by the energy intensity method. This is an example of a type I error in identifying inefficient buildings: declaring a building as inefficient when in fact it does not have significant potential for improvement. Type I errors lead to inefficient allocation of resources when planning energy efficiency programs.

5.5 Discussion on Level of Service (LOS)

The choice of Level of Service (LOS) impacts the results of SEEF. For instance, Figure 15 shows a comparison of efficiency scores of our data set when calculated using different LOS definitions. The upper-right part of the plot shows the correlation between efficiency scores of buildings when measured using different LOS's. The bottom-left part of the plot visually shows this relationship: the horizontal and vertical axes on each plot represent efficiency scores based on one LOS, and the local regression line shows the relationship of the efficiency scores. Finally, the plots on the diameter show histograms of efficiency scores with respect to each LOS definition.

The correlations of efficiency values range from 25% to around 63%, suggesting that efficiency scores of a building, when measured with respect to different LOS's, can be very different. In other words, different metrics encourage different behaviors. For example, choosing energy per square foot favors large buildings with less number of occupants, but tells very little about how efficiently the occupants use energy. For an energy efficiency planner, it is important to choose the right LOS definition to encourage the right behavior. Once the LOS is specified, the data requirements to benchmark the buildings become clear, too. In this paper, the goal is to develop a novel method to quantify energy efficiency once the LOS is defined. We chose the appliance ownership index as an example to illustrate the capabilities of SEEF. We also demonstrate the importance of the choice of LOS by comparing efficiency scores calculated using different LOS definitions. However, we argue that the choice of LOS is context-specific and needs to be made based on the goals of the energy efficiency analysis.

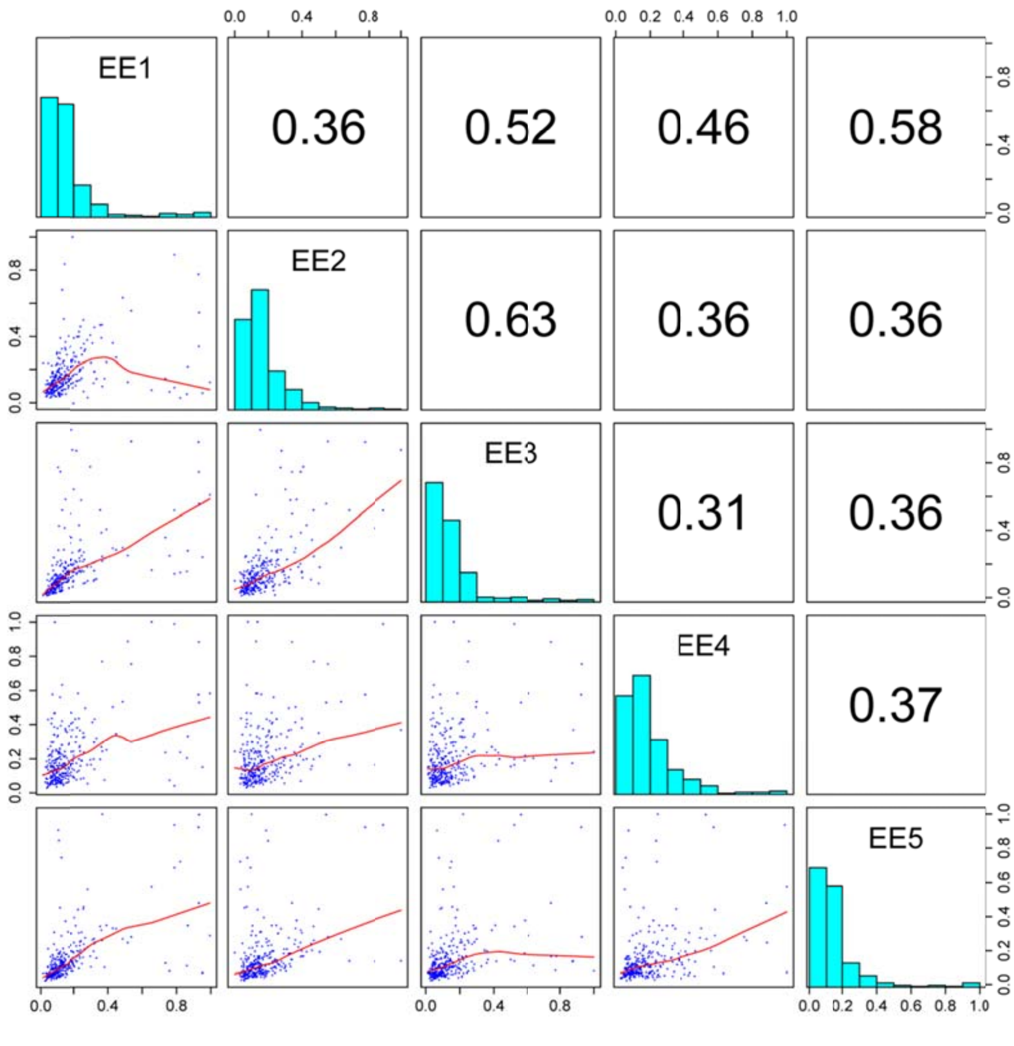


Figure 15 - The choice of Level of Service has considerable impact on efficiency scores. The upper-right part of the plot shows the correlation between efficiency scores of buildings when measured using different LOS's. The bottom-left part of the plot visually shows this relationship. The LOS definitions for each energy efficiency metric are as

follows: EE1: floor area; EE2: number of occupants; EE3: floor area and number of occupants; EE4: appliance ownership index; EE5: building characteristics (e.g., insulation level, type of building, etc.).

6 Conclusions

Frontier methods quantify the energy efficiency of buildings by forming an efficient frontier (i.e., best-practice technology) and comparing all buildings against that frontier. The efficient frontier specifies the lowest energy consumption observed at any Level of Service (LOS), defined as the utility that the users receive from using energy. Because energy consumption fluctuates with time, energy efficiency is a stochastic variable rather than a deterministic value. Thus, one needs to estimate the uncertainty of this stochastic variable when benchmarking building energy efficiency.

In this paper, we developed Stochastic Energy Efficiency Frontier (SEEF) method to find bias-corrected estimates and confidence intervals for energy efficiency of buildings based on repeated observations of energy consumption. We also presented a method for to rank buildings based on SEEF results. We showed how the functional form of the frontier can impact analysis results, and developed a method to choose the functional form based on the data. We also developed a method to calculate and compare the stochastic energy efficiency score, and introduced a ranking algorithm based on the stochastic efficiency scores.

Using a smart meter data set of 307 residential buildings in the US, we demonstrated how SEEF can be used to compare the energy efficiency of different populations at different times. To illustrate the method, we used the metric *Appliance Ownership Index* (a linear combination of the number of appliances, with coefficients proportional to the average consumption of each piece of appliance) as the LOS. The results of SEEF (efficiency scores and ranks) can be used in several follow-up analyses, including: comparing different sub-populations of buildings by comparing their efficient frontier and comparing the efficiency scores of buildings in different time-of-use periods to identify irregular behavior.

SEEF performed better than the energy intensity method (comparing kWh / sq. ft) for benchmarking energy efficiency of the smart meter data set. While SEEF identified efficient buildings across the entire spectrum of building sizes and types, the energy intensity method showed bias towards smaller buildings. The discrepancy between the results of SEEF and energy intensity method can be attributed to the non-linearity of buildings; i.e., the same ratio of kWh/Sq.Ft could not be used as a benchmark to compare buildings of all sizes and types. SEEF addresses the non-linearity of buildings by using a non-parametric frontier that is defined at each service level as the benchmark.

Stochastic Energy Efficiency Frontier (SEEF) offers an improvement over existing frontier methods to benchmark the energy efficiency of buildings. In sum, these advantages include:

- SEEF recognizes the stochasticity of energy consumption, and treats energy efficiency as a random variable. By doing so, SEEF finds bias-corrected efficiency estimates and their respective confidence intervals, which can then be used to compare the efficiency of different buildings.

- The ranking process identifies the probability distribution of the efficiency ranks of buildings. This is an improvement over the existing methods, which simply compare energy efficiency scores of buildings and deterministically declare one building more efficient than the other.
- SEEF uses actual data (i.e., repeated measurements) to estimate uncertainty of efficiency scores.

In short, the contributions of this paper to the field of energy efficiency include:

a) Calling for a new way to quantify efficiency based on input/output relationship;
b) Demonstrating the value of frontier methods in energy efficiency analysis;
c) Improving the way the efficient frontier is defined, by:
   I. formalizing a statistical method for choosing the functional form of the efficient frontier;
   II. developing a method to estimate the sensitivity of efficiency estimates to temporal variation in energy consumption using measured data as opposed to resampling methods (i.e., stochastic non-parametric frontier instead of deterministic);
   III. developing SEEF algorithm to rank buildings based on their energy efficiency given the Level of Service (LOS) and time-series data of energy consumption.

While this paper presented SEEF as a novel approach to benchmark building energy efficiency, the definition of Level of Service is still open to discussion. A possible follow-up study is to examine the results of energy efficiency ranking using different LOS definitions and compare the results critically. Since SEEF compares all buildings with the best-practice technology, it is prone to outliers. Thus, another direction for further development for SEEF is to devise a method to identify and remove outliers from the data set. We plan to develop a robust statistical ranking algorithm for buildings based on SEEF. As the first step, we are analyzing the performance of SEEF on a large data set of energy consumption. The results of this research can be used to improve the way we benchmark buildings, and ultimately, improve resource allocation for energy efficiency programs.

7 Acknowledgement

Amir Kavousian was funded through Advanced Research Projects Agency-Energy (ARPA-E) grant to Stanford University. Ram Rajagopal was supported by the Powell Foundation Fellowship and TomKat Center for Sustainable Energy. The authors would like to thank Professor Martin Fischer, Professor James Sweeney, Dr. June Flora, Dr. Carrie Armel, and Professor Pravin Varaiya for their helpful inputs and comments during this study. The authors would also like to thank Rene Morkos, Taewan Kim, Austin Becker, and Forest Peterson for their feedback on Level of Service discussion. The authors would also like to thank respected reviewers for their helpful comments on the paper.

8 References

Afriat, A. S. N. (1972). Efficiency Estimation of Production Functions. *International Economic Review*, *13*(3), 568-598.

Ahn, C., Lee, S., Peña-Mora, F., and AbouRizk, S. (2010) “Toward Environmentally Sustainable Construction Processes: The U.S. and Canada’s Perspective on Energy Consumption and GHG/CAP Emissions.” *Sustainability, Molecular Diversity Preservation International (MDPI)*, Basel, Switzerland, 2010, Vol. 2, No.1, pp. 354-370.

Aigner, D.J.; Lovell, C.A.K.; Schmidt, P. (1977) “Formulation and estimation of stochastic frontier production functions.” Journal of Econometrics, 6:21–37.

Atkinson, B. A. C. (1982). Regression Diagnostics , Transformations and Constructed Variables. *Journal of the Royal Statistical Society*, *44*(1), 1-36.

Azar, E. and Menassa, C. (2011). “Agent-Based Modelling of Occupants’ Impact on Energy Use in Commercial Buildings.” *Journal of Computing in Civil Engineering*, American Society of Civil Engineers. 26 (4), 506-518.

Berges, M., Goldman, E., Soibelman, L., Matthews, S. H., and Anderson, K., (2011) "User-centered Non-Intrusive Electricity Load Monitoring for Residential Buildings," *Journal of Computing in Civil Engineering*, vol. 25, iss. 1, 2011.

Carnes, D., Hunn, B., Jones, J. (1998) The Use of Data Envelopment Analysis for Evaluating Building Energy Consumption in Terms of Productivity. Proceedings of 1998 Summer Study in Building Energy Efficiency, Pacific Grove, CA.

Chen, J., Taylor, J. and Wei, H.H. (2012). "Modeling Building Occupant Network Energy Consumption Decision-making: The Interplay between Network Structure and Conservation," *Energy and Buildings*, 47: 515-524.

Clevenger, C., Haymaker, J., Ehrich, (2012) Design Exploration Assessment Methodology: Testing the Guidance of Design Processes, *Journal of Engineering Design.* ElectDOI:10.1080/09544828.2012.698256.

Farrell, M. J. (1957). The Measurement of Productive Efficiency. Journal of the Royal Statistical Society, 120(3), 253-290.

Goldstein, D. (2012) Personal communication.

Griffith, B., Crawley, D. (2006). Methodology for Analyzing the Technical Potential for Energy Performance in the US Commercial Buildings Sector with Detailed Energy Modeling. *Proceedings of SimBuild 2006 Conference*, Cambridge, MI, August 2-4, 2006.

Grosskopf, S. (1986). The role of the reference technology in measuring productive efficiency. *The Economic Journal*, 96(382), 499-513.

Houde, S., Todd, A., Sudarshan, A., Flora, J. A., & Armel, C. K. (2012). Real-time Feedback and Electricity Consumption: A Field Experiment Assessing the Potential for Savings and Persistence. *Energy*, 1-22.

Kavousian, A., Rajagopal, R., Fischer, M. (2012) A Method to Disaggregate the Impact of Behavioral from Structural Factors in Energy Consumption, CIFE Working Paper, Center for Integrate Facility Engineering, Stanford, CA.

Lutzenhiser, L., Moezzi, M., Hungerford, D., Friedmann, R. (2010). Sticky points in modeling household energy consumption. *ACEEE Summer Study on Energy Efficiency in Buildings*, Pacific Grove, CA.

Maile, T. (2010) Comparing Measured and Simulated Building Energy Performance Data. *PhD Dissertation*, Civil and Environmental Engineering Department, Stanford University, CA.

Onut, S., Soner, S., 2006. Energy efficiency assessment for the Antalya Region hotels in Turkey. *Energy and Buildings* 38, 964–971.

Oude Lansink, A., Bezlepkin, I., 2003. The effect of heating technologies on CO2 and energy efficiency of Dutch greenhouse firms. *Journal of Environmental Management* 68, 73–82.

Oude Lansink, A., Silva, E., 2003. CO2 and energy efficiency of different heating technologies in the Dutch glasshouse industry. *Environmental and resource economics* 24, 395–407.

Pacific Northwest National Lab (PNNL) (2010). Guide to Determining Climate Region by County, Building America Best Practice Series Volume 7.1. Prepared for Department of Energy's Office of Energy Efficiency and Renewable Energy.

Patterson, M. (1996) What is Energy Efficiency? Concepts, Indicators, and Methodological Issues. *Energy Policy* 24:5, 377-390.

Pisello, A.L., Taylor, J., Xu, X. and Cotana, F. (2012). "Inter-building Effect: Simulating the Impact of a Network of Buildings on the Accuracy of Building Energy Performance Predictions," *Building and Environment*, 58: 37-45.

Riani, M., & Atkinson, A. C. (2000). Analysis : in Transformations Regression Diagnostic Data. *Technometrics*, *42*(4), 384-394.

Savage, S. (2009). *The Flaw of Averages: Why We Underestimate Risk in the Face of Uncertainty*. Hoboken, N.J.: John Wiley & Sons Inc., 2009. 412 pp.

Shephard, R. (1963). *Theory of Cost and Production Functions*. Princeton University Press.

Silverman, B. W., Young, G. A. (1987). The Bootstrap: to Smooth or Not to Smooth?. *Biometrica*, 74, 3, pp 469-479.

Simar, L., & Wilson, P. W. (1999). Some Problems with the Ferrier / Hirschberg Bootstrap Idea. *Journal of Productivity Analysis*, *80*(93), 67-80.

Waltner, M., Goldstein, D. (2012) Update on Labeling, Rating, and Benchmarking Efforts in the US and China: How to Increase Collaboration and Effectiveness. Proceedings of 2012 Summer Study in Building Energy Efficiency, Pacific Grove, CA.

Xydis, G., Koroneos, C., Moropoulou, A. “Energy and exergy utilisation in the Greek building sector”, Int. J. *Energy Technology and Policy*, 2008, Vol. 6, No. 5/6, pp.474–494, doi:10.1504/IJETP.2008.020736.

Yezioro, A., Dong, B., Leite, F. (2008). “An applied artificial intelligence approach towards assessing building performance simulation tools”. In: *Energy and Buildings*, 40(4), p. 612-620.

Chapter 4

Estimating Appliance Energy Efficiency in Residential Buildings Based on Smart Meter Interval Data – An Application of Stochastic Energy Efficiency Frontiers (SEEF)

Amir Kavousian, Ram Rajagopal, Martin Fischer

Keywords: Home Energy Efficiency, Appliance Energy Efficiency, Stochastic Energy Efficiency Frontier (SEEF), Smart Meter Interval Data

1 Abstract

Residential buildings account for about 37% of total electricity consumption in the US, with more than 34% of residential electricity consumed by appliances. The efficiency of appliance usage depends on both the technical efficiency of the appliance and on users' behavior. Utility monthly bills provide limited data to infer the appliance usage efficiency. Increased availability of smart meter data enables more detailed analysis of appliance data; however, computational methods to perform such analyses are still missing. This paper offers a new process to estimate the efficiency of appliance usage by individual users in a large population of buildings. This process is based on Stochastic Energy Efficiency Frontiers (SEEF), a non-parametric method to estimate energy efficiency and its confidence interval. Our process also accounts for weather and seasonal effects, as well as lifestyle effects.

We validated the method using a smart meter data set of 4231 households in Ireland, supported by a 120+ question survey. In this data set, families with kids who have

full-time employment and are highly-educated showed higher appliance usage efficiency compared to families with no kids, or families with retirees or unemployed members. When normalized for appliance ownership, families with kids also have relatively higher efficiency values due to enjoying economies of scale. When the respondents expressed doubt in being able to motivate other household members to save energy, the energy efficiency was lower. Households with tendency to make major lifestyle changes to save energy were also more efficient on average. Conversely, households who express small chances of making major lifestyle changes were more likely to be in the least efficient group of households. Our results emphasize the role of education (to a wide audience) in achieving energy efficiency goals. Households with high penetration of efficient lightbulbs and double-glazed windows showed higher appliance usage efficiency. Owned residences had higher efficiency scores, most likely due to higher penetration of efficient features.

2 Introduction

Residential buildings consume about 22% of total primary energy in the US (US EIA, 2014). Appliances (including cooking, electronics, refrigeration, wet cleaning, and computers) account for more than 20% of the energy consumed in residential buildings (US Buildings Data Book, US DOE, 2014). Utilities spend millions of dollars annually to improve appliance energy efficiency. For example, in California alone in 2013, utilities spent $80M on appliance and plug load efficiency programs, the highest expenditure among all utility energy efficiency programs (CPUC, 2013). Thanks to improved manufacturing standards, the efficiency of newly-manufactured appliances has improved significantly over the past few decades (NRDC, 2014). However, this improvement has not resulted in equal reduction in appliance energy consumption in home. Such discrepancy can be attributed to lower than expected penetration of efficient appliances, increased use of efficient appliances (rebound effect), or even keeping old appliances such as refrigerators after purchasing a new more efficient system. As a result, even across similar buildings, the energy consumption levels for similar appliance saturation levels are widely variable (Kavousian et al., 2013). When planning energy efficiency programs, it is important to identify houses with higher potential for appliance energy reduction.

Several methods exist for estimating appliance saturation and individual appliance end use consumption; the most widely known methods are: Sub-Metering (Knight et al., 2007), Non-Intrusive Load Monitoring (NILM) (Berges et al., 2008), and Conditional Demand Analysis (CDA) (Parti & Parti, 1980). Sub-metering involves installing separate energy meters for appliances, and monitoring their loads individually. While highly accurate, sub-metering requires an extensive metering infrastructure that renders it impractical for large-scale applications such as utility energy efficiency programs. One level lower in data-intensity than sub-metering, NILM analyzes high-frequency energy consumption data (often on the sub-second frequency), and provides estimates of each appliance energy consumption, in a process called disaggregation. While some NILM applications claim successfully estimating individual appliance load with accuracies of higher than 90%, the data requirements (sub-second load) and high computational power needed to analyze the data make NILM impractical with the existing data collection and processing infrastructure at the utility level. The least data-intensive method of the three, CDA estimates appliance ownership using regression analysis of monthly bills. Due to its low data requirements, CDA is the most practical option of the three methods introduced above; however, the accuracy of its estimations are low due to its use of aggregate load. There is need for a method that works best with 15-min interval energy consumption data that most utilities collect using their smart meter networks.

Smart meter data have made it possible to make inferences about energy consumption components and underlying processes (such as lifestyle choices) that lead to energy consumption (Beckel et al., 2013; Kwac et al., 2014; Albert & Rajagopal, 2013; Kavousian et al., 2013). What is missing is a framework that makes use of the latest advances in analyzing smart meter data to identify efficient and inefficient users, and specify the root causes of such efficiency and inefficiency.

This paper presents a method to infer appliance energy efficiency using smart meter data. Our proposed method is based on our previous work on Stochastic Energy Efficiency Frontiers (SEEF), a dynamic method to estimate energy efficiency using an input-output framework. We also use previous work on the estimation of appliance ownership and household activities from smart meter data to control for factors outside the influence of users in determining their energy efficiency. We are not aware of any previous work that has taken advantage of smart meter interval data to estimate energy efficiency of appliances, without the need for disaggregating load into its components. To illustrate the method in a real-world setting, we use a database of 4000 households in Ireland to estimate their appliance energy efficiency, and examine the factors that contribute to appliance energy efficiency in households. The following sections first introduce the experimental data, followed by the model setup and results. Finally, we examine the results in the context of existing knowledge and practice of appliance energy efficiency.

3 Experimental Data

A data set of more than 4000 households in Ireland was used to illustrate the building ranking method. The data set was collected by Ireland's Commission for Energy Regulation (CER), with the goal of assessing the technical feasibility and potential effects of smart meter implementation in Ireland (CER, 2011). The experimental data set includes electricity consumption (30-min interval) from July 20th, 2009 through December 26th, 2010 for a total of 525 days. In addition to consumption data, the data set includes a detailed list of dwelling characteristics, as well as demographic and socioeconomic data about participant households. The household data describe (a) building structural information, such as building type, insulation level, vintage, (b) buildings systems features including type of water heating, cooking fuel, percent efficient light bulbs, (c) appliances ownership, for instance number of TV sets, refrigerators, (d) demographic and socioeconomic status, such as number of people, age, socioeconomic status, and (e) behavioral features, including motivation for energy saving and energy efficient habits (see Table 8 for more details). Details of the data collection methodology can be found at McLoughlin et al. (2012). Table 8 summarizes appliance ownership data. Immersion heaters are electric water heating devices that heat up water through a heating element put in direct contact with water. They are used commonly in Ireland for fast water heating.

No information on the location of individual homes was released. However, our analysis of weather data from weather stations across Ireland shows that the weather conditions are reasonably consistent across the geographic areas. Thus, we assumed that weather is similar across all experiment subjects, and used the average readings from three weather stations across Ireland to estimate users' weather for our analyses.

Table 8 - Summary of appliance ownership on CER data.

Appliance	Count (total n = 4231)			
	0	1	2	3+
Washing Machine	51	3516	25	
Clothes Dryer	1136	2451	5	
Dish Washer	1161	2422	9	
Instant Hot Shower (Electric)	1085	2298	189	20
Shower Pump	2529	968	82	13
Cooker	820	2761	10	1
Electric Heater	2594	800	162	36
Freezer	1805	1718	65	4
Water Pump	2893	684	14	1
Immersion heaters	845	2736	11	
TV	34	919	1333	1306
Computer	1867	1620	13	5
Laptops	1638	1546	296	112
Gaming Consoles	2352	828	305	107

Energy is a random variable. Daily and hourly energy consumption values of a single user may change dramatically over time, due to several reasons. Therefore, a building's energy efficiency rank is also a random variable, fluctuating over time in response to changes in energy consumption of the building itself, and the buildings that are used as comparison benchmark. A fair ranking system controls for factors that are outside the control and interest of the users to change, and ranks buildings based on remaining energy consumption values.

For our purposes, we consider three sources for temporal fluctuation in energy consumption of a user over time: climate and seasonal effects, lifestyle effects, and random effects. Climate and seasonal effects are due to fluctuations in outside temperature, humidity, cloud cover, and so on, as well as the change in daylight hours and seasonal activities over different times of the year. Lifestyle effects are individual temporal patterns of energy consumption of users that are not explained by external variables such as a change in seasons or outside temperature. Lifestyle effects impact both timing and end uses of energy. For example, the lifestyle of a family of four with kids will be different from the lifestyle of a graduate student, and the difference in lifestyle will impact their energy consumption patterns. The family with kids may use energy during the early morning and early evening hours most of the days and primarily use energy to cook, heat up the space, and for lighting. The single person may have a late-night schedule most often, with long periods of unoccupancy, and primarily use electricity for computers, watching TV and game consoles.

Climate and seasonal effects are outside the control of the households and should not be include in energy efficiency ranking. For our purposes, we also do not include lifestyle effects in our energy efficiency ranking. Other studies with a behavioral or conservation focus might find lifestyle effects particularly interesting for energy reduction purposes. After removing the effects of climate, season, day of week, and lifestyle effects, we incorporate the remaining fluctuation of energy consumption in energy efficiency estimation. The next sections briefly explain controlling for climate and seasonal as well as lifestyle effects.

3.1. Controlling for Weather and Seasonal Effects

Consumption shows both weather and seasonal components. For instance, Figure 16 shows the relationship of daily consumption with outside temperature and day of week in our data set. While Figure 16 (a) shows the relationship of total daily consumption versus outside temperature averaged over all homes in the study, each individual home is different in the way that it is affected by outside temperature. Figure 16 (b) shows the distribution of correlation of outside temperature with daily total consumption for individual houses. As seen the correlations range from -0.8 (highly negatively correlated) to +0.4 (moderately positively correlated). Finally, as Figure 16 (c) shows, consumption also has a strong relationship with day of week—the total consumption on the weekends is on average 0.83 kWh higher than weekdays.

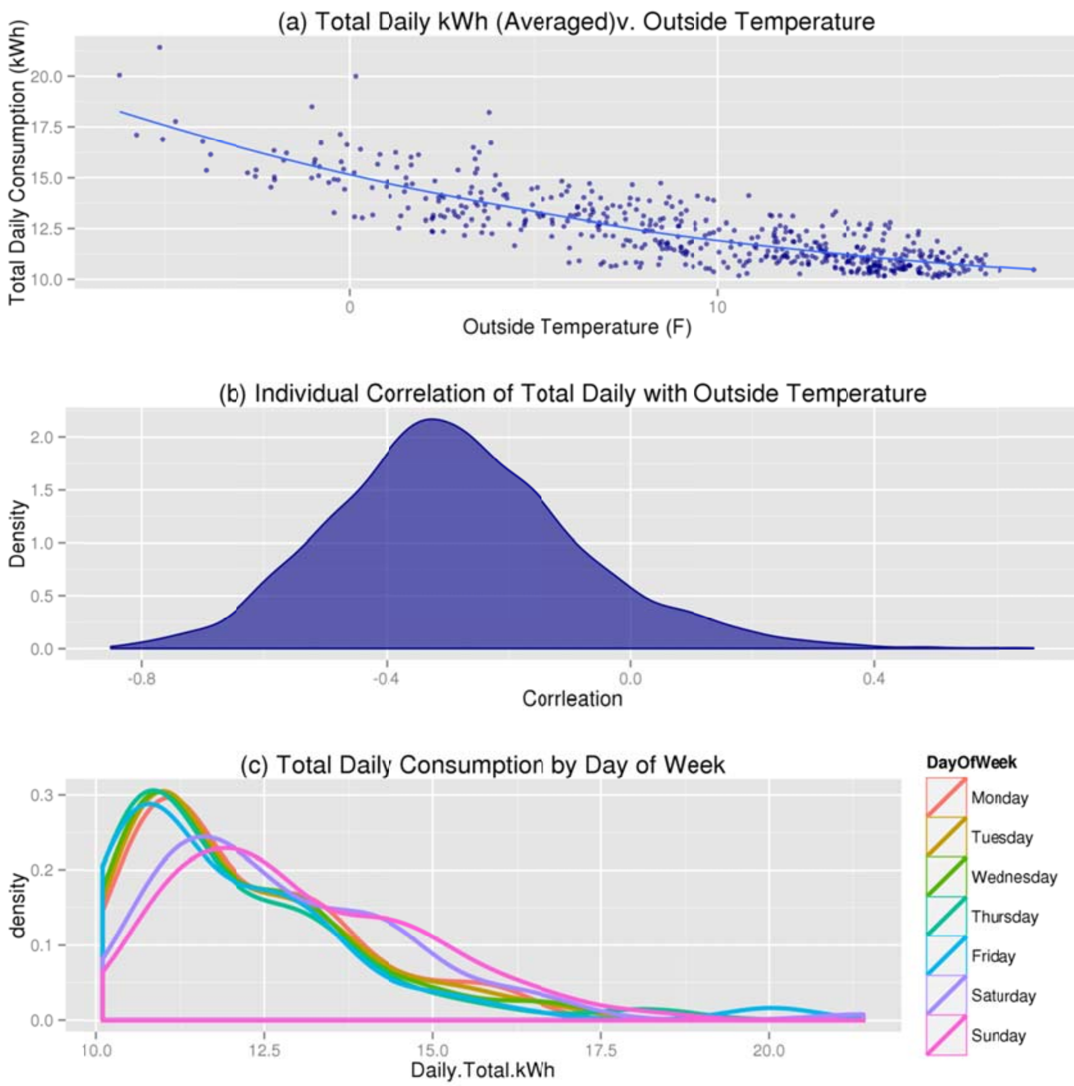


Figure 16 – (a) The relationship of daily consumption with outside temperature. Each dot represents one day of the experiment, with consumption values being averaged over all households. (b) The distribution of correlation values of daily total consumption with outside temperature for individual users. Each point on the density plots shows the frequency of a single household having certain correlation with outside temperature. In both plots a and b, temperature values are in Celsius. (c) Distribution of daily consumption values, broken down by day of week. None of the consumption data on these plots are normalized.

To account for weather and seasonal effects, an individual time-series regression model was fit to each house. The model has temperature, day of week, and month effects.

$y_i = a_i + \sum_j b_{ij}T_{ij} + \sum_j c_{ij}D_j + \sum_j d_{ij}M_j + \varepsilon_i$ *i=1,2,…,K buildings; j=0,1,2,…,525 days*

Where:

T_{ij}	outside temperature reading for building i at day j
D_j	day of week on day j of the data
M_j	month of year on day j

The estimated coefficients ($\hat{b}_{ij}$) are then used to calculate normalized energy consumption, by assigning a reference temperature (10'C) to all households, and calculating their hourly consumption using the regression model.

$\tilde{y}_i = a_i + \sum_j \hat{b}_{ij}\tilde{T}_{ij} + \varepsilon_i$ *i=1,2,…,K; j=1,2,…,24;* $\tilde{T}_{ij} = 65F \quad \forall i, j$

Figure 17 shows total daily consumption before and after normalizing for weather components.

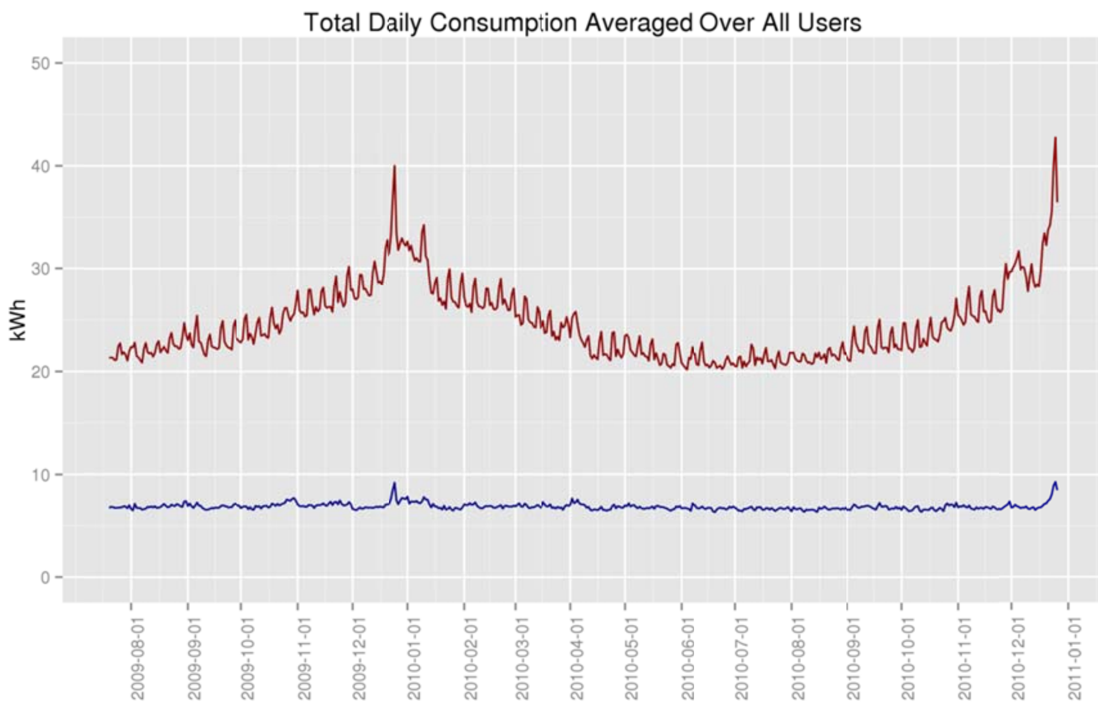


Figure 17 - Measured daily total consumption (averaged over all users) before and after normalizing for weather and seasonal effects.

The coefficient of determination (R2) of the temperature and seasonal sensitivity provides useful insights for energy efficiency planners. Especially, it is an indication of the ratio of the total load that goes towards electrical heating and cooling. A high value for R2 can be interpreted in several ways: (a) high heating and cooling load; (b) low non-heating and non-cooling loads, leading to a higher ratio of total load that goes towards heating and cooling; and (c) regularly and frequently-occupied house, resulting in a significant seasonal effect. In any of these cases, houses with high R2 values are good candidates for heating and cooling equipment retrofit, house sealing retrofits, or demand response programs.

Figure 18 shows the distribution of model R2, broken down by a few important household variables. As seen in the plots, outside temperature sensitivity is most impacted by building type and respondent age. This is expected since building type is a good proxy for heating and cooling load, and respondent age implies indoor air temperature preferences and lifestyle patterns.

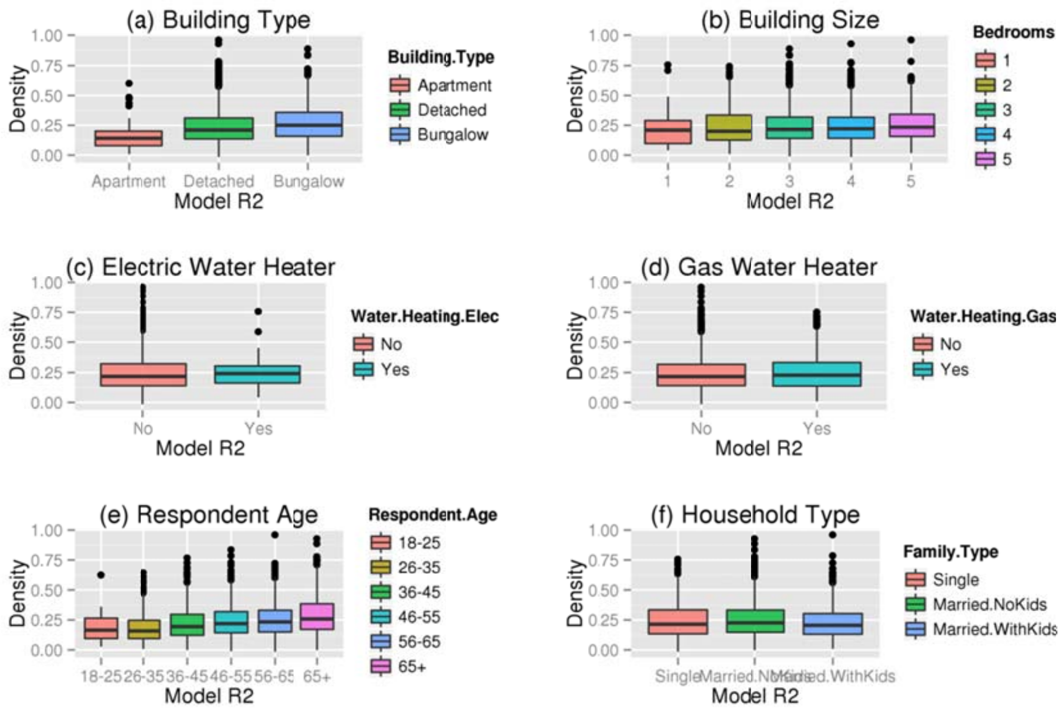


Figure 18 – Outside temperature sensitivity is most impacted by building type, respondent age, and water heating equipment.

A regression analysis of model R2s over household variables reveals more insights on the factors that contribute to temperature and seasonal effects. Using stepwise regression, we identified household variables with highest impact on temperature and seasonal sensitivity (as measured by R2 of the model explained above). Figure 19 shows these variables. This information can be used both in targeting users for energy efficiency programs, and for optimizing the message or incentive setup. For example, as Figure 19 shows, detached buildings and bugalows are more sensitive to outside temperature, while buildings with efficient features such as efficient lightbulbs and double-pane windows, or water heater insulation are less sensitive to outside temperature. Perhaps as a surprise, buildings that use renewable power for the water heater or have timer for their water heaters are more sensitive to outside temperature. This observation could be due to small non-heating or non-cooling loads; water heating is the second largest electricity consumer after space heating, and when the water heating load is removed, the effect of space heating is even more pronounced.

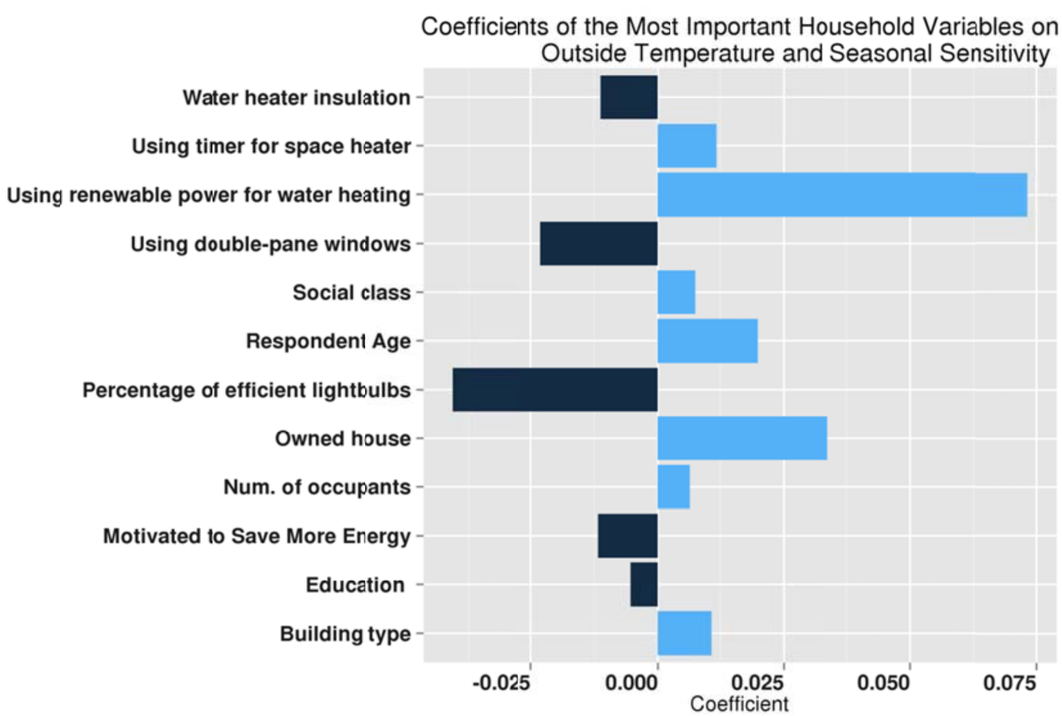


Figure 19 – Effect of various building characteristics on sensitivity to outside temperature and seasonal effects.

On the other hand, older households with higher income, residing in owned houses with larger families are more sensitive to outside temperature and seasonal effects, while higher education can reduce that sensitivity.

3.2. Controlling for Lifestyle Effects

Hourly load shapes (hourly consumption values of a single day of a single user) offer considerable information about daily activities of the user. For example, Figure 20 shows four different hourly load shapes of a user over four different days. The top-left plot belongs to a normal workday when the user wakes up in the morning, makes coffee, goes to work, comes back at around 6pm, makes dinner and watches TV, and goes back to bed. The other three plots in the figure infer different types of activities.

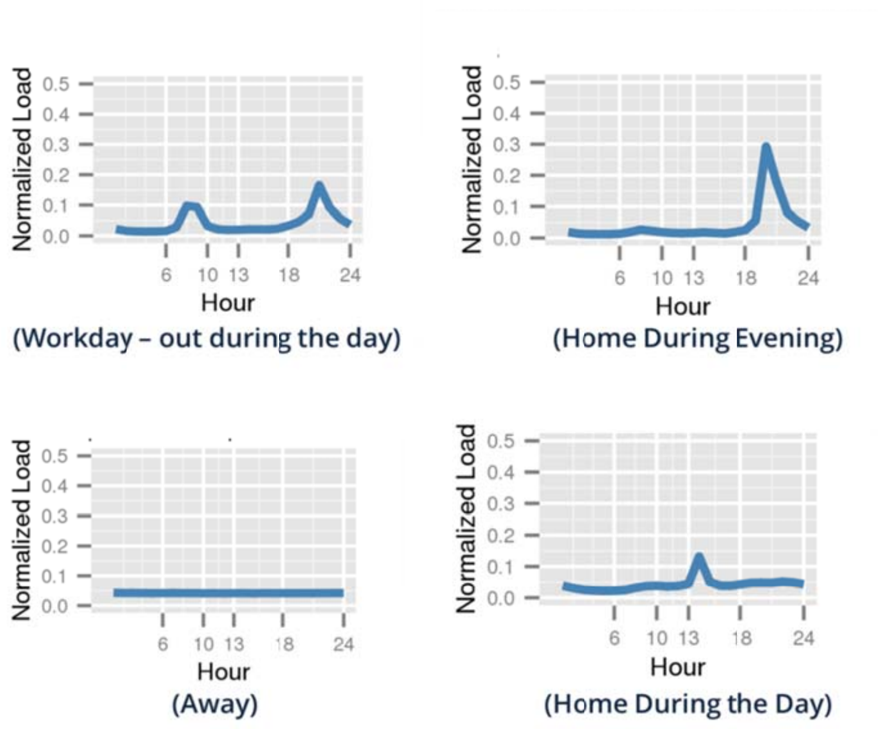


Figure 20 - Daily load shapes of a single user, over four different days.

In essence, each daily load shape acts as a “signature” of the daily activities of a user inside the house. These signature patterns can be used to label each day of each customer, using labels that correspond to the shape of the load; e.g., double-peak, evening-peak, unoccupied, etc. It has been shown that the entire load shapes of any user can be shown using a few recurring load shapes (Kwac at al., 2014). This enables us to compare energy consumption of users only within similar daily load shape clusters, effectively removing the lifestyle effects. For example, we can only compare double-peak and evening-peak days of customers against each other, knowing that these are the days when the customers are more likely to have used their appliances.

3.3. Incorporating Remaining Variation of Consumption in Energy Efficiency Estimates

After removing the effect of climate, seasonal effects, and lifestyle effects, we call the remaining energy component as "normalized" consumption. The fluctuations of normalized consumption can be incorporated in energy efficiency estimations. We use the bootstrapping algorithm explained in Kavousian & Rajagopal (2014) to estimate the confidence interval of efficiency scores based on the fluctuations of normalized consumption.

4 Model Setup for Appliance Efficiency Estimation

Efficiency is defined as "doing more using less". In a household context, the appliance ownership can be used as a proxy for "doing more"; i.e., a household that owns a washer and dryer is enjoying more utility (or as called in this paper, service) compared to a household without washer and dryer. Assuming that we are mainly interested in efficiency and not conservation, our approach to efficiency should be centered on keeping the same services in homes (e.g., washer/dryer, dishwasher, etc.), while minimizing the amount of energy that is spent on them. Therefore, we are interested in managing the tradeoff between the Level of Service (LOS, the sum of all services that a household enjoys) versus energy consumption (Kavousian & Rajagopal, 2014).

Frontier methods have been historically used to visualize and examine such tradeoffs, especially in settings with high degree of variability and uncertainty. For example, Figure 21 shows level of service (measured based on appliance ownership) versus energy consumption for the households of our case study. The energy consumption values are normalized for climate, seasonal, and lifestyle effects as explained in previous sections. As the plot shows, while energy consumption generally increases with increase in LOS, for any given LOS the range of energy consumption values is still wide. I.e., the energy-LOS plot is bounded on the left. This left boundary is the basis for specifying the energy efficiency frontier (Farrell, 1957; Carnes et al., 1998). The energy consumption of each user is then compared against the frontier as the benchmark: the closer a user is to the frontier, the higher its energy efficiency value is.

In the rest of this section, we explain how we used Stochastic Energy Efficiency Frontier (SEEF) to estimate appliance efficiencies of the households of our study. Later, we will explain the appliance efficiency scores using household data such as demographic and behavioral variables.

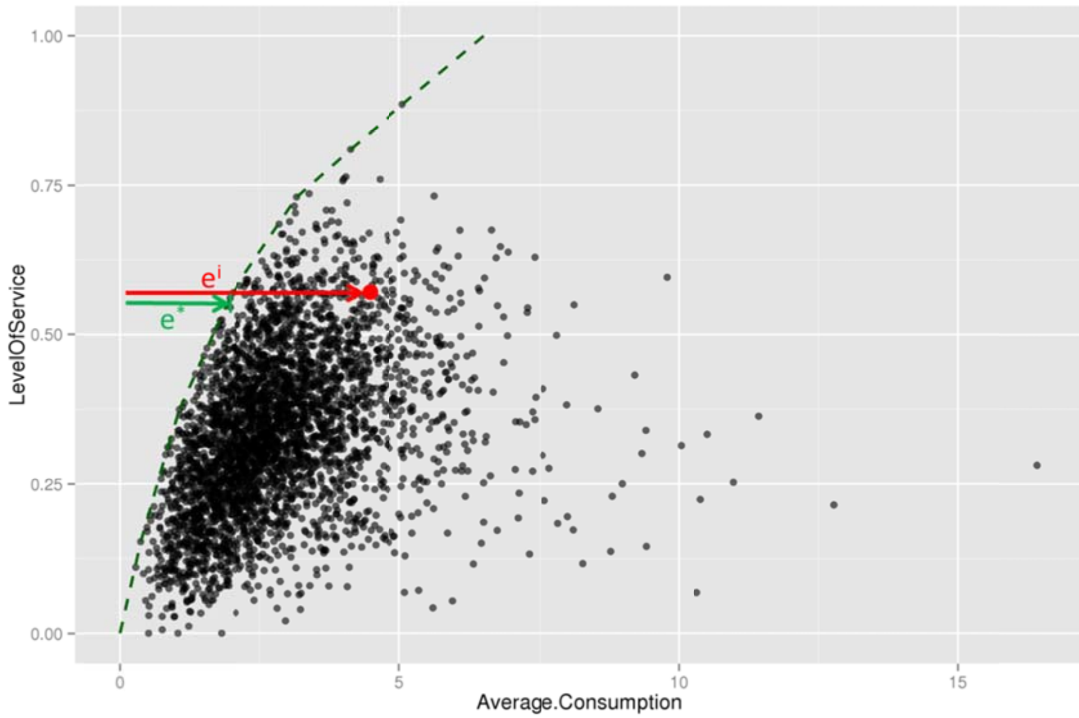


Figure 21 - Energy consumption versus level of service for a particular time of day and on particular days of the year. The energy consumption – LOS plot is bound on the left, which enables us to specify energy efficiency frontier (green dashed line). Each dot represents one home; energy consumption values are averaged over the experiment period. The y-axis (Service) shows the Appliance Ownership Index (AOI) values.

The efficiency of each household is estimated as:

$E_j = e_j / e^*$

Where:

E_j Efficiency of household j

e_j Energy consumption of household j

e^* Energy consumption of a household on the energy efficiency frontier, with the same LOS

Using SEEF algorithm, we then calculate the confidence interval of the efficiency score of each house. More details of the process are available in Kavousian and Rajagopal (2014).

4.1 Estimating Level of Service

To represent the appliance stock of each house, we use a metric called Appliance Ownership Index (AOI). It is a weighted sum of the number of appliances, with the weight of each appliance being proportional to its energy intensity. We use average annual consumption values of the appliances to represent energy intensity (Caposso et al., 1994). The AOI is normalized (setting the maximum value to one) to form LOS. In other words, the LOS of user u, denoted by LOS_u is calculated using the following formula:

$$AOI_u = \sum_{i=1}^{k} A_k \times N_{k,u}$$

$$LOS_u = \frac{AOI_u}{Max(AOI)}$$

Where A_k is the average electricity consumption for appliance k, and $N_{k,u}$ is the count of appliance k that user u owns.

4.2 Identifying Appliances Time of Use

For our case study, representative load shapes were found using adaptive K-means and summarized using hierarchical clustering, per Kwac et al., 2013. These representative load shapes formed a “dictionary”, by which we classified each day profile of each customer as one of the 7 main load shapes: morning peak, noon peak, afternoon peak, evening peak, night peak, double-peak, and inactive days (days with a flat load shape and low energy consumption).

While load shapes infer information about timing of activities, the intensity of activities reveal information about their types. For example, while a user may have a double-peak shape most nights of the week, he or she may use major appliances during certain nights (e.g., running the dishwasher or the washing machine), and only use lighting and other small activities during other week nights. For our purposes, we are interested in the higher-intensity energy use periods, when users are more likely to use their appliances. To identify those periods, we use a mixture Gaussian model, with two consumption regimes. We call these two regimes high-consumption and low-consumption days, each with average consumption of μ_1 and μ_2 respectively (Figure 22). Assuming a mixture Gaussian distribution for total daily consumption of any given user, we can find μ_1 and μ_2 and days corresponding to each regime using Expectation-Maximization algorithm (EM) (Dempster et al. 1977).

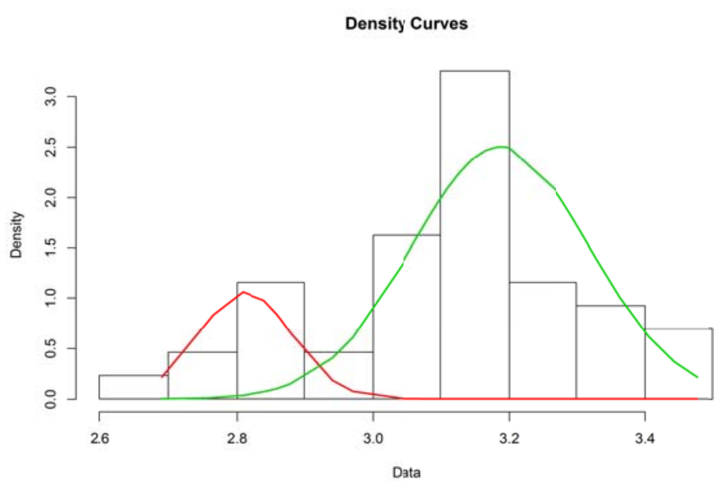


Figure 22 - A Mixture of Gaussian model to classify days of similar load shape into "low-consumption" and "high-consumption".

Once we have classified the load shape and magnitude of each day for each household, we can compare households based on their comparable day profiles.

4.3 Estimating Energy Efficiency Using Stochastic Energy Efficiency Frontier

As Figure 21 shows, the LOS-energy consumption plot is bounded from below—barring a few outliers. We identified 34 users as outliers, using a 3-standard deviation rule on the ratio of LOS on log of energy consumption. These outliers are unoccupied on average on 44% of the days, while the entire population is unoccupied on average 32% of the time.

4.4 Regression Model for Identifying the Most Important Household Variables in Driving Energy Efficiency

We regressed efficiency results on household variables to examine the drivers of energy efficiency. We assume that the energy efficiency score of each household can be explained using a linear combination of its household variables.

$$E_i = \beta_0 + \sum_{i=1}^{k} \beta_k H_{ik}$$

Where:

E_i Energy efficiency of $i-th$ user.

β_k Regression coefficient of $k-th$ household variable.

H_{ik} Value of $k-th$ household variable for $i-th$ user.

The data set includes 120+ household variables, many of which are highly correlated. Model selection techniques can reduce the number of variables, and reduce the risk of multi-collinearity. We chose stepwise regression for model selection. Stepwise regression is a greedy algorithm that starts with the variable that best describes the dependent variable, and subsequently adds to the model the variables that best improve the fit. The goodness of fit is measured using the Akaike Information Criteria (AIC):

$$AIC = -2.\log L + 2.edf$$

Where:
L The likelihood
edf Equivalent degrees of freedom (i.e., the number of free parameters)

Variables are added to the model until adding a new variable does not reduce residual deviance significantly. In short, stepwise regression identifies the variables with highest impact on goodness of fit, and ranks them in the order of their importance.

5 Results

For energy efficiency program management, it is important to understand the factors that contribute to higher or lower efficiency. This section presents the results of the regression model to explain the difference in energy efficiency using household variables (behavioral, demographic, and building variables).

Figure 23 shows the coefficients of the regression model. Variables with positive coefficients contribute to increased efficiency, while variables with negative coefficient are associated with decreased efficiency.

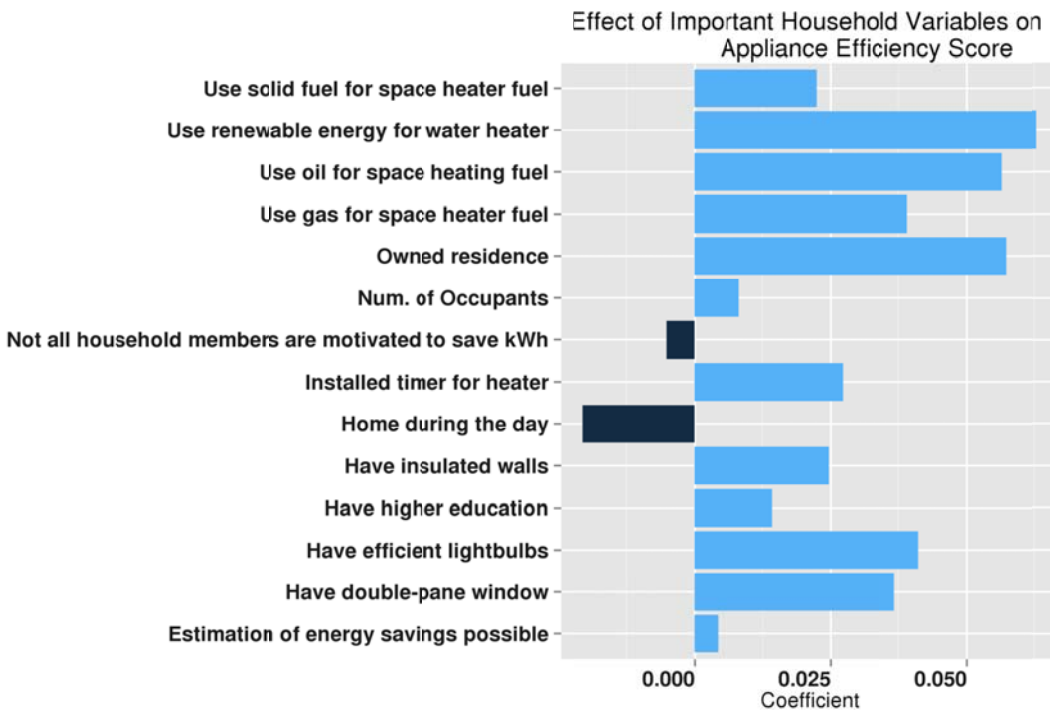


Figure 23 - Coefficients of the regression model of appliance efficiency score. The coefficients estimate the average increase in energy efficiency score as a result of one unit increase in the independent variable.

As seen in Figure 23, larger families (e.g., families with kids compared to single-person households) are on average more efficient. This is in line with previous research that has shown larger families enjoy economies of scale (Kavousian et al., 2013). Furthermore, being home during the day (captured by a variable identifying the families with unemployed, retired head of household and those with a care-giver) have demonstrated lower appliance efficiency in our data set. This suggests increased use of appliances when the house is more often occupied. Using renewable sources of energy (or sources other than electricity) for space heating and water heating contributes to higher efficiency, as does installing heater timers, efficient lightbulbs, wall insulation, and double-pane windows. Owned residences are on average more efficient, probably because of higher number of retrofits and energy efficient features that owned residences have. On behavioral variables, not being able to motivate others in the residence to save energy is correlated with lower efficiency scores. Households with the head of household having advanced degree are on average more efficient. Self-assessment of potential for energy savings has a positive but not very powerful impact on energy efficiency. Overall, as seen in Figure 23, more fundamental behavioral variables such as having higher education or living in an owned residence (implying higher socioeconomic status) have a larger effect on energy efficiency compared to self-assessment questions.

Demographic variables such as household type, education, employment showed significant correlation with energy efficiency. Generally, families with kids who have full-time employment and are highly-educated are more efficient compared to families with no kids, or families with retirees or unemployed members. This observation confirms the intuition that houses that are more active during the day are less efficient. The difference in demographics of efficient and inefficient users has important implications for communication purposes (e.g., use of local newspapers rather than on-line account). Families with kids consume the most, and also are the most responsive to incentives. When normalized for appliance ownership, families with kids also have relatively higher efficiency values due to enjoying economies of scale. Therefore, they are prime subjects for utility incentive programs.

Several behavioral variables especially around the motivation to improve energy efficiency showed correlation with energy efficiency scores. For example, when the respondents expressed doubt in being able to motivate other household members to save energy, the energy efficiency actually was lower. We also observed that households with tendency to make major lifestyle changes to save energy are indeed more efficient on average. Conversely, households who write off chances of making major lifestyle changes significantly increase their chances of being in the least efficient group of households. All these results emphasize the role of education (to a wide audience) in achieving energy efficiency goals.

Physical characteristics of the building, such as water heater fuel use, insulation, and use of efficient light bulbs showed significant correlation with energy efficiency. Households with high penetration of efficient lightbulbs and double-glazed windows showed higher efficiency. Owned residences had higher efficiency scores, most likely due to higher penetration of efficient features. Using renewable sources of energy for water heating did not impact appliance efficiency, but increased the sensitivity of electrical load to outside air temperature. This suggests that water heating load is a major load component in the homes of our study, and by removing the burden of water heating from electrical load, the most dominant load becomes space heating and cooling.

In short, larger families with higher education and both parents working full-time showed higher efficiency. People in owned residences with insulation and other efficient features also showed higher efficiency.

6 Discussion on Computational and Statistical Validation of SEEF

While SEEF offers advantages over existing methods for estimating energy efficiency, a few computational and statistical issues must be examined when using SEEF.

First, like any other frontier method, SEEF results are sensitive to outliers. While normalizing for weather and lifestyle effects reduces the chances of including outliers in the data set, it is still needed to run statistical tests to identify potential outliers. Furthermore, the outliers need to be examined separately before being completely removed from the data set. For example, in the data set used in this study, we found 32 outliers after applying common statistical outlier detection methods. Further analysis showed that these 32 outliers were mostly un-occupied residences, which validated our decision to remove them from the data set. However, it is also likely that some homes with extremely low consumption levels have on-site generation, in which case a separate treatment of that group of homes is necessary.

Next, as seen from the frontier plots, the data set becomes sparse as the LOS increases. This is expected; for example, the likelihood of observing homes with more than 4 TVs or three refrigerators is less than the likelihood of observing homes with one TV and one refrigerator. The sparse population of houses in higher LOS levels may introduce an upward bias in energy efficiency ranking of these houses, since, with the sparsity of data, it will be more likely to not observe more efficient homes that would move the frontier to the left.

The sparse population at high LOS levels also implies that a smaller number of unique buildings will be on the efficiency frontier throughout bootstrap iterations. When there are more homes close to the frontier (as is the case in lower LOS levels), at different bootstrap iterations different buildings may form the frontier; e.g., the chance that building A consumes less energy than building B during some months but not all the time is higher when both buildings overall use similar amounts of energy. The change in buildings that define the frontier can result in wider confidence interval in the mid-range of LOS values, while the confidence interval will get narrower at the higher or lower LOS values. While narrower confidence interval may not result in a bias in energy efficiency estimates, it will impact the results of comparison of energy efficiency values.

As explained above, the potential bias in efficiency estimation and narrow confidence interval of high-LOS buildings is caused by sparsity of data at high-LOS region of the energy-LOS plot. Thus, more data are needed to estimate the upward bias in efficiency estimates and the downward bias in confidence intervals. But an alternative solution is to use perturbation methods to simulate more high-LOS buildings. In this method, existing high-LOS buildings will be copied with a random added error term to form new simulated buildings. The distribution of the error term will be estimated based on the distribution of energy consumption and LOS at denser parts of the energy-LOS plot. Another alternative is to examine high-LOS buildings separately.

7 Conclusions

The goal of the majority of energy efficiency ranking (benchmarking) efforts is to extract actionable insights that can help the users or energy efficiency planners to reduce their energy consumption. When comparing energy consumption of different buildings against each other, one needs to control for the effect of variables that are outside the influence or interest of the users to change. For example, while the occupants may be able to change their lifestyle to save energy, not all of them may be willing to do that.

Energy consumption is driven by users' desire to use the services in a house (e.g., refrigeration or air conditioning), and is impacted by several internal and external variables, such as weather, building size, building systems, appliance ownership, and behavior of the occupants. Energy consumption is a highly fluctuating variable, hence energy efficiency is a stochastic variable. If one is interested in making statistical inferences about the efficiency of different users, the confidence interval for efficiency scores is needed. In this paper, we introduced a process to rank buildings' energy efficiency and calculate confidence interval for their ranks. We also demonstrated how to account for weather, seasonal, and lifestyle effects on energy consumption before ranking buildings. This framework is based on a few assumptions and principles:

(a) Energy consumption is stochastic, hence comparable periods of consumption should be used for comparison and benchmarking. Users have different consumption regimes in different days. These regimes are different both in pattern and magnitude of consumption. If the goal is to estimate how efficiently someone's appliances run, and how efficiently they use the appliances, we must focus attention on the periods that the user is home and uses the appliances.

(b) Frontier methods specify the boundary of relationship between LOS and energy consumption. Plots of energy consumption versus Level of Service (LOS) suggest that while the energy consumption generally increases with increase in LOS, for any given LOS a wide range of consumption values can be observed. While minimum possible energy consumption for a given house is conceivable (for example, having only refrigerator load for several days), the upper bound of energy consumption is not tractable. Therefore, a frontier model where the lower boundary of consumption for a given level of detail is specified is more appropriate than a regression model that is affected by all data points. Therefore, the relationship between energy consumption and LOS can be best specified using a frontier. To add stochasticity to efficiency estimates calculated using efficiency frontiers, we used Stochastic Energy Efficiency Frontier (SEEF) to estimate the energy efficiency of households.

(c) The role of household variables on appliance energy efficiency: The variation in energy efficiency can be partially explained by household variables. Examining the correlation of household variables with energy efficiency can help utilities in both planning and communicating the energy efficiency programs. For example, knowing that households with part-time employed or unemployed members as well as families with lower education levels have lower appliance usage efficiency can help utilities adjust their incentives and messaging of energy efficiency programs to obtain the best results.

The results of this work can help utilities in planning, targeting, and marketing energy efficiency programs. Furthermore, power systems researchers and policy makers can use the insights of our work in their market analysis.

In this paper, we analyzed the appliance load, and also briefly examined the thermal component of residential electrical load. Another important component of electrical load is lighting load. In our future work, we plan to examine lighting efficiency, and compare it to appliance and thermal efficiency of households.

While in this case study the appliance data were available, the methods described in the paper are not dependent on appliance data availability. In more general cases, estimates of appliance ownership can be obtained directly from the smart meter data using classification methods. The authors are involved in a follow-up study to integrate appliance ownership estimation techniques with the appliance efficiency estimation method used in this paper. Appliance ownership data can also be collected using or survey methods or inferred using nationally-available data (e.g., Residential Energy Consumption Survey or Residential Appliance Saturation Survey in the US).

8 Acknowledgements

The first author was funded by Center for Integrated Facility Engineering (CIFE) at Stanford University. The authors would also like to thank colleagues at Stanford Sustainable Systems Lab (S3L), especially Dr. June Flora, Dr. Chi-Woo Tan, Adrian Albert, and Jungsuk Kwac for their valuable comments and inputs throughout this research.

9 References

Albert, A.; Rajagopal, R., "Smart Meter Driven Segmentation: What Your Consumption Says About You," Power Systems, IEEE Transactions on , vol.28, no.4, pp.4019,4030, Nov. 2013. doi: 10.1109/TPWRS.2013.2266122

Beckel, C., Sadamori, L., Santini, S. (2012) Towards automatic classification of private households using electricity consumption data. Proceedings of the Fourth ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings. Pages 169-176.

Beckel, C., Sadamori, L., Santini, S. (2013) Automatic socio-economic classification of households using electricity consumption data. Proceedings of the fourth international conference on Future energy systems. Pages 75-86.

Berges, M., Goldman, E., Matthews, H. .S., Soibelman, L. (2008) Training Load Monitoring Algorithms on Highly Sub-Metered Home Electricity Consumption Data, Tsinghua Science & Technology, Volume 13, Supplement 1, October 2008, Pages 406-411, ISSN 1007-0214, http://dx.doi.org/10.1016/S1007-0214(08)70182-2.

Capasso, A., Grattieri, W., Lamedica, R., Prudenzi, A. (1994) A Bottom-Up Approach to Residential Load Modeling. IEEE Transactions on Power Systems, 9 (2). May 1994.

Caves, D., Christensen, L. (1980) Econometric analysis of residential time-of-use electricity pricing experiments, Journal of Econometrics, Volume 14, Issue 3, December 1980, Pages 287-306, ISSN 0304-4076, http://dx.doi.org/10.1016/0304-4076(80)90029-9.

Commission for Energy Regulation of Ireland (CER) (2011) “Smart Metering Information Paper 4: Results of Electricity Cost Benefit Analysis, Customer Behavioral Trials, and Technology Trials”. Information Paper CER11080, May 16, 2011. Published by the Commission for Energy Regulation, Dublin, Ireland.

Dempster, A.P., Laird, N.M. , Rubin, D.B. (1977): "Maximum Likelihood from Incomplete Data via the EM algorithm", Journal of the Royal Statistical Society, Series B, vol. 39, 1:1-38.

Kavousian, A., Rajagopal, R., Fischer, M. (2013) Determinants of residential electricity consumption: Using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants' behavior. Energy, Volume 55, 15 June 2013, Pages 184-194, ISSN 0360-5442, http://dx.doi.org/10.1016/j.energy.2013.03.086.

Kavousian, A. and Rajagopal, R. (2014). "Data-Driven Benchmarking of Building Energy Efficiency Utilizing Statistical Frontier Models." Journal of Computing in Civil Engineering, 10.1061/(ASCE)CP.1943-5487.0000327.

Kwac, J., Rajagopal, R., Flora, J. (2014) Household Energy Consumption Segmentation Using Hourly Data. Smart Grid, IEEE Transactions on , vol.5, no.1, pp.420,430, Jan. 2014. doi: 10.1109/TSG.2013.2278477.

Guy R. Newsham, Brent G. Bowker, The effect of utility time-varying pricing and load control strategies on residential summer peak electricity use: A review, Energy Policy, Volume 38, Issue 7, July 2010, Pages 3289-3296, ISSN 0301-4215, http://dx.doi.org/10.1016/j.enpol.2010.01.027.

Karen Herter (2007) Residential implementation of critical-peak pricing of electricity, Energy Policy, Volume 35, Issue 4, April 2007, Pages 2121-2130, ISSN 0301-4215, http://dx.doi.org/10.1016/j.enpol.2006.06.019.

Knight I, Kreutzer N, Maaning M, Swinton M, Ribberink H. (2007) European and Canadian non-HVAC electric and DHW load profiles for use in simulating the performance of residential cogeneration systems. International Energy Agency, Annex 42–FC + COGEN-SIM, Subtask A; 2007.

Margossian, B. (1994) Deriving End-Use Profiles Without End-Use Metering: Results of Recent Validation Studies. Proceedings of the ACEEE 1988 Summer Study on Energy Efficiency in Buildings, Pacific Grove, CA. Volume 2, pp. 2.217-2.223. American Council for an Energy-Efficient Economy, Washington, D.C.

Massimo Filippini, Short- and long-run time-of-use price elasticities in Swiss residential electricity demand, Energy Policy, Volume 39, Issue 10, October 2011, Pages 5811-5817, ISSN 0301-4215, http://dx.doi.org/10.1016/j.enpol.2011.06.002.

McLoughlin, F., Duffy, A., Conlon, M. (2012), Characterizing domestic electricity consumption patterns by dwelling and occupant socio-economic variables: An Irish case study, Energy and Buildings (2012), doi:10.1016/j.enbuild.2012.01.037

Natural Resources Defense Council, NRDC (2014) Efficient Appliances Save Energy and Money. Last accessed via http://www.nrdc.org/air/energy/fappl.asp

Parti, M., Parti, C. (1980) The Total and Appliance-Specific Conditional Demand for Electricity in the Household Sector. The Bell Journal of Economics 11(1) 309-321.

Richard J. Sexton, Nancy Brown Johnson and Akira Konakayama (1987) Consumer Response to Continuous-Display Electricity-Use Monitors in a Time- of-Use Pricing Experiment. Journal of Consumer Research, Vol. 14, No. 1 (Jun., 1987), pp. 55-62.

US DOE in New York Times (2010) Last accessed on 05/07/2014 at
http://green.blogs.nytimes.com/2010/03/19/consumers-buy-more-efficient-refrigerators-but-keep-the-old-ones-humming/?_php=true&_type=blogs&_r=0

US Department of Energy (2014) Building Energy Databook, Table 2.1.16 Operating Characteristics of Electric Appliances in the Residential Sector. Last accessed 2014/02/24 via
http://buildingsdatabook.eren.doe.gov/TableView.aspx?table=2.1.16

Zeifman, M., Roth, K. (2011) Nonintrusive appliance load monitoring: Review and outlook. IEEE Transactions on Consumer Electronics. 57 (1).

Appendix 1 – Summary Statistics

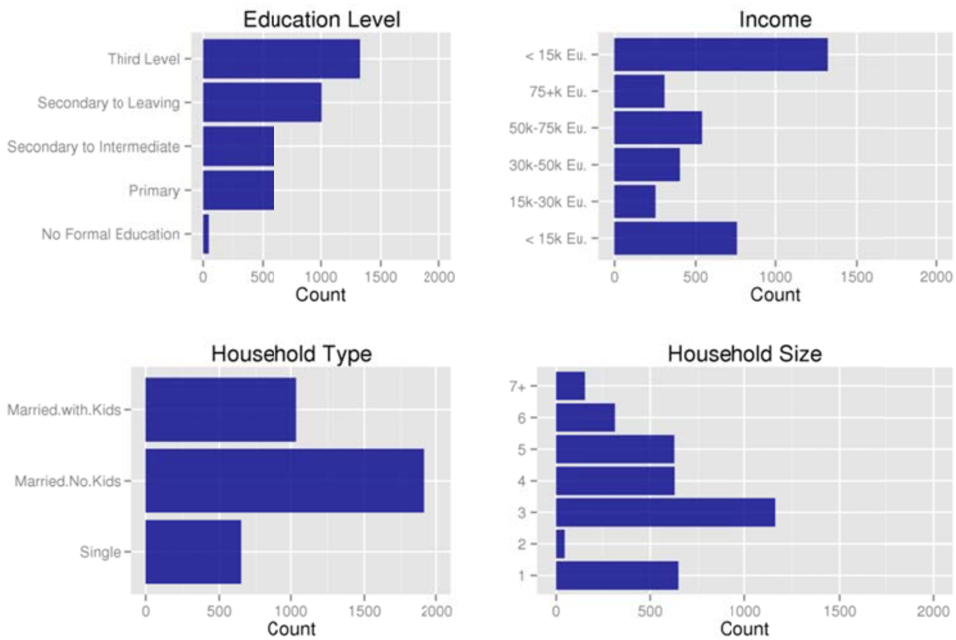


Figure 24 – Summary of demographic characteristics of study participants.

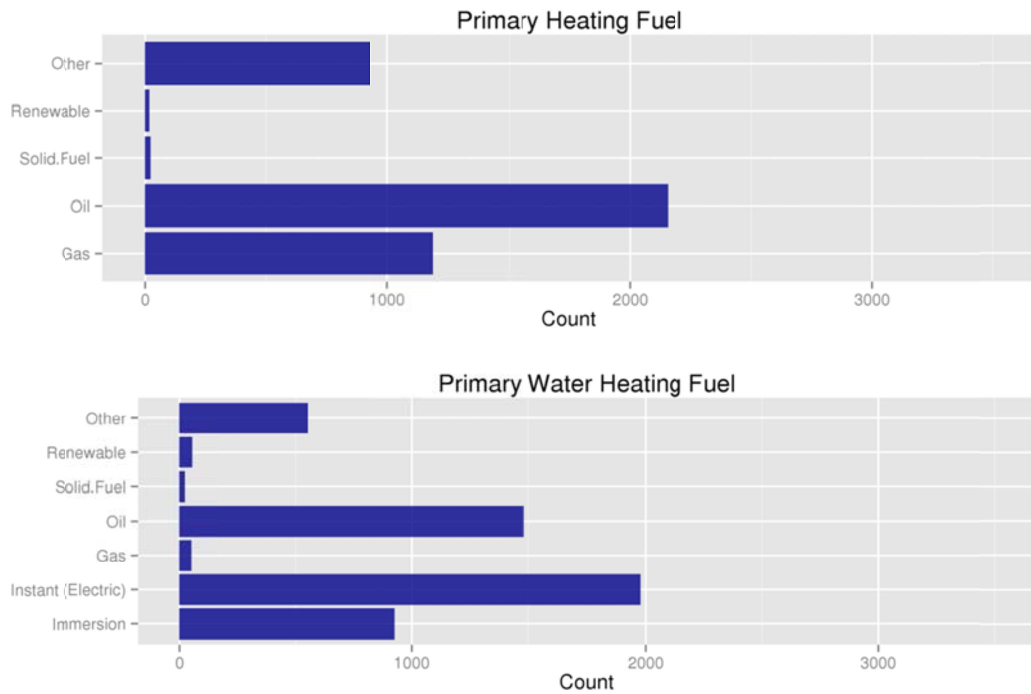


Figure 25 - Primary heating and water heating fuel.

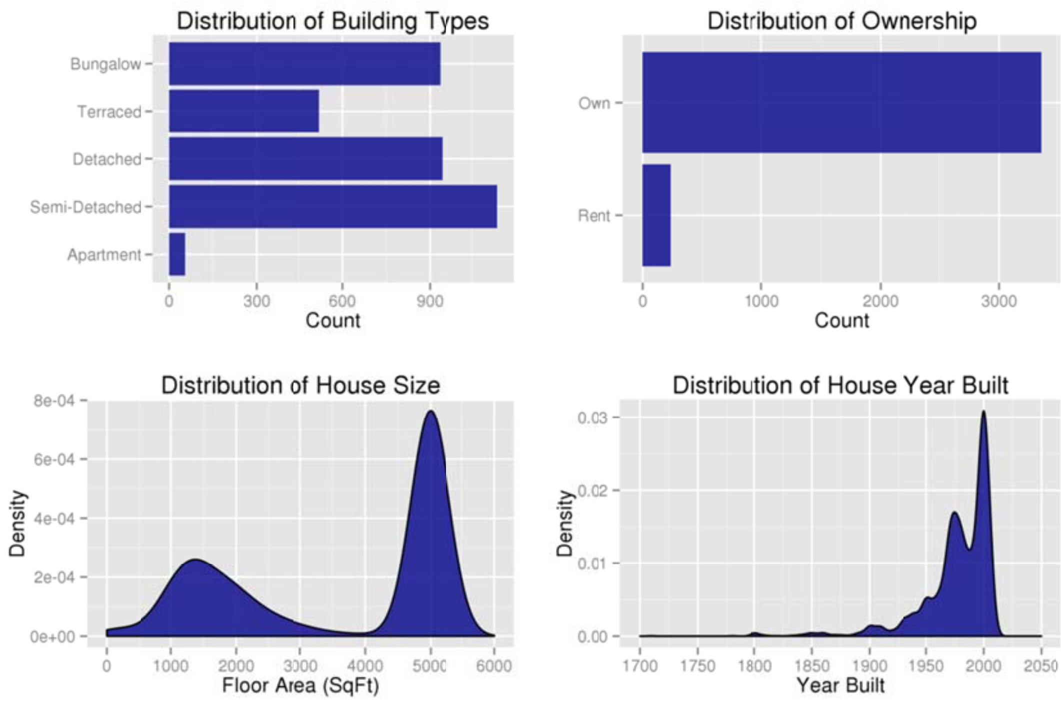


Figure 26 - Summary of building characteristics

Table 9 – Factors contributing to home energy sensitivity to outside temperature and seasonal effects.

Variable	**Estimate**	**Standard Deviation**	**P-Value**
Respondent's Age	0.019887	0.002138	< 2e-16
Building type (apartment, detached bungalow)	0.010813	0.001967	4.11e-08
Percentage of efficient lightbulbs	-0.040585	0.006546	6.29e-10
Using renewable power for water heating	0.073243	0.019097	0.000128
Social class (based on employment; higher value for higher income)	0.007641	0.002283	0.000824
Number of occupants	0.006556	0.001557	2.62e-05
Ownership (or rent)	0.033777	0.009590	0.000434
Using double-pane windows	-0.023038	0.008340	0.005771
Education (higher value for higher education)	-0.005328	0.002300	0.020570
Using timer for water heater	0.011850	0.004873	0.015076
Using timer for space heater	-0.013123	0.006632	0.047907
Using solid fuel for heating	-0.011065	0.005599	0.048223

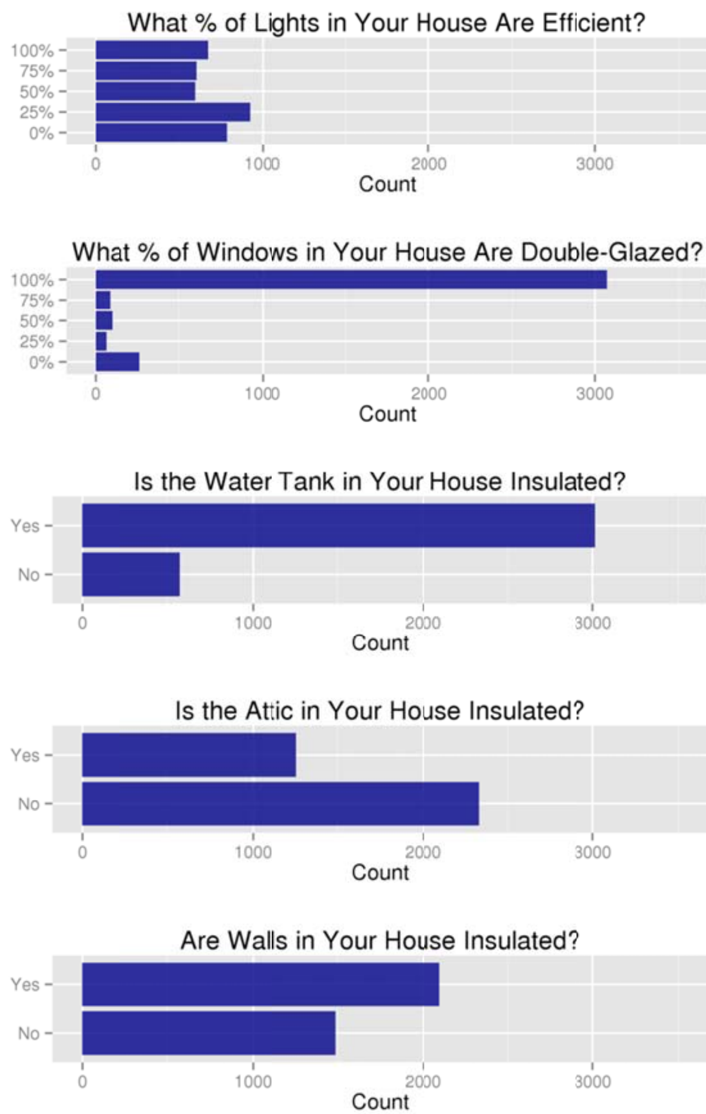
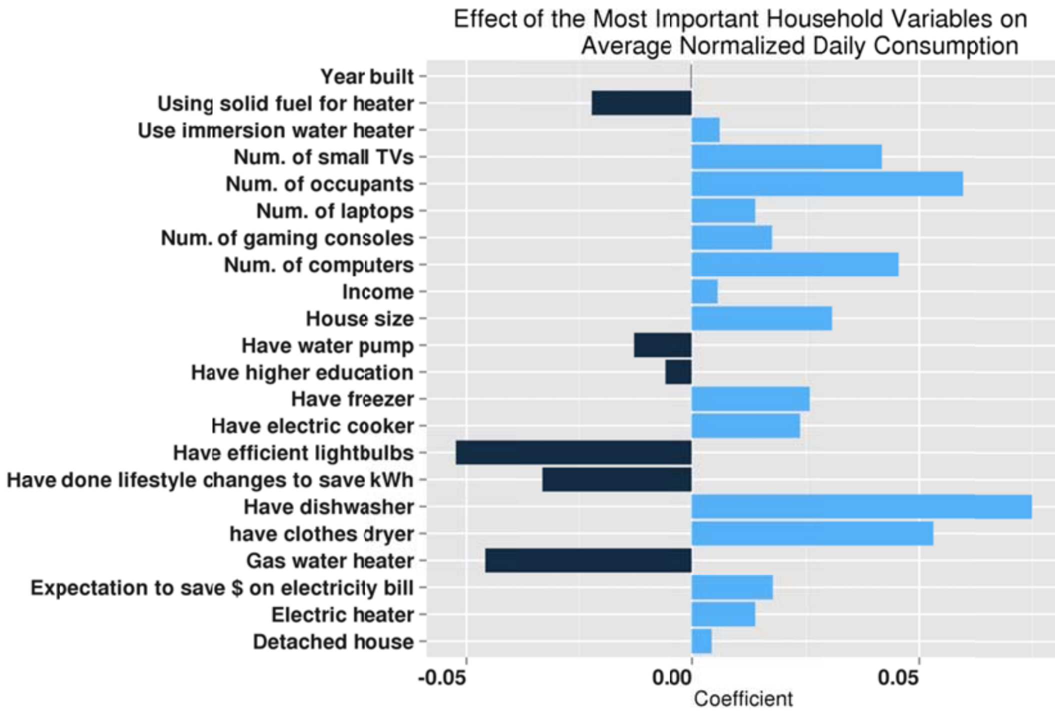


Figure 27 - Summary of efficient features of homes.

Appendix 2 – Other Model Results



Chapter 5

Actionable Benchmarking of Commercial Buildings: Examining the Effectiveness of Energy Conservation Measures and the Impact of Operational and External Factors on Energy Efficiency Scores Utilizing Stochastic Energy Efficiency Frontiers

Amir Kavousian, Martin Fischer, Ram Rajagopal

Keywords: Actionable Benchmarking, Energy Star, IPMVP Option C, Stochastic Energy Efficiency Frontier (SEEF), Energy Conservation Measures (ECMs), Energy Management

1 Abstract

Energy efficiency benchmarking methods help energy managers monitor energy performance of buildings and prioritize energy efficiency investments. Several benchmarking methods are available to compare a building against its own, against buildings of a portfolio, or against a large number of buildings of the same class. However, a critical review of different benchmarking methods based on how they support ongoing improvement of buildings' energy efficiency is missing. In particular, there is a need to understand how different benchmarking scores change as a result of changes in operational and external parameters that affect building performance to be able to interpret benchmarking scores.

This paper first reviews different classes of benchmarking methods using three specific examples: IPMVP Option C, SEEF, and Energy Star. These methods

respectively compare a building with its own, with a portfolio of buildings, and with all buildings of the same class. The interactions of each benchmarking method with two different energy factors are examined. First, the effect of internal changes such as implementation of retrofits or a change in number of occupants on benchmarking scores is examined. Next, the sensitivity of each benchmarking method to external changes such as change in outside temperature is analyzed. To demonstrate these interactions, a data set of monthly energy consumption and operational data from 120 buildings on the campus of a technology firm is used.

Our results suggest that general benchmarking methods such as Energy Star are useful for reporting and public purposes, while individual-buildings methods such as IPMVP are valuable in detailed analysis of the impact of interventions on energy consumption. Portfolio comparison methods such as SEEF are useful in monitoring the energy efficiency of a portfolio of buildings, and in inferring retrofit options for buildings of interest. Furthermore, portfolio methods such as SEEF are more flexible in the choice of metrics and in specifying the expected relationship of energy factors and energy consumption (e.g., linear versus non-linear relationship between kWh and area).

The results of our research helps practitioners choose the benchmarking method that is most appropriate for their purposes, and to better interpret the benchmarking results.

Finally, we observed that the quality of models and analyses can considerably benefit from more detailed data, such as interval consumption data and Building Management System (BMS) data.

2 Introduction

Commercial buildings account for nearly a fifth of US primary energy use and more than a third of US electricity consumption (US DOE, 2014). With existing technologies, it is possible to economically reduce energy consumption of commercial buildings by 30%. When emerging technologies are included, this figure reaches 80% by 2020 (Farese et al., 2012). In 2010 alone, energy utilities in the US spent $7B in energy efficiency projects in buildings; however, research shows that most energy efficiency projects do not achieve their goal of energy reduction, or even result in increased consumption (Deutsche Bank & Living Cities, 2011; Shapiro, 2011). An important step in ensuring that investments achieve expected energy reduction goals is to select the buildings that collectively offer the highest energy reduction. Energy benchmarking can help in selecting buildings for retrofit.

While several widely-used benchmarking methods exist, they mostly serve for demonstration purposes rather than as decision tools (Peterman, 2012). In the absence of formal decision making processes, energy managers and property owners usually resort to personal opinion and heuristics for selecting the buildings to retrofit and the type of retrofit to apply to the buildings. This practice normally results in implementing cooling, heating, or lighting retrofits on the largest and the most energy intense buildings, while attractive retrofits such as insulation and infiltration improvement are often ignored (Shapiro, 2011). To achieve desired outcomes, energy managers need to select proper benchmarking method. However, a comparison of the most important benchmarking methods is missing.

To address this gap, this paper offers an overview and comparison of the most common methods for benchmarking, and proposes a method to examine the results of Energy Conservation Measures (ECMs) on the benchmarking scores. The interaction of the benchmarking results with operational, structural, and external factors is examined, and practical recommendations for energy managers on how to interpret benchmarking results is offered. Furthermore, using a real-world case study, this paper compares the performance of the benchmarking methods, and identifies their advantages and weaknesses in different settings. Particular attention is given to minimal data situations, since most existing buildings still use monthly data to monitor building performance.

The following section introduces the most common benchmarking methods chosen for comparison in this paper. Then, real-world data from a case study is used to compare these benchmarking methods, and examine their correlations with ECMs. Based on the results of the comparisons, the pros and cons of each method and specific conditions under which to select each benchmarking method are discussed.

2.1 Longitudinal Benchmarking

Primarily used for Measurement and Verification (M&V), longitudinal benchmarking compares the actual performance of the building with an empirical model developed based on past performance data of the building itself. Any deviation of future energy consumption from values projected by the empirical model is regarded as an improvement or decline in energy efficiency, depending on the direction of the diversion (Figure 28).

A common method for longitudinal benchmarking is the International Performance Measurement and Verification Protocol – Option C Whole Building Measurement, referred to as IPMVP Option C in this paper. The IPMVP fits a least-squares regression model of the actual consumption against independent variables such as Cooling Degree Days and Heating Degree Days (CDD and HDD) (Porter et al., 2005). The IPMVP regression model for building k (assuming the baseline period included m observations) is as follows:

$$E_B^k = X_B^k . \beta^k + \varepsilon^k$$

Where:

E_B^k : $(m \times 1)$ matrix of observed energy consumption during the baseline period

X^k : $(m \times (n+1))$ matrix of observed independent variables, such as CDD, HDD, Headcount, Number of computers, etc.

β^k : $((n+1) \times 1)$ matrix of model parameters, which will be estimated by the least squares method

ε^k : $(m \times 1)$ matrix of model residuals, which are assumed to be other effects and trends not identified using the simple regression model.

Once the baseline model is fitted, the projected values are estimated by the following equations:

$$E_P^k = X_P^k . \beta^k + \varepsilon^k$$

Where:

E_P^k : $(m \times 1)$ matrix of predicted energy consumption for the post-baseline period

X_P^k : $(m \times (n+1))$ matrix of n observed independent variables during the post-baseline period. The intercept is assumed to be the baseload; i.e., the load that does not depend on independent variables such as CDD and occupancy.

β^k : $((n+1) \times 1)$ matrix of model parameters, already estimated by the least squares method using the baseline period data.

ε^k : $(m \times 1)$ matrix of model residuals of the baseline model.

Finally, the avoided energy consumption (i.e., energy savings) is calculated as:

$$E_A = E_P - E_M \qquad (1)$$

Where E_P is the estimated (projected) energy consumption as explained above, and E_M is the actual measured energy consumption during the post-baseline period.

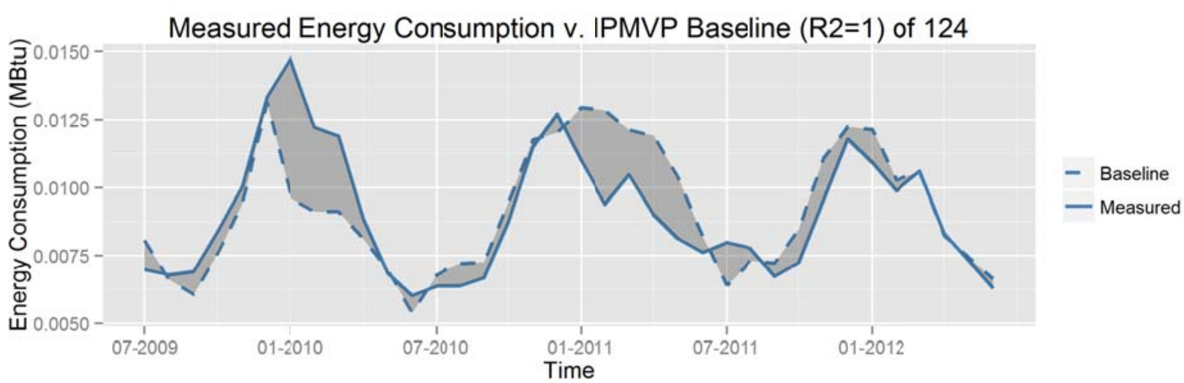


Figure 28 - Example of longitudinal benchmarking using the IPMVP method, using a rolling baseline starting on 07-2008 through 06-2009 and rolling forward, making each year's consumption the baseline for the following year. Baseline is IPMVP projected; Measured is actual observed consumption.

Simulation-based methods are essentially similar to longitudinal methods explained above, with the difference that the building energy model is simulated from historical data of a representative set of buildings, rather than the building of interest. Griffith & Crawley (2006) explain a method for estimating the technical potential for energy efficiency in building using detailed energy modeling. However, high data and computational requirements of simulation models, as well as problems with model verification have made them a less practical method for benchmarking of large number of buildings (Maile, 2009). In this paper we do not examine simulation-based methods.

2.2 Cross-Sectional Benchmarking

Cross-sectional benchmarking methods compare a building with its peer buildings, normalizing for factors such as floor area, occupancy, energy fixtures, and climate. The most common cross-sectional benchmarking method is Energy Star, developed by the US Environmental Protection Agency (EPA) under the national energy performance rating program. Energy Star is an external (cross-sectional) benchmark to assess how efficiently buildings use energy, relative to similar buildings nationwide. The rating system is based on a 1-100 scale for each building type; a rating of 50 indicates average energy performance while a rating of 75 or better indicates top performance. EPA evaluates and modifies the rating system separately for each building type. For office buildings, the following seven variables are currently used in the model: square foot, number of occupants, operating hours, number of workers per square foot, location and the amount of conditioned and heated space (EPA, 2011). Once the predicted EUI is calculated, the ratio of actual source EUI to predicted source EUI can be calculated, and compared with the distribution of this ratio over all buildings. The Energy Star score is calculated as the quantile of the calculated ratio with respect to the industry distribution; e.g., an Energy Star value of 75 means that only 25% of buildings have a ratio of actual to predicted energy consumption of more than that building. Table 10 summarizes model parameters for Energy Star calculations.

Table 10 - Energy Star predicted EUI calculation summary

Variable	Mean	Coefficient
Constant (intercept)	--	186.6
Natural log of square foot	9.535	34.17
Number of computers per 1000 ft^2	2.231	17.28
Natural log of weekly operating hours	3.972	55.96
Natural log of number of workers per 1000 ft^2	0.5616	10.34
Heating Degree Days x Percent heated	4411	0.0077
Cooling Degree Days x Percent cooled	1157	0.0144

2.3 Non-Parametric Methods

Frontier methods quantify the energy efficiency of buildings by forming an efficient frontier (i.e., best-practice technology) and comparing all buildings against that frontier (Figure 29). The efficient frontier is formed by examining how much energy a building consumes in relation to the Level of Service (LOS) that it provides. Frontier methods divide into two categories: parametric and non-parametric. Parametric frontier methods, also called Stochastic Frontier Analysis (SFA), assume that the efficient frontier can be presented using a mathematical formula plus a random component (Aigner et al., 1977). Non-parametric frontier methods such as Data Envelopment Analysis (DEA) and Stochastic Energy Efficiency Frontier (SEEF) make no assumptions about the mathematical formula of the frontier, and instead form the frontier using measured data (Kavousian & Rajagopal, 2013). Our focus in this paper is on non-parametric frontier methods, and in particular SEEF. Thus, whenever we mention frontier models in the paper, we are referring to non-parametric frontier models.

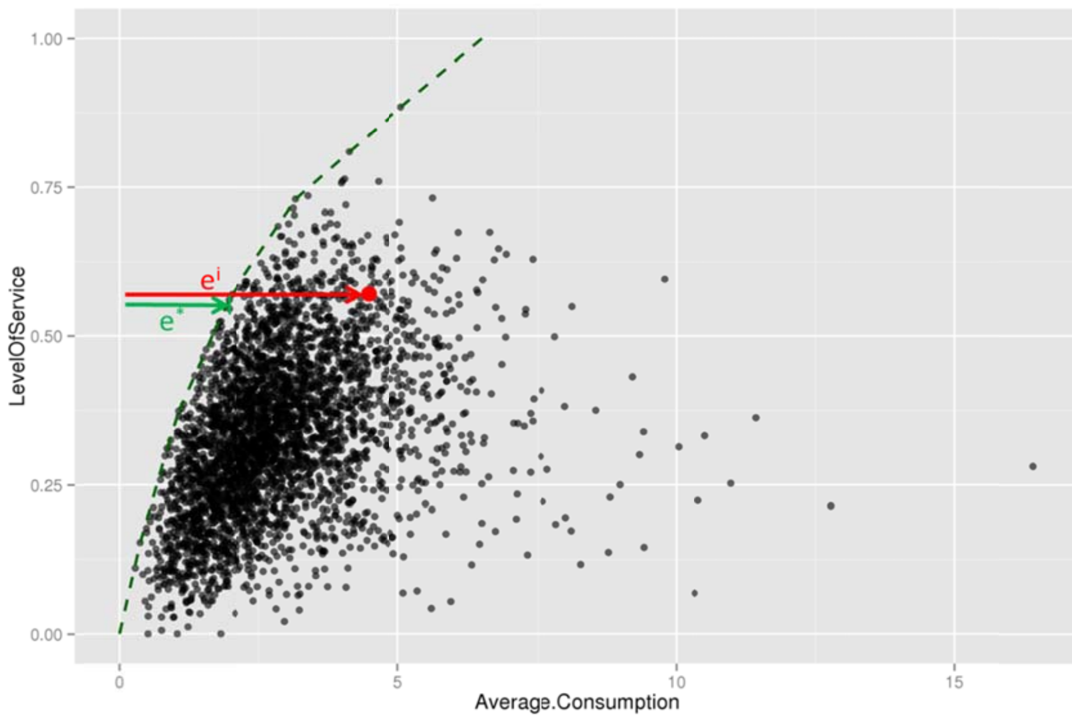


Figure 29 - Frontier methods quantify the energy efficiency of buildings by forming an efficient frontier (i.e., best-practice technology) and comparing all buildings against that frontier.

The efficiency of each building is estimated as:

$$E_j = e_j \,/\, e^*$$

Where:

E_j Efficiency of building j

e_j Energy consumption of building j

e^* Energy consumption of a building on the energy efficiency frontier, with the same LOS

Huntington (1995) enumerates the shortcomings of existing methods, such as strong assumptions that parametric models make to develop the energy model, and suggests the use of frontier methods to deal with those complexities and shortcomings. Some examples of applying DEA in energy efficiency are found in Carnes et al., 2008; Lee & Lee, 2008; Grosche, 2009; Onut & Soner (2006); Ferrier & Hirschberg (1992). For reviews of using this method in environmental studies, see Ang & Zhang (2000) and Zhou et al. (2008). Kavousian & Rajagopal (2013) describe the shortcomings of existing applications of DEA in energy efficiency, and introduce Stochastic Energy Efficiency Frontier (SEEF) for quantifying the energy efficiency of buildings using high or low resolution consumption data.

2.4 Other Benchmarking Methods

Some benchmarking methods focus on a particular aspect of a building's performance, e.g., heating and cooling load. For example, in some studies the correlation of outside temperature with cooling and heating loads has been used as benchmarking metrics. (Shiel, 2014). In this paper, we examined the correlation of electricity consumption with Cooling Degree Days (CDD), the correlation of electricity consumption with Heating Degree Days (HDD), and the correlation of gas consumption with HDD.

3 Case Study

A data set of monthly energy consumption and operational data from 120 buildings on the campus of a technology firm is used to compare energy benchmarking methods and energy efficiency projects on commercial buildings. The next sections describe the experimental data and the results of the case study, followed by a discussion and recommendations for practitioners.

3.1 Experimental Data

The experimental data includes energy consumption and operational data for 120+ buildings on the campus of a technology firm in Washington, US, spanning from July 2008 to July 2012. The data set includes monthly energy consumption (electricity and gas) of individual buildings, as well as their headcount, number of computers, data center load, HVAC system type, and a breakdown of floor area (office space, data center space, café space, and warehouse). The data set also includes energy management data including dates and descriptions of Energy Conservation Measures (ECMs), retrofits, and Retro-commissioning (RCx) projects. Several building types (including developer buildings, administrative buildings, data centers, cafeterias, and sales buildings) were included in the data set. For our analyses, we only focused on developer and administrative buildings, since they constitute the largest group of buildings, and are similar to most common commercial buildings.

3.2 Methods

Data for 44 developer and administration buildings that had sufficient amount of verifiable data were used for the analyses. Three different benchmarking methods were applied on the data set: Energy Star, IPMVP Option C, and Stochastic Energy Efficiency Frontier (SEEF).

IPMVP procedure uses the difference between predicted and measured energy consumption as a performance metric. To convert this metric to a score, we used the empirical quantiles of IPMVP scores (48 months x 44 buildings = 2112 values), and ranked each building at each month based on their quantile. Figure 30 shows the Empirical Cumulative Distribution Function (ECDF) of monthly values of avoided energy consumption of all buildings.

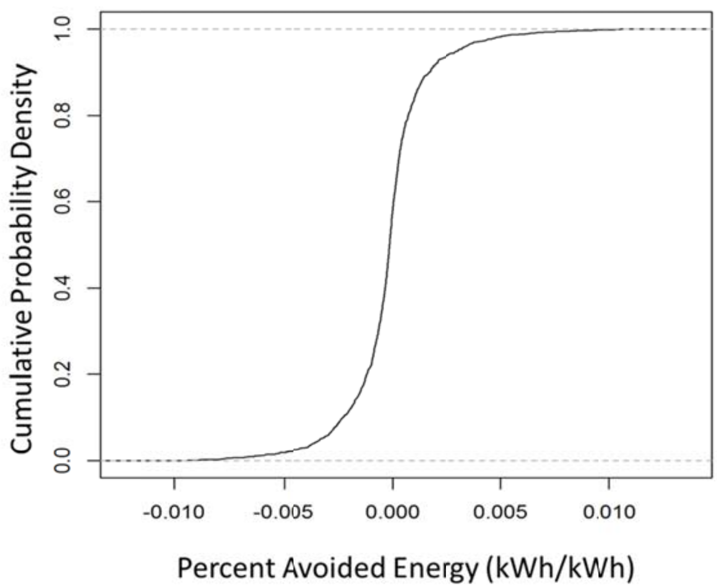


Figure 30 - ECDF of avoided energy (MBtu).

SEEF values are calculated per Kavousian & Rajagopal (2014). Figure 31 shows the results of SEEF analysis. Four buildings consistently outperform other buildings, and the lowest efficiency score is around 20%.

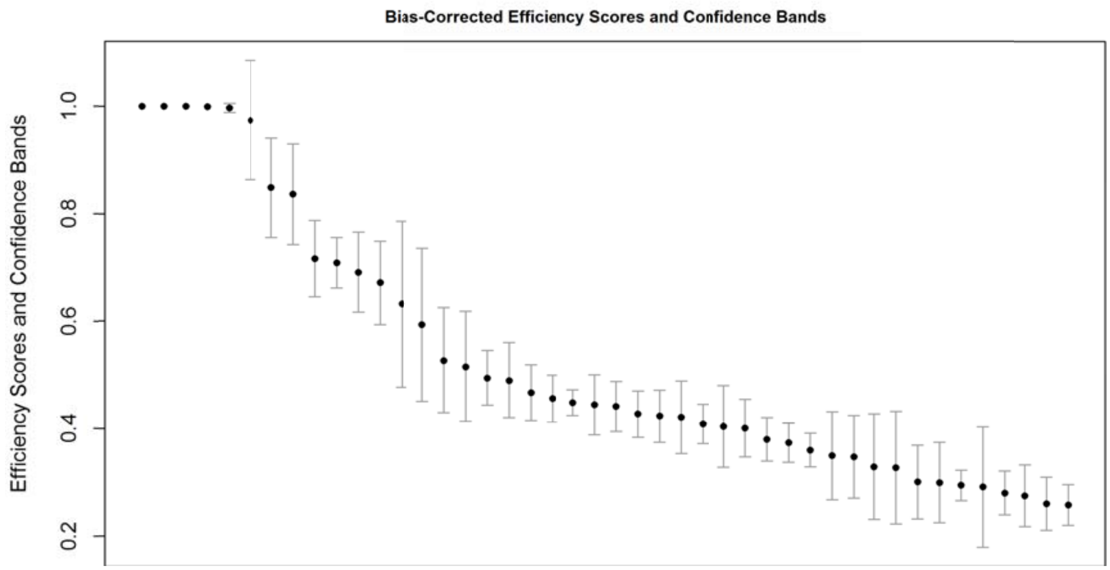


Figure 31 - Results of Stochastic Energy Efficiency Frontier (SEEF). Each bar represents a building (building names removed to maintain anonymity). Dots are the average efficiency scores over the 36 months of analysis, and the grey lines show the 95% confidence interval.

Finally, Energy Star scores are calculated using Energy Star Performance Ratings Technical Methodology for Office Buildings (US EPA, 2011). After calculating buildings' ranks using three different methods as explained above, we compared the efficiency scores across buildings and across different methods.

4 Results

We compared different benchmarking methods against each other, and against factors such as ECMs, external factors such as outside temperature, structural factors such as system type, and operational factors such as number of occupants.

4.1 Comparison of Benchmarking Methods

Figure 32 shows the results of correlation analysis on efficiency scores of buildings. Each point on the lower-left plots represents one building. Some plots have less than

44 buildings, due to invalid data on one or more efficiency scores. The values on the upper-right cells of the matrix are correlations of the efficiency scores. Finally, the histograms on the diameter show the distribution of each score. From left to right and top to bottom, the boxes represent: ECM.sz (total volume of ECMs in MBtu), SEEF, Energy Star, and IPMVP Option C score. Volume of ECMs is the expected annual savings as the result of the ECM, estimated by the energy management team.

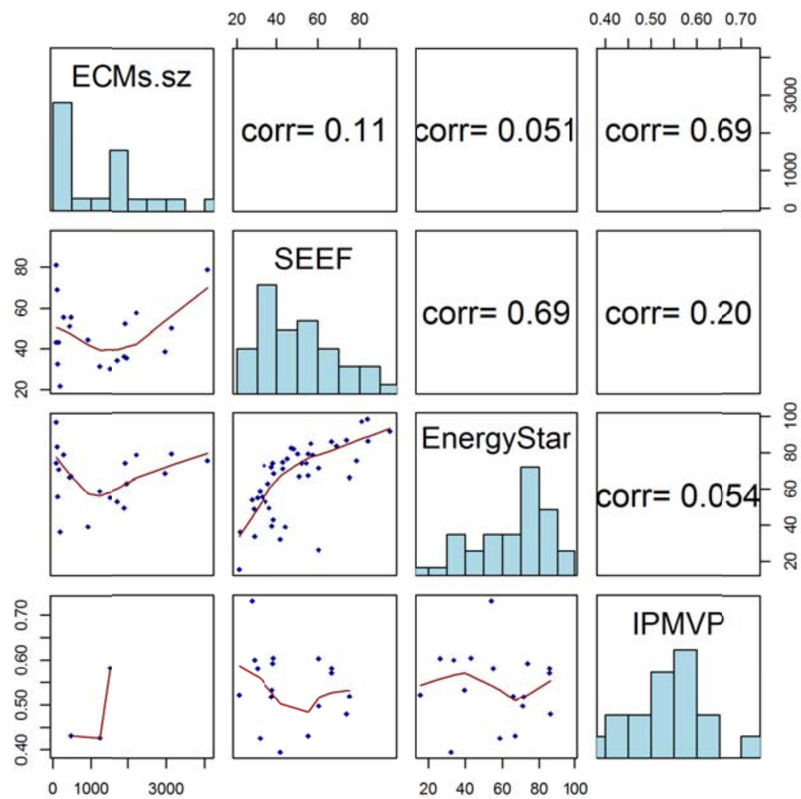


Figure 32 - Correlation of efficiency scores for different benchmarking methods. From left to right and top to bottom, the boxes represent: ECM.sz (total volume of ECMs in MBtu), SEEF, Energy Star, and IPMVP Option C score. Red lines on the lower-left boxes are LOWESS fit.

As Figure 32 suggests, most relationships between efficiency metrics are non-linear. Particularly, the relationship between volume of ECMs and SEEF/Energy Star scores is interesting: while smaller ECMs do not considerably contribute to improving SEEF/Energy Star scores, the higher the volume of ECMs get, the more the efficiency scores improve. This is expected since most lower-volume ECMs are small retrofits such as changing a fraction of building light bulbs. On the contrary, larger ECMs such as retrofitting HVAC systems or optimizing control sequence much bigger impact on benchmarking scores, with a rate noticeably higher than that of smaller retrofits.

Comparison of Energy Star and SEEF

SEEF and Energy Star show moderate correlation (0.69). This is expected since, for comparison purposes, we used similar variables in calculating the two metrics. The relationship between SEEF and Energy Star is also non-linear: Energy Star values tend to be higher than SEEF in general, and the difference becomes even higher after a certain threshold. A few buildings have considerably higher Energy Star scores than SEEF scores (highlighted in Figure 33). A closer analysis of the differences in Energy Star calculation against SEEF reveals more insights about the differences in scores.

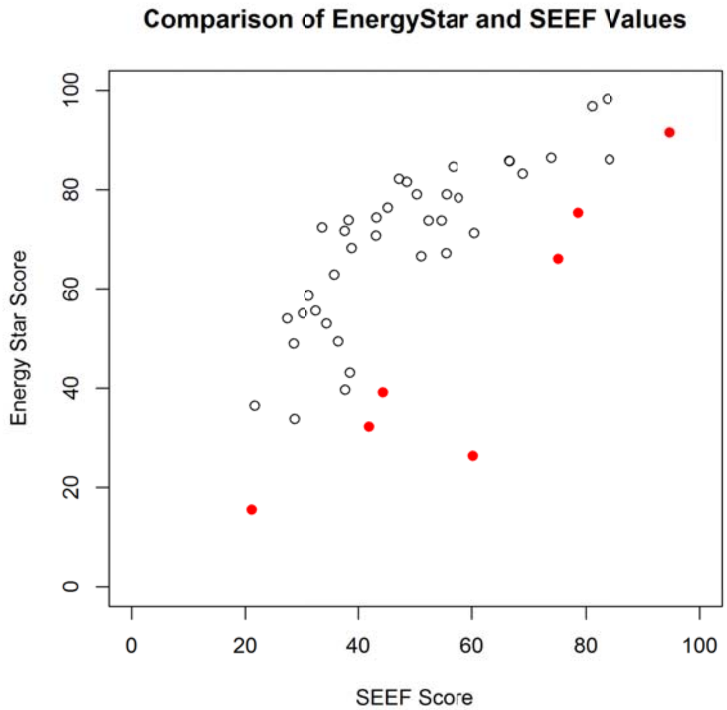


Figure 33 - Comparison of Energy Star and SEEF scores. Buildings with significant difference in SEEF and Energy Star scores are highlighted.

The difference in SEEF versus Energy Star scores is the result of how the two metrics are calculated: while Energy Star forms a parametric relationship between consumption and consumption factors, SEEF forms a non-parametric frontier to compare buildings. The parametric model makes several assumptions about the relationship of the energy factors and energy consumption. In this case, the discrepancy is the result of an assumption that Energy Star makes about the relationship of data center energy consumption and cooling load. In particular, Energy Star assumes that all data centers consume 0.92 times their computing power for cooling purposes. Such relationship between cooling power demand and computing power might be true for the industry in average; however, the amount of cooling power requirement can be considerably different from one building to another, depending on the climate and type of air conditioning system. Previous studies have reported cooling loads of 0.14 to 1.42 times the computer loads, just as a result of the difference in system types (US Department of Energy, 2011).

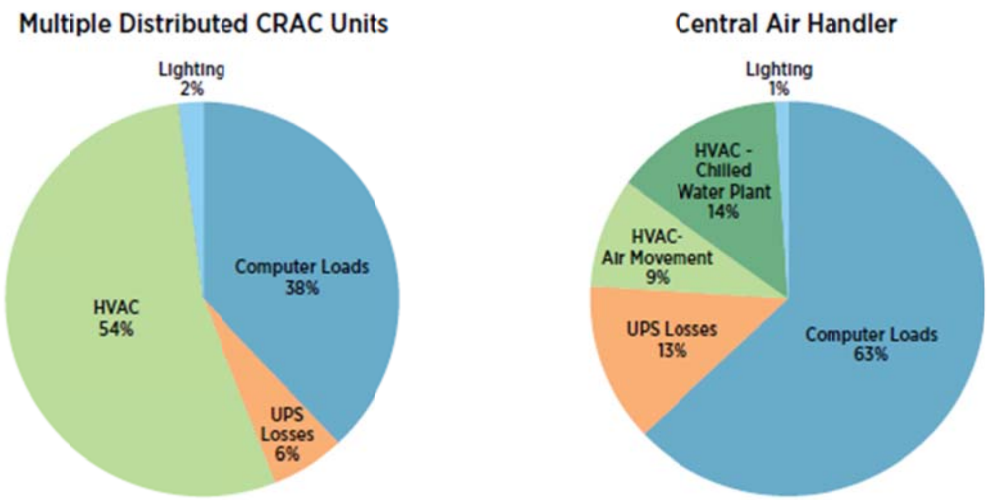


Figure 34 - Comparison of the breakdown of total data center electricity usage. Both data centers are operated by the same company, are located in adjacent buildings, and operate similar computer loads. The data center associated with the left plot uses a multiple-distributed unit system based on air-cooled CRAC systems, while the system on the right uses a central air handling system. [Image by Rumsey Engineers for US Department of Energy, 2011]

In our case study, since the buildings are located in a cold climate (Washington state in Pacific Northwest, US), and the HVAC system uses outside air for cooling data centers, the cooling load is much lower than the national average (closer to the 14% of computer load figure mentioned above). This drives up the Energy Star artificially high: the building is getting an unfair advantage for being able to use outside air to cool down its data centers. While the higher Energy Star score in this case acknowledges correctly that the building is using less energy, but it does not mean the building is more “efficient” or estimate its potential for further energy reduction.

Implications for continuous performance tracking

When benchmarking scores are used for continuous improvement of buildings, the increase or decrease in the scores offers feedback to practitioners on whether the energy efficiency actions were helpful. Thus, it is important to choose the benchmarking metric that encourages the proper change in buildings. A correlation analysis of SEEF and Energy Star scores of each building over time (36 values, one for each month), will reveal whether the two metrics increase or decrease together or not. A positive correlation implies that during the months that SEEF had a higher than average value, Energy Star also had a higher than average value. A negative correlation implies the contrary situation. Figure 35 shows the correlation of monthly SEEF and Energy Star values for each building of our analysis.

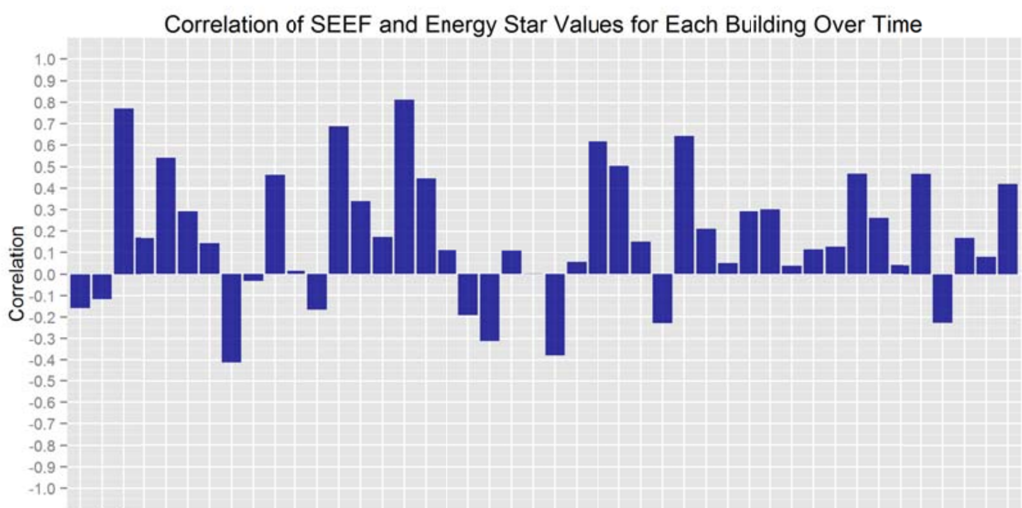


Figure 35 - Correlation of monthly SEEF and Energy Star scores for each building for the 36 months duration of the experiment. Each bar represents one buildings. Building names are removed to preserve anonymity.

As seen in Figure 35, ten buildings have negatively-correlated values of Energy Star and SEEF. A follow-up analysis reveals that the buildings with negative correlation have slightly more variation in their occupancy rates, compared to the buildings with positive correlation.

Comparison of Longitudinal Benchmarking against Cross-Sectional Benchmarking

Longitudinal benchmarking methods such as IPMVP measure the improvement of buildings compared to their pre-retrofit or normal states. As Figure 32 suggests, the IPMVP score is not correlated with any of the two cross-sectional benchmarking scores. This implies that whether a building is improving compared to its own baseline does not necessarily result in an improvement over the entire population. In many cases, the baseline model of the building is already low-performing, hence considerable improvement is needed before the building can be compared against the entire population. Thus, longitudinal and cross-sectional benchmark scores should be regarded independently.

Based on our validated data set of three buildings (Figure 32), high-volume ECMs improved IPMVP scores. However, we need more data to establish a general relationship based on this observation.

4.2 Effect of ECMs on Benchmarking Scores over Time

Energy Conservation Measures (ECMs) are implemented with the goal of improving the performance of buildings and reducing their energy waste. Actionable benchmarking of buildings provides feedback to energy managers on whether ECMs have achieved their goal of energy reduction or not. By actionable, we refer to the ability to infer specific actions and decisions to improve the energy efficiency of buildings. For example, benchmarking can reveal the features that lead to higher efficiency in a group of buildings, or retrofit projects that resulted in improved efficiency. Therefore, in this section we examine the relationship between ECMs versus different benchmarking scores. ECMs are measured using their "volume", which is the estimate of energy managers from potential effect of each ECM on energy consumption of the buildings.

To compare the correlation of benchmarking scores with ECMs, we tracked both the efficiency scores and month-over-month improvement of the scores and regressed them against the size of ECMs. As a metric to measure month-over-month improvement of scores, we used a log-ratio model. We explain the two models in the following sections.

Simple efficiency score model: the simple model analyzes the effect of ECMs on the efficiency score, assuming the following relationship holds between efficiency score and cumulative volume of ECMs:

$$e_t^k = f(m_t^k)$$

Where the variable definitions are the same as the log-ratio model.

For simple efficiency score model, 20 buildings (out of 25 with ECMs) showed significant correlation between the cumulative volume of ECMs and SEEF efficiency scores. Out of these 20 buildings, 16 had positive slopes of the regression model, implying that ECMs had a positive impact on efficiency scores of the buildings, while 4 buildings had negative slopes. Similarly, 20 buildings showed statistically significant correlation between ECM volume and Energy Star score, with 15 buildings having a positive correlation while 5 buildings having a negative correlation.

Continuous improvement model: The log-ratio model examines the effect of ECMs on continuous improvement of buildings. The model regresses the log of the ratio of subsequent efficiency scores on cumulative volume of ECMs.

The underlying assumption of the model is that the log-ratio can be explained as:

$$\log(e_{t+1}^{k} \;/\; e_{t}^{k}\;) \;= f(m_{t}^{k})$$

Where:

e_{t+1}^{k} Efficiency score of building k at time $t+1$

e_{t}^{k} Efficiency score of building k at time t

m_{t}^{k} Cumulative volume of ECMs performed on building k up to time t

The log-ratio model was statistically significant for only 5 buildings out of 35 buildings, and the coefficients of the models were very small. This implies that while the ECMs had a positive impact on the efficiency score, they did not significantly improve the continuous month-over-month improvement of buildings.

The type of ECM might have an effect on the slope of the efficiency score regression model. In our data set, buildings with a positive relationship between ECMs and SEEF score had more HVAC equipment retrofits, while buildings with negative relationship between ECMs and SEEF scores had more lighting ECMs. Such insights can help practitioners compare the effect of their energy management practices on their building portfolios in real time.

As Figure 36 shows, Control Sequence Revision is the largest-implemented ECM on the buildings (and second most frequent ECM).

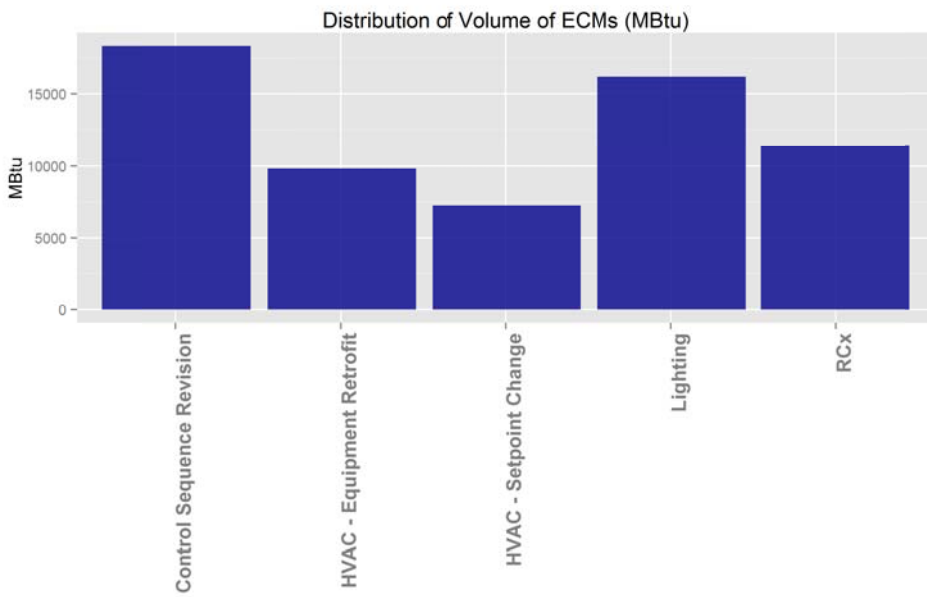


Figure 36 - Distribution of total volume of expected energy saved as a result of ECMs.

4.3 Sensitivity of Efficiency Scores to Operational Parameters of Buildings

To examine the sensitivity of benchmarking scores to changes in operational parameters of buildings, estimated the change in SEEF and Energy Star score of buildings following a 20% reduction and increase in three major operational parameters (headcount, number of computers, and data center load). For all cases, we assumed the energy consumption levels stay the same as before.

Results of SEEF sensitivity analysis show that, when energy consumption level is constant, an increase in number of computers has the most significant impact on improving SEEF score, while a decrease in lab load, headcount and number of computers have almost equal but negative impact on SEEF scores. In terms of size of effect, the positive effect of number of computers is larger in magnitude than the negative effect of both headcount and lab load. This suggests that, for example, if the number of computers increases by 20%, to maintain the efficiency scores, lab load should decrease by more than 20%. Overall, the sensitivities range from 0.6 to 1.4; the efficiency scores may change up to 40% as a result of a 20% change in one operational value. SEEF is more sensitive to headcount, while Energy Star is more sensitive to data center load. For example, if a decrease in lab load is accompanied by an increase in occupancy (for instance, a change of building usage from developer to administrative), the energy managers need to reduce the energy consumption to keep the benchmarking scores at pre-change levels.

4.4 Effect of External Parameters on Efficiency Scores

Other than operational factors and ECMs, external factors such as outside temperature also impact the energy consumption of buildings. Both SEEF and Energy Star normalize for outside temperature; thus, it is likely that any change in the efficiency score of the buildings across different temperatures is due to the efficiency of the building systems and its structure (Kavousian et al., 2013). For example, Figure 37 shows three different scenarios for the relationship between Total Degree Day (sum of CDD and HDD) with efficiency scores. In each case, the change in efficiency over different outside temperature ranges imply the efficiency of the building system, and provides insights into which systems work best under which conditions.

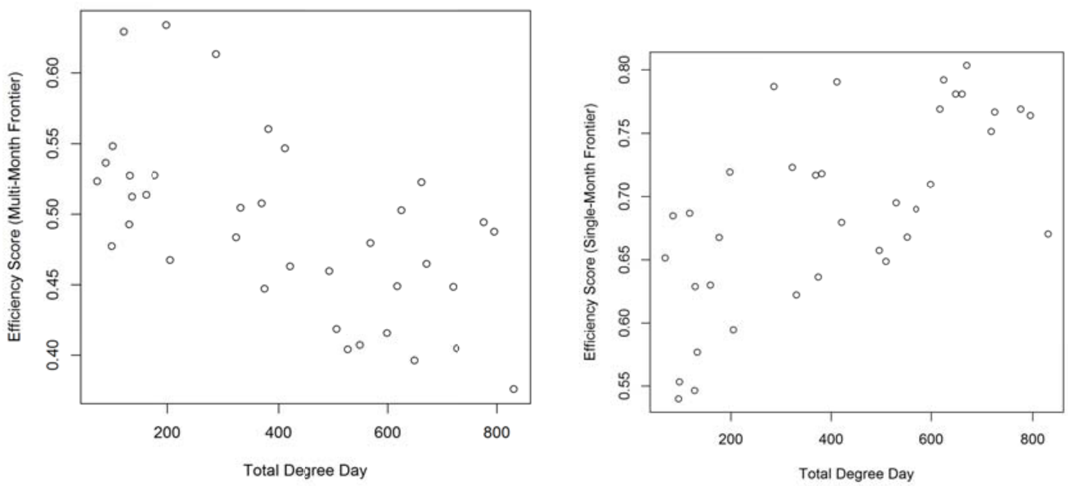


Figure 37 – Correlation of SEEF scores with Total Degree Days (sum of Cooling and Heating Degree Days). Different system types perform differently under a range of outside air temperature. (a) left: for buildings with system type D, efficiency declines with increase in outside temperature; (b) right: for buildings with system type F, efficiency improves with more extreme weather.

Table 11 shows the sensitivity values for both SEEF and Energy Star scores. Energy Star correlations are much lower than those of SEEF, mostly due to Energy Star normalizing for weather. Furthermore, all correlations are positive, implying that Energy Star has a bias to over-compensate for outdoor temperature. SEEF correlations range from -0.70 to 0.67, implying that SEEF is able to identify buildings that become more efficient with change in outside temperature.

Table 11 - Sensitivity of buildings to outside temperature, by system type.

System Type	Average SEEF Efficiency Score	Correlation of SEEF Score with Total Degree Days	Average Energy Star Efficiency Score	Correlation of Energy Star Score with Total Degree Days
A (Central Chilled Water Plant, not located in the building; Chilled Water AHU's, Chilled Water Labs)	64.47	-0.38	74.87	0.05
B (Chillers in building; Chilled Water AHU's, Chilled Water Labs)	69.48	0.08	66.96	0.13
C (Chillers in building; Direction Expansion AHU's, Chilled Water Labs)	54.40	0.37	58.33	0.10
D (Direction Expansion AHU's, Direction Expansion Labs)	65.86	-0.11	48.97	0.07
E (Chillers in building, Direction Expansion Labs)	69.08	0.67	73.75	0.12
G (Underfloor Air Distribution Building with Labs)	86.07	-0.70	81.75	0.06
H (Chilled water serving building, No Labs)	44.35	0.01	66.53	0.15
I (Direction Expansion AHU's, No Labs)	54.38	-0.46	44.45	0.04

4.5 Effect of Building Characteristics on Efficiency Scores

Several building characteristics impact energy efficiency. For example, Figure 38 shows SEEF scores of all buildings, averaged over groups of the same system type, presented over 36 months of the experiment (starting from July 2008 through June 2012). As seen in the figure, buildings with system type G (underfloor air distribution) have the highest efficiency scores, while buildings of type H (chilled water serving the building) have the lowest efficiency scores.

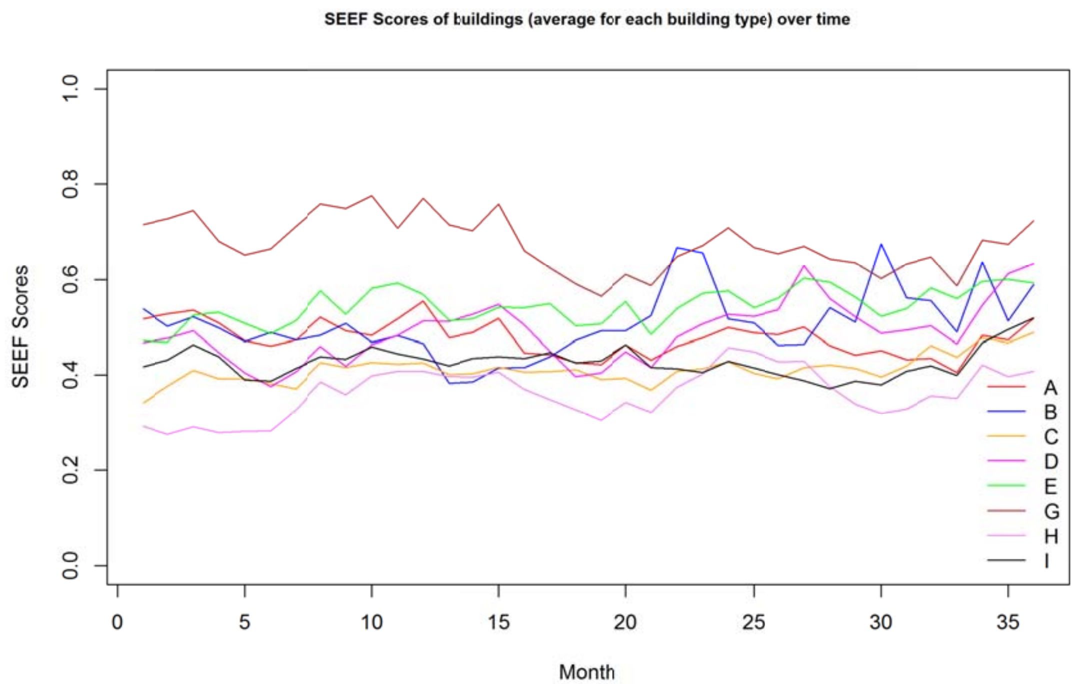


Figure 38 - SEEF scores of buildings, averaged over building system types, over the course of the experiment. For a description of system types, see Table 11.

Figure 39 shows the same plot, for Energy Star scores. SEEF and Energy Star agree on the most efficient system (type G); however, system type H which is the least efficient system as measured by SEEF, is mid-to-top performer in Energy Star. In both efficiency measures, building with system type I are less efficient. Overall, Energy Stat shows less variation than SEEF.

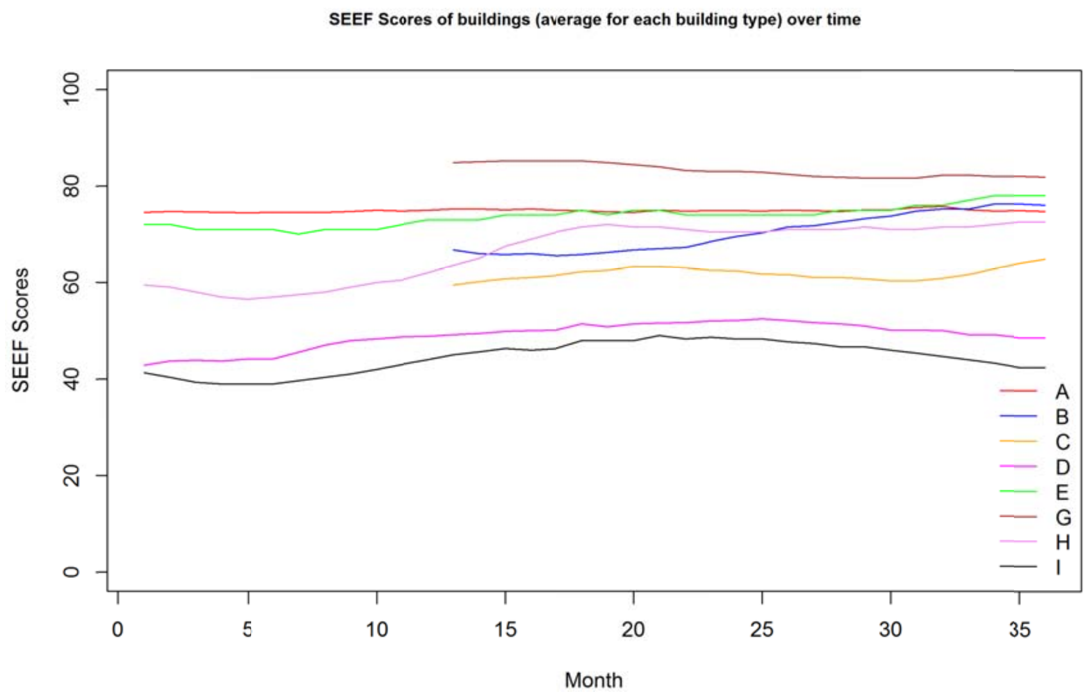


Figure 39 – Energy Star scores of buildings, averaged over building system types, over the course of the experiment.

5 Discussion

Energy Star, SEEF, and IPMVP Option C each measure the efficiency of buildings differently. Energy Star and SEEF are cross-sectional benchmarking methods; they compare a building with other buildings of similar characteristics, while normalizing for most important factors such as weather. Energy Star uses industry averages of the US buildings to create a parametric model of the energy consumption of buildings, and compare measured performance of the building with the model. SEEF compares the buildings by its peer group buildings by forming an efficient frontier and comparing the population against the frontier. IPMVP Option C is a longitudinal benchmarking method, which compares the building with itself over time. Hence, it is expected that the three methods would not provide identical results. Instead, each method should be used for its intended use and under proper circumstances. However, similarities or dissimilarities between performance measures still provide insights into the efficiency of buildings.

The choice of the benchmarking metric has several significant effects. Each performance metric encourages a specific behavior. For example, in our case study, the energy managers used to solely track Energy Star scores. As a result, the Energy Conservation Measures (ECMs) that were associated with increased Energy Star score were more frequently and widely implemented than other ECMs.

We observed a dual-mode relationship between the volume of ECMs and improvement in benchmarking scores: the volume of ECMs should reach a certain threshold to expedite the energy efficiency improvements. Smaller ECMs such as lighting retrofits improve the benchmark scores; but the rate of increase in benchmarking score increases considerably when deeper ECMs such as HVAC retrofits are implemented.

Different benchmarking scores also respond differently to changes in the operational characteristics of the building. For example, Energy Star is more sensitive to changes in the data center load, while SEEF is more sensitive to changes in number of computers. Longitudinal models are even more sensitive to changes in operational factors that drive energy consumption. Hence, when a major change in the building happens, an entirely new benchmark needs to be formed. For practitioners, an understanding of the factors that impact benchmarking scores is essential. Such insights will help energy managers to turn the benchmarking scores into actionable feedback.

While most benchmarking methods normalize the energy consumption for outside temperature, it is still likely that a building is less efficient during a certain season—depending on the efficiency of its cooling versus heating system. Benchmarking methods such as Energy Star that create the model based on a full-year cycle of the building are less likely to identify such changes in the seasonal efficiency of buildings. However, methods such as SEEF that develop separate comparative models for each month can be useful in unraveling such trends.

Finally, structural and mechanical features of the building have a considerable impact on efficiency. We observed considerable differences in the efficiency scores of buildings with different system types. The effect of building systems also depends on their temporal patterns.

Ideally, the goal of benchmarking should be to continuously monitor and improve the energy performance of buildings. Previous studies have shown the benchmarking tools such as Energy Star have been used as public tools in the past. Even in such cases, simply measuring the performance of buildings is associated with improved performance. For example, a 2012 study by the EPA examined 35,000 buildings that consistently used the Energy Star Portfolio Manager from 2008 to 2011 to track energy use. The study found that energy consumption in those buildings decreased by an average of 7 percent (EPA, 2012).

5.1 The Choice of Benchmarking Method

The choice of the benchmarking method depends on several factors, including organization's priorities, energy mangers' goals from benchmarking, and availability of data. For example, since Energy Star uses industry averages to predict energy consumption of buildings, it is best suited for buildings with characteristics similar to average buildings. However, once the building is un-conventional or in a climate that is not well-represented in Energy Star's calculations, then the scores' accuracy considerably decline.

SEEF demonstrated more sensitivity to operational factors of the building, external factors, building systems, ECMs, and an overall faster feedback rate. Thus, when the data for SEEF analysis are available, and when the goal of the analysis is continuous and fast feedback system for energy efficiency improvement, SEEF is a more viable option. SEEF particularly responds to the need for a metric that has the following features: (a) is tailored to the organization rather than based on industry-wide averages; (b) does not assume certain relationship between energy factors and energy consumption (i.e., is non-parametric); and (c) enables cross-sectional benchmarking among a finite building stock. Furthermore, frontier methods offer a new paradigm in improving efficiency: the efficiency of a firm can be improved both by reducing the consumption for a given output and increasing output for a given level of consumption. Frontier methods enable comparing to best practice rather than average practice; it is often more interesting to learn from the best rather than to try to improve upon mediocre performance. Finally, by defining the efficiency as the ratio of the energy consumption of the most efficient comparable building to the actual consumption of the building of interest, SEEF enables calculating the potential gain from increasing efficiency.

Longitudinal benchmarking methods are best suited to examine the effectiveness of certain one-time ECMs that were performed at a particular point in time.

5.2 The Choice of Energy Retrofits

Another important decision that energy managers face is the choice of energy retrofit(s) to implement on certain buildings to improve their energy efficiency. Benchmarking scores provide insights on the buildings that are less efficient, and can also be used to examine drivers of energy efficiency, as seen in the paper. In some cases, several competing retrofits are available, and in some other cases budget constraints limit the number of options for energy managers. In such cases, a follow-up analysis similar to the log-ratio analysis or simple regression analysis explained in this paper can help energy managers to select the projects that improve the energy efficiency scores of their buildings the most, while staying within budget.

6 Conclusions and Future Work

The choice of benchmarking method has a profound impact on energy performance of buildings. Different benchmarking tools are better suited for different purposes. This paper provides an overview of three most significant categories of benchmarking tools, and compares them using a data set of 120 buildings on the campus of a technology company.

Energy Star is a general tool that will continue to be used by practitioners, but mostly for reporting and public relations purposes. While the generality of Energy Star makes it very useful for comparison of a wide range of buildings, it is not specific enough to properly pinpoint problem areas in buildings. On the other end of the specificity spectrum, longitudinal benchmarking methods such as IPMVP Option C develop models that are specific to buildings of interest, and help uncover deviations from the normal operations. However, since longitudinal analysis is normally isolated to individual buildings, it does not provide a method to estimate the potential for energy reduction, based on lessons learned or using different systems or strategies. Therefore, longitudinal benchmarking is the most useful as a diagnosis tool, or as a measurement and verification tool for retrospective analysis. Finally, frontier methods such as SEEF are suitable for cases where a specific relationship between the energy factors and energy consumption cannot be made, or where non-conventional energy factors need to be included in the model. SEEF also requires availability of data from a small to mid-size population of similar buildings. Our analyses also showed that frontier methods are more responsive to changes in internal and external variables that are associated with energy consumption. Hence, frontier methods are better suited for continuous monitoring and improvement of energy performance of buildings.

Data availability can influence the choice of benchmarking, too. This paper used monthly data as the common denominator of data sets used by energy managers. However, more detailed data are needed for more detailed analysis, especially if more rigorous analysis of methods such as longitudinal benchmarking is desired. Additional data that can be useful for benchmarking include interval (e.g., hourly) consumption data and Building Management System (BMS) data. Our future work will uncover the insights that energy managers can achieve from such additional data. Our intuition is that the added cost of collecting interval data and setting up a system for easy access to BMS data will be more than paid off by added efficiency gains from such efforts.

7 References

Ang, B., Zhang, F. (2000) A Survey of Index Decomposition analysis in energy and environmental studies. Energy, 25, 1149-1176.

California Energy Commission (2014) Non-Residential Building Energy Use Disclosure Program (AB 1103). Last accessed on 2/14/2014 via http://www.energy.ca.gov/ab1103/

Carnes, D., Hunn, B., Jones, J. (2008) The Use of Data Envelopment Analysis for Evaluating Building Energy Consumption in Terms of Productivity. Proceedings of 1998 Summer Study in Building Energy Efficiency, American Council for an Energy Efficient Economy, Washington, DC.

Deutsche Bank and Living Cities. 2011. Recognizing the Benefits of Energy Efficiency in Multi-Family Underwriting. New York.

EPA (Environmental Protection Agency). (2011). ENERGY STAR Performance Ratings Technical Methodology for Office, Bank/Financial Institution, and Courthouse. Environmental Protection Agency. Retrieved from www.energystar.gov/ia/business/evaluate_performance/office_tech_desc.pdf

EPA (Environmental Protection Agency). (2011). ENERGY STAR Performance Ratings Technical Methodology, General Overview. Environmental Protection Agency. Retrieved from http://www.energystar.gov/ia/business/evaluate_performance/General_Overview_tech_methodology.pdf

Farese, P., Gelman, R., Hendron, R. (2012) A Tool to Prioritize Energy Efficiency Investments. National Renewable Energy Lab (NREL) Technical Report NREL/TP-6A20-54799, August 2012, Golden, CO, USA.

Ferrier, G., Hirschberg, J.G. (1992) Climate Control Efficiency. Energy Journal, 13 (1), 37-54.

Grosche, P. (2009). "Measuring residential energy efficiency with DEA." Journal of Productivity. Analysis , 31 , 87–94.

Huntington, H. (1994) Been Top Down So Long It Looks Like Bottom Up to Me. Energy Policy, 22 (10), 833-839.

IPMVP 2001. IPMVP Committee, International Performance Measurement & Verification Protocol: Concepts and Options for Determining Energy and Water Savings, Vol. 1, U.S. Dept. of Energy, DOE/GO-102001-1187, 86 pp., January.

Kavousian, A., Rajagopal, R., Fischer, M. (2013) Determinants of residential electricity consumption: Using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants' behavior. Energy, Volume 55, 15 June 2013, Pages 184-194, ISSN 0360-5442, http://dx.doi.org/10.1016/j.energy.2013.03.086.

Kavousian, A. and Rajagopal, R. (2014). "Data-Driven Benchmarking of Building Energy Efficiency Utilizing Statistical Frontier Models." Journal of Computing in Civil Engineering, 10.1061/(ASCE)CP.1943-5487.0000327.

Lee, W., Lee, K. (2009) Benchmarking the Performance of Building Energy Management Using Data Envelopment Analysis. Applied Thermal 29 (16), 3269-3273, doi:10.1016/j.applthermaleng.2008.02.034.

Onut, S. , and Soner, S. (2006). "Energy efficiency assessment for the Antalya Region hotels in Turkey." Energy Build. , 38 (8), 964–971.

Peterman, A., Kourula, A. Levitt, R. (2012) A roadmap for navigating voluntary and mandated programs for building energy efficiency. Energy Policy, 43 (1) 415-426

Porter, T., Weaver, K., Vaughn, K. "Effective Implementation of "IPMVP Option C- Whole Building Measurement" Measurement and Verification Plans". Proceedings of the Fifth International Conference for Enhanced Building Operations, Pittsburgh, Pennsylvania, October 11-13, 2005.

Shapiro, I. (2011) Ten Most Common Problems with Energy Audits. ASHRAE Journal, Debruary 2011, 26-32.

US Department of Energy, Office of Energy Efficiency and Renewable Energy (2011) Best Practices Guides for Energy-Efficient Data Center Design. Published by Federal Energy Management Program. http://www1.eere.energy.gov/femp/pdfs/eedatacenterbestpractices.pdf

US Department of Energy (2014) Building Energy Databook. Last accessed on 2/10/2014 via http://buildingsdatabook.eren.doe.gov/TableView.aspx?table=1.1.3

Zhou, P., Ang, B., Poh, K. (2008) A Survey of Data Envelopment Analysis in Energy and Environmenta Studies. European Journal of Operational Research. 189, 1-18.

Chapter 6

Summary of Contributions, Conclusions, Limitations, and Future Work

My doctoral research contributes to buildings' energy efficiency ranking by introducing computational methods to (a) remove from the efficiency estimation the effect of external and internal factors that are outside the control or interest of users to change, and (b) incorporate stochasticity of energy consumption in energy efficiency estimation. Both (a) and (b) considerably influence the energy consumption over time (see Chapter 2 and Chapter 4 for evidence), and can highly skew the energy efficiency rankings. To estimate the stochasticity of energy efficiency, I introduce a method called Stochastic Energy Efficiency Frontier (SEEF) that extends existing applications of efficient frontiers in energy efficiency by calculating confidence intervals for energy efficiency estimations based on repeated measurements. SEEF reduces potential biases in energy efficiency ranking that are introduced due to fluctuations in energy consumption.

Ranking building energy efficiency poses challenges, some of which are caused by the complexity of modeling building energy consumption (see Chapter 2 and Chapter 4 for more details on these challenges). A summary of these challenges are:

- Energy consumption characteristics (e.g., daily base load, peak load, load shape) vary widely across buildings, both in magnitude and patterns.
- Different energy consumption characteristics are explained by different factors. For example, drivers of base load are different from drivers of peak load, and

drivers of morning consumption are different from factors that influence evening consumption.

- Demographics weakly predict energy consumption characteristics.
- Limited building features data are available, but with more large-scale databases for real estate business and the growth trend in utilities' data collection efforts, lack of building features data is becoming less limiting.
- It is hard to keep track of un-occupancy states with aggregate data.
- As a result of all the above points, explaining what drives energy consumption is not an easy task. It is possible to write an algorithm that works great on 20% of buildings, but there are so many special conditions and nuances that cannot be explained by simple methods and limited data.

In addition to the above, challenges specifically posed by building energy efficiency ranking include:

- Because energy consumption is highly variable, it is hard to make a general case about a building being more efficient than another one using limited data.
- A tradeoff exists between consuming energy and enjoying building services. The goal of most programs is improving the efficiency of existing operations, rather than reducing the ability of users to perform activities or operations that they desire to do. I.e., the assumption is that users will continue to use their existing services (e.g., keep the same number of appliances or air-conditioned area).

- The choice of metric in building energy efficiency ranking is very important, because the wrong metric will encourage the wrong behavior.

The challenges summarized above are partly due to lack of data in traditional settings. For example, monthly billing data in residential buildings settings is not adequate to identify the variation of energy consumption or to differentiate between consumption characteristics. With interval data becoming more accessible through the advent of Advanced Metering Infrastructure (AMI) and smart meters, lack of data has become less problematic. However, methods to analyze smart meter data to better estimate energy efficiency of different buildings are still missing.

1 Contributions and Impacts

My research extends frontier models by (a) formalizing a process to account for the effect of climate and lifestyle in energy efficiency ranking, and (b) improving how the efficient frontier is defined by calculating confidence intervals for efficiency estimates. The contributions of my dissertation to the field of data-driven energy efficiency include:

a) Demonstrating the value of frontier methods in energy efficiency analysis;
b) Developing a method to normalize consumption for weather, lifestyle and building services, and incorporating the remaining fluctuations of energy consumption in energy efficiency estimates;

c) Improving how the energy efficient frontier is defined, by developing a method to incorporate the stochasticity of energy consumption in energy efficiency estimates using measured data as opposed to resampling methods.

In this way, SEEF improves upon existing building benchmarking methods by improvements upon:

- Scalability: SEEF supports several independent outputs (services) and can be extended to several inputs (e.g., electricity and gas consumption), as demonstrated in the case studies.
- Verifiability (the ability to perform statistical tests): By estimating a confidence interval for efficiency values, one can make statistical inferences when comparing efficiency of buildings.
- Generality: SEEF is non-parametric, which means it does not make any assumptions about the relationships between outputs and inputs.

2 Limitations and Future Work

In the case studies presented in this thesis, the data on appliance ownership of residential buildings were available. However, in real-world scenarios such as large-scale utility energy efficiency planning, individual appliance data are not available. Smart meter data can be used to infer appliance ownership from electricity consumption profiles (Becker et al., 2013). I am currently collaborating with a team of researchers at Stanford and ETH Zurich to integrate appliance ownership estimation methods with SEEF, to enable ranking of building energy efficiency using merely smart meter data.

Furthermore the effect of Level of Service (LOS) definition on ranking results can be studied in more detail. In the residential case studies, I chose the Appliance Ownership Index (AOI) since appliances are normally the most important service that building occupants receive from the building (after accounting for air conditioning through weather normalization). In the case of commercial buildings, LOS can have a broader set of definitions. For example, in the case study presented in chapter 5, I chose LOS based on occupancy, data center load (as a proxy for level of occupancy in the building), and outside temperature since these metrics were of highest importance to the building owner. However, as explained in the thesis, different researchers have used different definitions for LOS. Further research is needed to analyze the impact of choice of LOS on future performance of buildings and success of energy conservation measures.

Further studies are needed to examine the form of the frontier across different populations. For example, a linear frontier implies a constant returns to scale technology. Therefore, if a population of buildings show a linear frontier, it implies that increasing the scale of services (in the residential context this means adding more appliances to existing buildings or moving to larger buildings), does not considerably reduce the efficiency of energy consumption. Thus, one can make a hypothesis that the farther from linear the frontier is, the more “wasteful” the larger buildings are. Further case studies and analyses are needed to validate or reject this hypothesis.

As explained in Chapter 4, sparsity of data from high-LOS buildings can produce biases in energy efficiency estimates. In particular, a small number of observations from high-LOS buildings may result in more buildings being closer to the frontier, hence introducing an upward bias in energy efficiency estimations. The sparsity also implies that the confidence interval for the efficient frontier may be narrower, because a limited number of buildings form the frontier throughout bootstrap iterations. More data are needed to estimate the upward bias in efficiency estimates and the downward bias in confidence intervals. But an alternative solution is to use perturbation methods to simulate more high-LOS buildings. In this method, existing high-LOS buildings will be copied with a random added error term to form new simulated buildings. The distribution of the error term will be estimated based on the distribution of energy consumption and LOS at denser parts of the energy-LOS plot.

Finally, to close the loop on energy efficiency estimation and improvement, an extension to my doctoral research is to develop a recommendation system that tracks energy consumption and LOS of buildings over time and presents real-time, actionable feedback to users and energy managers. For example, SEEF enables the identification of unusually high or unusually low consumption values (after removing the effects of immutable factors). Thus, a potential feature of the recommendation engine could be to generate alerts for users when their energy consumption exceeds normal values.

Based on my findings and validations of my thesis, future studies on energy efficiency ranking should use the following lessons learned:

- Energy consumption is a stochastic variable, and so is energy efficiency. Using aggregate numbers to compare energy consumption of users can result in a biased view about their energy efficiency, affected by periods of abnormally high or abnormally low consumption.
- Several factors impact energy consumption. Some of these factors are within the control and interest of the users to change, but many are not. An actionable ranking method accounts for immutable energy drivers and finds the correlation of energy efficiency with changeable factors.
- Energy consumption is driven by services that the building offers. A view that compares energy consumption based on merely structural factors (e.g., energy per square foot) can be very misleading, as different buildings with similar structures can provide very different levels of service.
- Frontier methods specify a boundary for the relationship between factors such as energy consumption and LOS, when an ordinary regression does not adequately represent their relationship.

While my research addressed all points mentioned above, it did not examine the scenarios for inferring or simulating more data from available smart meter data. In particular, two applications for immediate development are inferring appliance ownership data from smart meter data and simulating more users at higher LOS values to identify and remove potential statistical bias of efficiency and confidence interval at high LOS values. Through these extensions, SEEF can be expanded into a building energy efficiency ranking and recommendation engine which can be used by energy users or energy managers to track and improve energy efficiency of their buildings.

3 References

EIA (2014) Peak-to-average electricity demand ratio increasing in New England and many other US regions. Latest accessed on 05/10/2014 via: http://www.eia.gov/todayinenergy/detail.cfm?id=15051

Beckel, C., Sadamori, L., Santini, S. (2012) Towards automatic classification of private households using electricity consumption data. Proceedings of the Fourth ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings. Pages 169-176.

Kavousian, A. and Rajagopal, R. (2014). "Data-Driven Benchmarking of Building Energy Efficiency Utilizing Statistical Frontier Models." J. Computing in Civil Engineering, 10.1061/(ASCE)CP.1943-5487.0000327.

Kavousian, A., Rajagopal, R., Fischer, M. (2013) Determinants of residential electricity consumption: Using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants' behavior. Energy, Volume 55, 15 June 2013, Pages 184-194, ISSN 0360-5442, http://dx.doi.org/10.1016/j.energy.2013.03.086.

Kavousian, A., Ho, J., Win, L., Yip, H. (2013) Deploying Battery Storage in Commercial Buildings in the US: Opportunities and Challenges. Stanford Energy Journal, June 2013, 3(1).

Kissock, J., Haberl, J., Claridge, D. (2003) Inverse Modeling Toolkit, Numerical Algorithms. ASHRAE Transaclions, 109 (2).

National Renewable Energy Lab (NREL) by Polly, B., Kruis, N., Roberts, D. (2011) Assessing and Improving the Accuracy of Energy Analysis for Residential

Buildings. NREL Technical Report. NREL, Colorado, US.

http://www.nrel.gov/docs/fy11osti/50865.pdf