Whose fault is it anyway? Evaluating the Energy Efficiency Gap in Commercial Buildings & Measuring Energy Savings Associated with Fault Detection and Diagnostics

by

Danielle Dahan

Bachelors of Science, Engineering

Brown University, 2011

SUBMITTED TO THE INSTITUTE FOR DATA, SYSTEMS, AND SOCIETY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN TECHNOLOGY AND POLICY AT THE MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2017

Signature of Author ...... Signature redacted ....
Institute for Data, Systems, and Society
August 31, 2017

Certified by ..... Signature redacted ....
Christopher Knittel
Professor, Applied Economics
MIT Sloan School of Management
Thesis Supervisor

Accepted by................ Signature redacted ....
Munther Dahleh
Professor, Electrical Engineering and Computer Science
W. Coolidge Director, Institute for Data, Systems, and Society

Whose fault is it anyway? Evaluating the Energy Efficiency Gap in Commercial Buildings & Measuring Energy Savings Associated with Fault Detection and Diagnostics

by

Danielle Dahan

Submitted to the Institute for Data, Systems, and Society on August, 31 2017 in Partial Fulfillment of the Requirements for the Degree of Master of Science in Technology and Policy

According to the International Energy Agency, energy efficiency programs make up 72% of global greenhouse gas abatement strategies. However, there is extensive literature that shows compelling evidence for an "energy efficiency gap" in which expected energy savings from energy efficiency programs are not realized. Due to the importance of energy efficiency in global climate mitigation, as well as the significant federal, state, and local budgets for energy efficiency, there is a clear need for further research in this domain to evaluate the energy efficiency gap and prioritize methods for reducing the gap. Further, there is significantly less research on the gap as it applies to commercial buildings; the majority of research does not take advantage of advancements in available statistical modeling techniques; and there is very limited research evaluating the gap as it applies to the new field of fault detection and diagnostics (FDD). With FDD, building owners are able to closely monitor on an ongoing basis any faults that begin to occur in a commercial building that can waste energy and lead to the gap in energy efficiency. However, there has been very little research evaluating these systems in real buildings and calculating the energy efficiency impact. This thesis proposes and tests a modeling approach using novel machine learning algorithms to estimate counterfactual energy usage in real buildings and calculate the energy efficiency savings associated with an existing FDD system.

In this thesis, I propose a modeling technique using novel machine learning algorithms to estimate counterfactual energy usage of commercial buildings. I take advantage of high-frequency 15-minute interval electricity, chilled water, and steam energy usage data over several years in four campus buildings. I then compare the accuracy of these models applied to brand-new data using three different machine learning modeling techniques, the Lasso Model, Ridge Regression, and an Elastic Net Model. Finally, I applied these models to 8 time periods in which the existing FDD system identified a fault, thus isolating the energy impact of the fault. With this approach, I found that each of the three modeling techniques outperformed the other two techniques in at least one of the models, indicating that there is

likely a benefit from using three approaches in building energy modeling. Further, I found that the models are likely able to isolate the energy increase associated with these faults, with some models yielding a higher confidence level than others. In addition to the overall average increase in energy, the faults showed consistent results in the daily load profile shifts after the fault occurred. Overall, the faults yielded monthly energy cost increases of $800-$1600 each. This methodology could therefore be used in more buildings and with different types of fault detection diagnostics systems to better evaluate the benefits of FDD software across applications. By using this method more extensively, we can better inform policy that can in turn aim reduce the energy efficiency gap in commercial buildings.

Thesis Supervisor: Christopher Knittel
Title: Professor of Applied Economics

Acknowledgements

I am so grateful to so many people for all of the amazing support I have received during my time in the Technology & Policy Program (TPP) and in the Center for Energy & Environmental Policy Research (CEEPR) at MIT. TPP, CEEPR, and MIT at-large have provided amazing communities for me to grow my career and develop my skills over the last two years. Below are just a few of the people that I am greatly indebted to for all of their support.

First, my thesis research supervisor, Professor Christopher Knittel has taught me so much during my time at MIT. He has helped me develop my research from a vague idea to something I am excited and proud of. His expertise in energy evaluation projects, economics and statistics has helped me tremendously in my research. Not only is Professor Knittel a leader in the industry, but his excitement in this work shines through every speech, presentation, or conversation he has. His interest and excitement in the material is almost contagious, and I left each review meeting more energized and excited than before. Joshua Hodge, my second research supervisor, has helped me on several research projects in CEEPR and provided me with multiple opportunities to present my research and research ideas. Joshua has been extremely supportive and accessible for both my research and my career development. I greatly enjoyed many conversations in his office about the energy industry and academia world, and I learned so much insight from his many years in the field. Manuel Mundo is an undergraduate that worked in CEEPR during the Spring and Summer of 2017 funded by the MIT Undergraduate Research Opportunity Program (UROP) Office. Manuel has been an amazing help in my research and thesis work. He has a very clear passion for machine learning and statistical analyses, and he was eager to always read as much as possible and learn as much as he could on the topics that he worked on.

Of course, I am very grateful for the support I was able to receive from MIT Facilities and the MIT Office of Sustainability. I met with many different people within MIT Facilities over the last two years to better understand the systems, collect data, and receive answers to many of my questions. Wade Berner, Ken Packard, Mark Mullins, and Josh Raines each provide me with many different resources. The MIT Office of Sustainability was able to fund my last two months of research and provide an additional level of support.

Further, Tony, Fannie, and Dimonika have been a huge support in the CEEPR office, from managing the common spaces; to helping me set up my workstations and meetings; to many enjoyable conversations in the office. Arun, Leila, Manuel, Danielle W., Beth, Emil and Katie helped to further create an enjoyable atmosphere in the office and in many fun lunches and teatime hours. The TPP team provided a wonderful community with so many different fascinating skills and interests in one group. Paige and Esther helped answer so many of the questions I had on my research, from methodology to formatting suggestions. Barb and Frank allowed me to pop by any time to answer my many questions and helped me stay focused and organized throughout my career at MIT.

My partner, Joe has helped me talk through different aspects of my coursework and research ideas; allowed me to take over the living room and dining room with hundreds of scattered research papers; and cooked me my favorite vegetarian meals so I could work longer hours during my "homestretch" period. My dog, Basil, has grown up during my time at MIT, from a four-month-old puppy chewing all of my homework assignments, to the perfect late-night work partner cuddled by my side. Finally, I am so lucky for all of the support and advice my parents have given me over the years, prioritizing education from a young age. My dad helped me develop my interest and skills in the math and sciences, and my mom showed me how a woman can champion the business world and be whoever she wants to be, despite the lack of women in both of our fields. My parents have given me everything I could ever ask for, opening a world of opportunities for me at every turn.

Thank you so much to all who have supported me through my education, my Master's degree, and in my next steps.

Table of Contents

List of Figures

List of Tables

1 The Energy Efficiency Gap

1.1 Introduction

Energy efficiency programs account for a significant portion of most climate action plans around the globe. In fact, the IEA World Energy Outlook of 2011 identified energy efficiency programs to make up 72% of greenhouse gas abatement by 2020 and 44% of greenhouse gas abatement by 2035, as seen in Figure 1.1.[1] Energy efficiency opportunities are seen as the "low hanging fruit" of abatement programs, as they often provide a lower upfront cost than other strategies and include high rates of return. However, there has been a lot of literature surrounding the "energy efficiency gap," which suggests that energy efficiency programs do not achieve their full potential of energy savings. There is extensive literature that defines the gap and attempts to explain why there are so many missed opportunities in energy efficiency. This chapter outlines reasons for the energy efficiency gap, existing literature that attempts to explain it, and the need for new research in this domain.

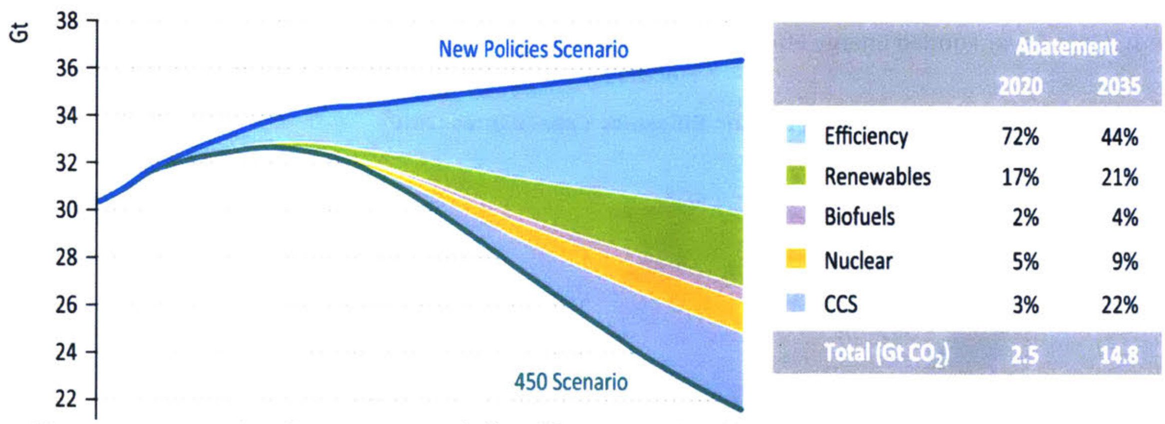


Figure 1.1 International Energy Agency's "World energy-related CO2 emissions abatement in the 450 Scenario relative to the New Policies Scenario" Source: IEA 2011

1.2 Energy Efficiency in the United States

In addition to the nationally determined contributions (NDC), there is a significant amount of international, federal, and state regulation that promote the use of energy-efficient products and installation in the United States. In addition, many utility companies in the United States provide economic incentives for energy efficiency investments in commercial, industrial and residential

[1] International Energy Agency. 2011. "World Energy Outlook 2011." <https://www.iea.org/publications/freepublications/publication/WEO2011_WEB.pdf>

buildings. These incentives are typically regulated by the state's department of public utilities. Figure 1.2 shows the sharp increase in total utility funding in the U.S. for energy efficiency from 2007 to 2011, according to the Institute for Electric Efficiency.[2] Figure 1.3 shows energy efficiency incentive variation by state.[3]

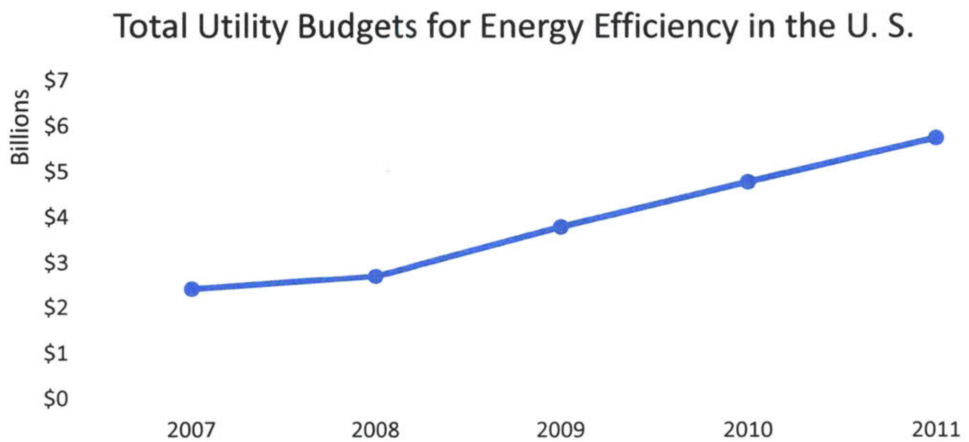


Figure 1.2 Rate Payer Funded Energy Efficiency Budgets in the U.S. from 2007-2011. Source of data: Institute for Electric Efficiency

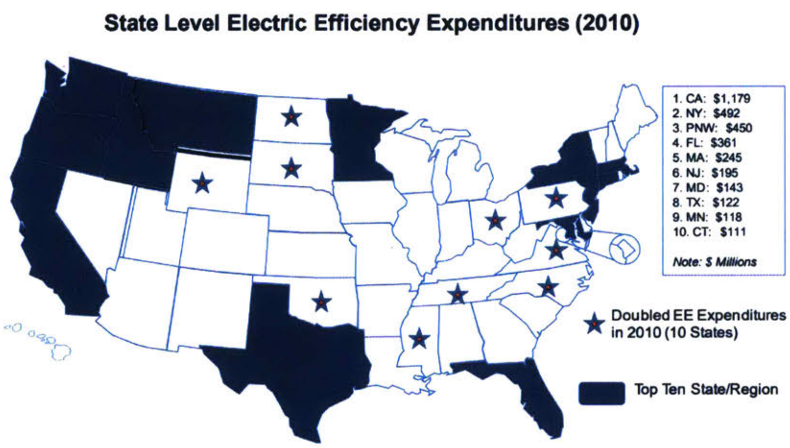


Figure 1.3 Electric Efficiency Expenditures by state. Source: Institute for Electric Efficiency

---

[2] Cooper, Adam, and Lisa Wood. 2012. "Summary of Ratepayer-Funded Electric Efficiency Impacts, Budgets, and Expenditures." *Institute for Electric Efficiency. The Edison Foundation.*

[3] Ibid.

In addition to energy efficiency incentives from utility companies, billions of federal and state dollars have been spent on other energy efficiency financing mechanisms.[4] Figure 1.4 shows major federal and state cost allocations for various energy efficiency initiatives across sectors since 1978.

Significant U.S. Energy Efficiency Policies

Name	*Year*	*Magnitude*
Corporate Average Fuel Economy Standards	1978–	$10 billion annual incemental cost from tightened 2012 rule (NHTSA 2010)
Federal Hybrid Vehicle Tax Credit	2006–2010	$426 million total annual credit (Sallee 2010)
Gas guzzler tax	1980–	$200 million annual revenues (Sallee 2010)
Federal appliance energy efficiency standards	1990–	$2.9 billion annual incremental cost (Gillingham, Newell, and Palmer 2006)
Residential and commercial building codes	1978–	
Electricity Demand-Side Management programs	1978–	$3.6 billion annual cost (US EIA 2010)
Weatherization Assistance Program (WAP)	1976–	$250 million annual cost (US DOE 2011a)
2009 Economic Stimulus	2009–2011	$17 billion total (U.S. DOE 2011b)
Additional WAP funding		$5 billion
Recovery Through Retrofit		$454 million
State Energy Program		$3.1 billion
Energy Efficiency and Conservation Block Grants		$3.2 billion
Home Energy Efficiency Tax Credits		$5.8 billion credit in 2009 (U.S. IRS 2011)
Residential and Commercial Building Initiative		$346 million
Energy Efficient Appliance Rebate Program		$300 million
Autos Cash for Clunkers		$5 billion

Figure 1.4 Significant Energy Efficiency Policies from 1978-2011. Source: Alcott and Greenstone (2012)

1.3 Energy Efficiency Gap: Unexploited Opportunities

Energy efficiency strategies are not adopted at the rate that they would be expected to, given the positive financial rate of return on energy efficiency investments. This phenomenon is often called the "energy efficiency gap" or "energy paradox," and there is a significant amount of literature that attempts to explain this paradox. One of the more widely utilized research papers in this domain is a paper by McKinsey & Company that provides the cost curve in Figure 1.5. This figure shows a mid-range greenhouse gas abatement curve with many different technologies. The graph shows a goal of abating over 3 Gt of carbon per year by 2030. As shown in Figure 1.5, many of the energy efficiency investments have a negative cost once the energy savings are taken into account over the lifetime of the measure. This negative cost is an intriguing finding, as one would not expect negative costs to exist in the absence of market failures. Many different papers have attempted to explain the energy efficiency gap, which

[4] Allcott, Hunt, and Michael Greenstone. 2012. "Is There an Energy Efficiency Gap?" *Journal of Economic Perspectives* 26.1 3–28.

likely exists for several different reasons. This chapter walks through several explanations of the gap that have been identified in the literature on this topic.

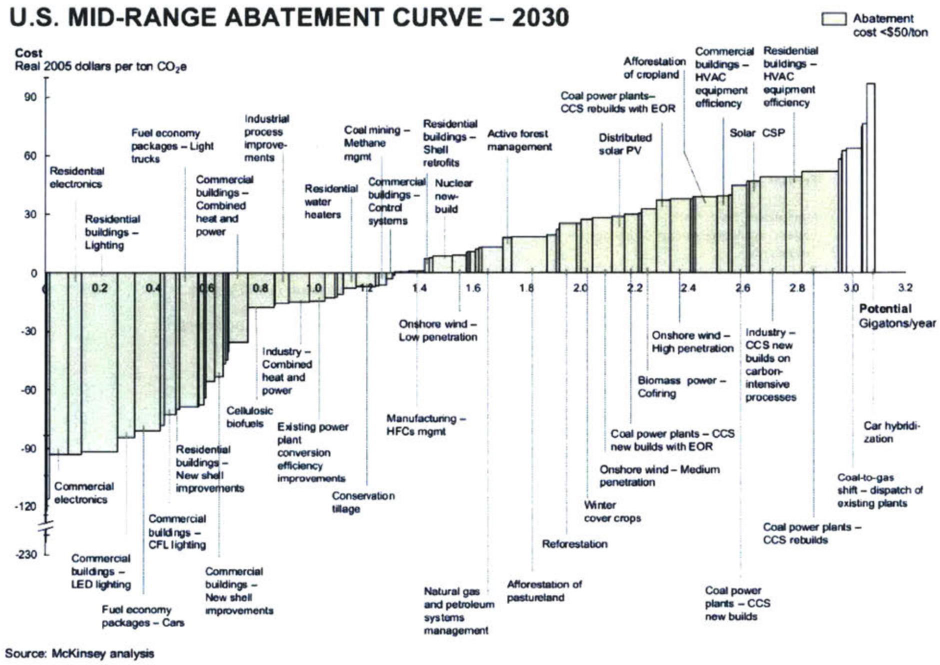


Figure 1.5 Mid-Range Abatement Curve for 2030 Source: Mckinsey & Company

1.3.1 The Energy Rebound Effect

When a consumer invests in energy efficiency, a certain percentage of the energy savings may be lost due to consumer behavior. For example, once the consumer saves on energy efficiency, they may use the cost savings to increase their energy demand. This phenomenon is termed the "rebound effect." An example of the rebound effect can be illustrated through home weatherization strategies. If a homeowner loses large amounts of heat to the outdoors due to the absence of insulation in their walls and drafty windows, they may choose to set the thermostat at 65°F and compromise occupant comfort in order to save on their energy bills. If the same homeowner insulates all the walls and air seals the windows, they may now decide to set the thermostat at a comfortable 70°F. The weatherization measures reduced the energy bill, and it now may be economically viable to set the thermostat at a more comfortable temperature. Thus, due to the rebound effect, the original projected carbon savings will not be fully achieved. Some studies suggest that 100% of the predicted energy savings will not be realized, thus deeming the strategy useless. Economists term this phenomenon "backfire" when energy efficiency strategies have no impact on energy reduction.

The existing literature on the rebound effect exhibits a wide array of definitions and types of rebound. This thesis adopts the definitions set forth by the International Panel on Climate Change (IPCC). The IPCC defines direct rebound as the rebound that occurs when there is a direct relationship between the energy usage that is reduced and the rebound of energy that is increased.[5] For example, if a consumer installs LED lights and subsequently becomes less diligent in turning off the lights when they leave the room since their energy bill is now fairly affordable, the IPCC would consider this a direct rebound effect. Alternatively, if the consumer installs LED lights, saves a significant amount of money on their energy bill, and subsequently uses the savings to buy an iPad or another energy intensive device, the IPCC would consider this an indirect rebound effect.[6] Direct and indirect rebound effects are often studied separately, but both are scrutinized for increasing overall energy usage.

Michael Shellenberger and Ted Nordhaus published a literature review, "Energy Emergence, Rebound and Backfire as Emergent Phenomena"[7] in 2011 through the Breakthrough Institute, which was the basis for their much debated New York Times op-ed article, "The Problem with Energy Efficiency."[8] The team reported in their literature review a 10 to 30% rebound effect for direct rebound in developed nations; 50% for direct and indirect rebounds at national and global scales; and 100% rebound or "backfire" in certain studies.[9] The literature review cited the 10 to 30% figure from a 2000 literature review which found a 10 to 30% rebound effect for residential space heating in developed countries.[10] However, this literature review focused only on studies which compared the change in energy use to

[5] Blanco G., R. Gerlagh, S. Suh, J. Barrett, H. C. de Coninck, C. F. Diaz Morejon, R. Mathur, N. Nakicenovic, A. Ofosu Ahenkora, J. Pan, H. Pathak, J. Rice, R. Richels, S. J. Smith, D. I. Stern, F. L. Toth, and P. Zhou. 2014. "Drivers, Trends and Mitigation. In: Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change." [Edenhofer, O., R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel and J.C. Minx (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

[6] Ibid.

[7] Jenkins, Jesse et al. 2011, "Energy Emergence." *Breakthrough Institute.* <http://thebreakthrough.org/blog/Energy_Emergence.pdf>

[8] Shellenberger, Michael. 2014. "The Problem with Energy Efficiency". *New York Times.* <http://www.nytimes.com/2014/10/09/opinion/the-problem-with-energy-efficiency.html?_r=1>

[9] Jenkins, Jesse et al. 2011, "Energy Emergence." Breakthrough Institute. <http://thebreakthrough.org/blog/Energy_Emergence.pdf>

[10] Greening, Lorna et al. 2000. "Energy efficiency and consumption — the rebound effect — a survey" *Energy Policy.* <http://ac.els-cdn.com/S0301421500000215/1-s2.0-S0301421500000215-main.pdf?_tid=c30d6dac-943f-11e5-9b6f-00000aab0f02&acdnat=1448543908_ad7cfe1e5fcaafd3d53fabcd39d87007>

engineering calculations. This approach relied on the engineering calculations to be 100% accurate and ignored any other possible explanations for the difference in energy use.

Nonetheless, energy rebound and backfire do help to explain the energy efficiency gap in certain circumstances. Davis and Fuchs (2014) evaluated an energy efficiency appliance program in Mexico and found significant evidence of the rebound effect. [11] The program provided refrigerators and air conditioners to 1.9 million households over a four-year period. After receiving high efficiency air conditioners, participants actually increased their electricity usage, likely from using more air-conditioning after the price to operate air-conditioning was reduced. Not only did electricity usage increase after receiving the new air conditioners, but the increase only occurred during summer months, further substantiating the claim that the energy rebound effect was occurring.[12]

Although energy rebound reduces the energy savings potential of a project, it should be noted that energy rebound still allows for an overall improvement in welfare. For example, in the Davis and Fuchs (2014) paper, occupants were able to increase their air-conditioning usage and therefore improve comfort. Perhaps some occupants were able to spend more time in their home, improve productivity while at home, or even improve their health. Although some energy efficiency programs are geared only towards energy savings and would not benefit from welfare improvements, some programs are geared towards both welfare improvements and energy savings.

1.3.2 Bias in Ex Ante Assumptions

In addition to the energy rebound effect, some energy efficiency programs do not meet expected energy savings, simply because the ex-ante predictions are overconfident. Most studies on the energy efficiency gap compare actual savings with predicted savings, even though predicted savings are based simply on engineering analyses. These analyses can include various biases if the engineers have an incentive to over predict; if the engineer simply does not have any incentive to accurately predict the savings; or other biases that might exist. In fact, Alcott and Greenstone in a 2012 paper found that "much of the evidence of energy cost savings from energy efficiency comes from engineering analysis or

[11] Davis, Lucas W., et al. 2012. "Cash for Coolers" *Cambridge, Mass.: National Bureau of Economic Research, NBER working paper series: no. 18044.*
[12] Ibid.

observational studies that can suffer from a set of well-known biases.[13] For example, in the Davis and Fuchs (2014) paper, the authors found evidence of over prediction in the case of the refrigerator installations. Although the installation of new refrigerators did show an 8% decrease in electricity consumption, this number was still approximately a quarter of the original ex ante predictions.[14] The authors found this discrepancy due in part to the fact that "ex ante predictions were overly optimistic." For example, the predictions did not take into account the increase in size of refrigerators in the market at the time, as compared with the older models that were being replaced. The paper found that according to the July 2009 Mexican Consumer Protection Office Report, the average size of refrigerators was 13.5 ft.[3] However, the ex ante predictions predicted that refrigerators would stay between 9 and 13 ft.[3]

The Shellenberger literature review of the energy rebound effect omitted other methodologies that would capture overly optimistic ex ante predictions, with the argument that:

> "Other methods, such as direct measures of changes in thermostat set point, have been used in the analysis of the rebound effect (Nadel, 1993). However, many of those studies suffer from sample bias resulting from the selection techniques used for participants, and many other factors were not controlled for (Greening and Greene, 1998)."[15]

Although there may be a sample bias in these previous studies, the 2015 research paper "Do Energy Efficiency Investments Deliver? Evidence from the Weatherization Assistance Program" successfully utilized this methodological approach of zeroing in on the thermostat set point.[16] In the case of residential space heating rebound effects, zeroing in on the thermostat is an effective method to isolate the rebound effect. The rebound effect says that occupants will set a higher temperature set point after weatherizing homes, so the rebound effect can be evaluated by focusing on the before and after thermostat set points over a period of time. This approach isolates the rebound effect from other

[13] Allcott, Hunt, and Michael Greenstone. 2012. "Is There an Energy Efficiency Gap?" Journal of Economic Perspectives 26.1 3–28.
[14] Davis, Lucas W., et al. 2012. "Cash for Coolers" Cambridge, Mass.: National Bureau of Economic Research, NBER working paper series: no. 18044.
[15] Ibid.
[16] Fowlie, Meredith. 2015. *"Do Energy Efficiency Investments Deliver? Evidence from the Weatherization Assistance Program." University of Chicago.*
<http://econresearch.uchicago.edu/sites/econresearch.uchicago.edu/files/paper_draft_06_15_clean.pdf>

changes to energy use. This study evaluated 30,000 homes checking both energy usage and space temperatures, and the study did not find any evidence for the rebound effect. The study did find that the original predicted savings through engineering calculations were 2 ½ times that of the actual savings, but the study attributed the difference to poor engineering calculations. There was no evidence for a significant increase in space temperature in the weatherized homes for this specific application.[17] However, the savings from this research study were still below the ex ante predictions, and the average rate of return was -9.5% annually.[18] This study suggests that in the case of the weatherization assistance program, the engineering estimates were overly optimistic.

As noted previously, most papers on the energy efficiency gap focus on comparing engineering estimates to actual energy savings. For example, the literature review "Energy Emergence, Rebound and Backfire as Emergent Phenomena" focused only on comparing ex ante calculations to actual energy savings. This methodology fails to take into account overly optimistic predictions and calls for additional research methodologies. Most FDD software use simple engineering calculations to predict energy loss, and most evaluation programs of FDD software rely on engineering calculations. However, this thesis provides an alternative approach to estimating energy waste associated with FDD software utilizing actual energy data.

1.3.3 Other Reasons for the Gap

There have been many other reasons identified in the literature that help to explain the energy efficiency gap. Split incentives occur when the individual or entity that pays the energy bills are separate from those that are utilizing the energy in buildings. For example, in a landlord-tenant scenario, the landlord might pay the energy bill, but the tenant is utilizing energy. In this scenario, the tenant does not have an incentive to reduce energy usage. Further, some literature cites imperfect information as another reason for the existence of the gap. Consumers do not always have the right information about the long-term energy impact of various technologies and may not make rational economic decisions. Fault detection diagnostics software addresses the issue of imperfect information and energy efficiency by identifying issues in commercial HVAC equipment that would otherwise not be identified by staff. For

---

[17] Ibid.
[18] Ibid.

example, a common issue found by fault detection software is simultaneous heating and cooling. If a heating control valve fails open and overheats the air supply, the cooling valve controller may respond by providing additional cooling, such that the supply air is discharged at a comfortable 68°F. This problem can easily go unnoticed without fault detection software.

1.4 The Need for Improved Energy Modeling Techniques

Because of the large gap in energy efficiency, there is a need for more robust methodologies to better estimate energy savings and optimize energy efficiency programs. If energy efficiency is indeed going to be such a large component of climate action plans, and continue to be an integral component of federal and state regulations, there is a strong need for improved methods to verify savings and quantify methods to improve these programs. As noted earlier, most papers compare engineering estimates to actual savings, but these engineering estimates are often flawed and over or under predict actual savings. While there is some literature that provides improved methods for the evaluation of energy savings programs, the research is very limited in scope and application. This section outlines the existing literature on the topic and the many gaps in literature. As seen in the following subsections, most papers on the topic identify a strong need for more research in this domain.

1.4.1 The Need for Improved Linear Regression Modeling Techniques

After an energy efficiency measure is implemented in a building, and at least 12 months of data is generated, there is often some sort of "measurement and verification" (M&V) phase in which performance data is evaluated to determine how effective the installation was. M&V often occurs when there is a performance-based contract in place in which the energy efficiency installer is paid all or in part based on the savings that occur. The International Performance Measurement and Verification Protocol (IPMVP) is often widely used, as well as and American Society of Heating, Refrigerating, and Air Conditioning Engineers (ASHRAE) Guideline 14, Measurement of Energy and Demand Savings. There are several options outlined for estimating the impact of an energy conservation measure, one of which includes using energy data. However, interval energy data is often not used, and the full benefits of a multiple linear regression model are not often used. For example, in the US Department of Energy's "M&V Guidelines: Measurement and Verification for Performance – Based Contracts Version 4.0," the report recommends not using 15-minute interval data and instead aggregating interval data into daily or monthly data sets.[19] Although the variability of 15-minute interval data can be difficult to analyze in simplified models, there is a wealth of information in this data that can be uncovered through more complex linear models. With 15-minute interval data, and more complex algorithms, it is possible to better predict what the energy use would have been absent the energy savings measure – often called "baseline energy use" or "counterfactual energy use."

A paper by Lawrence Berkeley National Lab successfully built a multiple linear regression model in order to evaluate a demand response program.[20] In a demand response program, customers are paid to reduce energy consumption during peak energy events on the electric grid. For this reason, it is very important to effectively determine the counterfactual energy consumption on a granular hourly and daily level. Despite this need for granularity, the paper found that most utilities use only simplified approaches to estimate baseline energy consumption, such as averaging electric energy over recent days with high energy usage. The paper instead offers a methodology for building a regression model to better estimate baseline energy consumption during demand response events. This paper provides useful methods for understanding and utilizing energy data for C&I buildings. The paper offers new ways of visualizing electrical load data (mostly for demand response) and methods for building a regression model with 15-minute interval load data from dozens of commercial and interior facilities. The regression model uses a "time of week indicator variable" and a "piecewise linear and continuous outdoor air temperature." The time of week indicator variable is a categorical value for every 15-minute period in a week. The paper was able to provide a more robust option for evaluating demand response participation. The paper found that a multiple linear regression model offers "negligible computational burden," as compared with a more sophisticated methodology, such as machine learning techniques.[21] However, in many applications in which the extra computational power is both possible and cost-effective, the higher accuracy and granularity of a machine learning model can be a strong asset.

---

[19] U.S. Department of Energy Federal Energy Management Program. 2015. "M&V Guidelines: Measurement and Verification for Performance-Based Contracts Version 4.0." <https://energy.gov/sites/prod/files/2016/01/f28/mv_guide_4_0.pdf>

[20] Mathieu, Johanna L. et al. 2010. "Quantifying Changes in Building Electricity Use, with Application to Demand Response." *IEEE Transactions on Smart Grid.*

[21] Ibid.

1.4.2 Opportunities for Further Improvements in Statistical Modeling

While there is limited usage of multiple linear regression models for evaluating energy usage, especially in industry, more robust machine learning models are not used in industry at all and are used very little if at all in academic or industry research. In fact, David Hsu identified in his paper on statistical models for energy usage that unlike other topics or applications, "the energy consumption literature seems not to have used regularization methods very much at all. Searching for the exact phrases "penalize regression" or "regularization" in Google scholar yields only one paper in the core energy journal… Searching for the names of specific methods such as "lasso" and "Ridge" yields only a few more papers, mostly in economic models of energy use."[22] This paper applied hierarchical group – Lasso regularization to evaluate energy consumption of New York City multifamily homes. However, the paper focuses on creating building benchmarking in order to effectively compare buildings to each other, rather than evaluating one buildings energy consumption over time. Hai-xiang Zhao identified a similar gap in the literature on energy use in buildings in his paper on energy consumption production through statistical models. The paper offers methods for determining which factors are most necessary and useful for building a detailed energy model.[23]

1.4.3 Applications to Commercial Buildings

In addition to the need for improved modeling techniques for energy efficiency evaluation, there is a strong need in the literature for more evaluations of commercial building energy efficiency programs. For example, the Fowlie (2015) paper mentioned previously on the Weatherization Assistance Program found that the energy efficiency financing for this particular program in Minnesota did not pay off after the energy savings were evaluated. However, the paper was widely referenced in many different media sources and widely debated. Some argued that because this paper focused only on the residential market, and only on low-income customers (as the federal Weatherization Assistance Program is targeted for) it does not paint a full picture of energy efficiency programs. Some further argue that low-income residential customers are the hardest to target, and therefore significant

[22] Hsu, David. 2014. "Identifying key variables and interactions in statistical models of building energy consumption using regularization." *Energy.* 83 (1) 144–155

[23] Zhao, Hai-xiang, and Magoules, Frédéric. "Feature Selection for Predicting Building Energy Consumption Based on Statistical Learning Method." *Journal of Algorithms & Computational Technology.* 6 (1) 59-77

financing is used to advertise to this community. For example, in a Scientific American article, the author commented that the findings from the paper:

> "didn't surprise Steven Nadel, the American Council for an Energy-Efficient Economy's executive director. 'It is tough to reach these people,' he said. 'They're working multiple jobs, they're not sure they want to trust the government ... there are a lot of reasons people don't participate.' Nadel cautioned against drawing broad conclusions about energy efficiency's cost and potential from the University of Chicago study. "They looked at one single program, and they're trying to generalize,' he said, arguing that weatherization efforts like WAP are among the 'least likely [efficiency programs] to be cost-effective.'[24]

A major takeaway from the debate following this paper is that there is a need for more similar studies done on the commercial sector to see if the same results can be replicated, or if the commercial sector results in higher returns. This thesis targets the commercial building sector in part because of these findings.

One successful application to commercial buildings can be found in a 2017 paper, "Evaluating Energy Efficiency Upgrades to K-12 Public Schools in California Investor-Owned Utility Territories."[25] This paper successfully utilizes interval energy data and machine learning techniques, while addressing the absence of research identified above. This paper estimated counterfactual energy consumption using machine learning techniques in 1105 K-12 schools in California that underwent 4354 upgrades between January 2008 and December 2014. Rather than only compare the energy savings to the ex-ante predictions, this paper also compared the savings to a control group of schools that did not receive energy efficiency treatments. The paper found a 2 to 4% reduction in energy use during daytime hours at highest temperatures for the HVAC energy efficiency upgrades, accounting for 60 to 85% of the expected energy savings. On the lighting side, the paper found 5 to 7% reductions during daytime hours. The authors argue "there is a clear need for techniques to estimate returns to energy efficiency programs that can be applied in a wide set of contexts."[26] This thesis builds off some of the methodology

---

[24] Detrow, Scott. "Energy-Efficiency Efforts May Not Pay Off." Scientific American. <https://www.scientificamerican.com/article/energy-efficiency-efforts-may-not-pay-off/>

[25] Knittel, Chris and Wolfram, Catherine. 2017. "Evaluating Energy Efficiency Upgrades to K-12 Public Schools in California Investor-Owned Utility Territories." *MIT working paper.*

[26] Ibid.

utilized in the K-12 paper, but applies it to the evaluation of fault detection diagnostics in commercial university buildings in New England.

1.4.4 Improvements in Data Access

With the increased availability of interval energy data, there are more opportunities to derive more information and build more accurate models of energy consumption. As cities, academic institutions, companies, and other entities develop climate action goals and energy efficiency goals, an integral component of these plans is often to add more submetering and interval metering. While there is some literature that utilizes 15-minute interval electric data to improve modeling, none of the papers described above use 15-minute interval chilled water or steam data. As noted in more detail in Chapter 4, some buildings utilize chilled water for cooling loads and steam for heating loads. By metering and evaluating each of these utilities separately, there is significant increase in information that can be utilized for a model. For example, by building separate models for each utility, a model can analyze a heating system. This would not be possible for a building that uses gas or district heat but is evaluated only on electricity consumption.

1.4.5 The Need for Research Evaluations of Fault Detection & Diagnostic

Finally, there is very little research done to evaluate the energy that could be saved by making repairs to faults that are identified by a fault detection and diagnostics system (FDD). Most studies evaluating FDD and estimating energy savings do not use actual energy metered data, but rather make estimations based on the equipment affected. Chapter 2 builds on this section; provides further details and examples of literature that has been done in this domain; and provides information to show the need for this type of research.

1.5 Conclusion

This chapter provides information on the energy efficiency gap and the need for more research done to evaluate what causes the gap in certain domains and why there is a strong need for improved methodologies to build baseline statistical models with higher accuracy. This chapter also provides a literature review of existing research that has been done using various statistical models applied to energy consumption, and to outline several gaps in this research. This thesis addresses each of these gaps, by evaluating both multiple linear regression and machine learning models; focusing on commercial buildings; taking advantage of advancements in available building energy usage data; and applying these models to faults identified by an FDD system. Chapter 2 elaborates on section 1.4 of this chapter by providing more information on the energy efficiency gap at as it applies to building automation systems, as well as more details on the need for applying these models to energy data to evaluate FDD systems.

2 Building Automation Systems & the Energy Efficiency Gap

2.1 Introduction

This chapter provides background on building automation systems (BAS) and fault detection & diagnostics (FDD) systems. Building automation systems are found in many commercial buildings in the United States and around the world. These systems are able to control all HVAC equipment in a building utilizing sophisticated control strategies that can save significant amounts of energy. FDD systems represent a new field of techniques to continuously monitor building HVAC systems, often through the BAS, in order to find any issues that arise that waste energy. This chapter outlines a brief history and background on building automation systems and FDD systems; the benefits that building automation systems can provide in terms of energy savings; and existing literature around the energy efficiency gap as it pertains to building automation systems.

2.2 Background

This section provides background on both the building automation system industry and the relatively new fault detection industry, as well as brief overviews on how both systems work. There are significant savings that can be found through building automation systems, and fault detection and diagnostics is one method to ensure proper functionality and savings from these systems.

2.2.1 A Brief History of Building Automation Systems

In the early 1900s, the HVAC controls industry emerged as building owners began using pneumatic controls to control heating and cooling in buildings. Building owners could now condition spaces and control to a specified temperature set point. The pneumatic systems consisted of tubes and compressed air running through a building from air compressors to a network of HVAC devices. In response to the energy crisis of 1973, the HVAC controls industry shifted from a pneumatic-based system to an improved electronic-based system. Johnson Controls introduced the first direct digital control computer in 1980, and others followed suit.[27] With the help of electronic systems, building

[27] Newman, Michael. 2013. "BACnet: The Global Standard for Building Automation and Control Networks." *New York, Momentum Press.*

owners were able to control HVAC equipment with more precision and more sophisticated control techniques, subsequently saving substantially more energy.

Unlike with the pneumatic system, now building owners had the ability to control an entire network of HVAC equipment with a computer. However, unlike the pneumatic system, the new platform did not exhibit interoperability between manufacturers' products. Each manufacturer rushed to come out with an electronic control technology, and each system was proprietary. The pneumatic control systems were "standardized, and several different companies manufactured interchangeable pieces of equipment to control buildings' internal environments."[28] However, the electronic controls industry developed quite differently. Once a building owner purchased a Building Automation System (BAS) with electronic controls, the owner was locked in to that manufacturer's product. The manufacturer had market power in the building, and the owner was forced to purchase new components and service directly from the original manufacturer. Each time the owner made small upgrades to the HVAC system, or had any servicing needs, the owner did not have a competitive market to turn to.

In January of 1987, The Association of Heating Refrigeration and Air-Conditioning ("ASHRAE") began addressing this issue and formed the Standard Project Committee 135P to create the voluntary BACnet standard protocol for communicating between systems. ASHRAE published the first public draft for review of BACnet in August 1991, and the American National Standards Institute (ANSI) approved the BACnet standard in December 1995. Eventually, BACnet technology saturated the market and both producers and consumers welcomed the technology. By 2000, a BACnet Manufacturers Association study found that BACnet systems had been installed in 82 countries around the world.[29]

Although there are still issues being reported regarding interoperability as the BAS industry rapidly expands and innovates, the development of the BACnet protocol was a huge landmark for the industry. Now, with almost 100 systems on the market, the industry faces new opportunities and challenges. As systems become more complex and incorporate more functionality, the systems require more resources and more data sources to take advantage of. Often these systems are underutilized and the data is not fully taken advantage of. As mentioned earlier, new field has begun to emerge, often termed fault detection and diagnostics ("FDD"), that pulls data from the BAS, analyzes the data, and offers insightful information into opportunities for energy savings and maintenance saving opportunities.

---

[28] Lee, Jim. 2000. "Controlling Your Environment." *Bacnet*.
[29] Swan, William 2004. "BACnet Retrospective" *Networked Controls*.

2.2.2 Energy Saving Control Strategies in a Building Automation System

A BAS has the potential to save significant amounts of energy in commercial buildings through numerous different control strategies. For example, through demand controlled ventilation, a BAS can collect information from carbon dioxide sensors to determine how much fresh air is actually needed for the space (e.g. roughly how many people are actually in the space), and adjust the ventilation rates accordingly. Demand controlled ventilation can save significant amounts of energy, especially during periods of high or low outdoor temperatures that require substantial conditioning of outdoor air. Through simple scheduling techniques, a BAS can turn off systems or set back temperatures based on the scheduled usage of each space. Further, through economizer control, a BAS can control HVAC such that outdoor air is used to cool spaces when the outdoor air is colder and less humid than indoor spaces. This can occur in certain climates during the fall and the spring when many people, computers, and other technologies add heat to the building. With each of these control strategies and many more, a BAS is able to save significant amounts of energy if the full potential of the system is met. However, this full potential is not often met, and fault detection and diagnostics can help address this energy efficiency gap.

The eSource schematic Figure 2.1 shows a schematic of a typical building automation system.[30] As seen in the figure, sensors throughout the HVAC system, such as temperature sensors, humidity sensors, valve positions and other data points provide data back to the controller. The controller then uses this information to send outputs back to the system. For example, a temperature sensor might indicate that a room in the building is below the temperature set point for that space. If this is the case, the controller will then signal to the system to increase temperature, either by opening a heating valve, increasing airflow, or another method to increase temperature. A BAS often controls valve positions, damper positions, fan speeds, and many other components of an HVAC system. There is often a browser-based interface in which building operators can visualize data from the system; change set points and schedules; or even change the control schemes.

---

[30]Sustar, John and Goldschmidt, Ira. 2007. "Saving Energy with a Building Automation System." *Energy Managers' Quarterly.* <http://bea.touchstoneenergy.com/sites/beabea/files/PDF/Tech/SavingEnergywithaBuildingAutomationSystem.pdf>

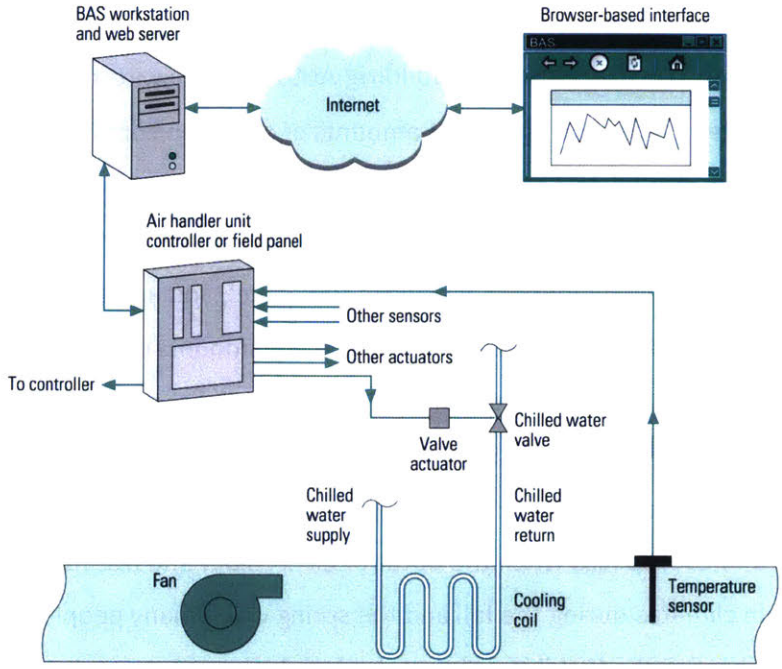


Figure 2.1 Schematic of typical Building Automation System network. Source: Esource.

As one could imagine, with all of these data points and components that can be controlled in each room and mechanical space of a commercial building, there is often considerably more data coming out of the BAS then can possibly be monitored by building operator. Thus, with this wealth of data coming from the BAS, and with the rise of available machine learning techniques and other big data mining techniques in other industries, we are now seeing an abundance of FDD systems on the market. These systems can utilize this large source of data from a BAS and can continually analyze the systems for any potential faults that arise. The next section provides a brief background on FDD systems.

2.2.3 Background on Fault Detection & Diagnostics

As mentioned previously, FDD systems are able to take advantage of big data from building automation systems to analyze the system and provide useful insight to building operators. These systems monitor data in order to find potential faults that have arisen on the system. For example, Figure 2.2 provides a basic schematic of a typical HVAC system. [31] The cooling fan provides conditioned

[31] Sathyananda M R. 2017. "Air side, VAV systems, HVAC." *LinkedIn*. <https://www.linkedin.com/pulse/air-side-vav-systems-hvac-sathyananda-m-r>

air to the variable air volume (VAV) terminal boxes throughout the building. These variable air volume terminal boxes typically sit in the ceiling space of a room, and the term "variable" defines air terminal boxes that can modulate flow with the use of a damper that opens and closes anywhere between a minimum position and 100% open. These terminals often have a reheat valve inside the unit in order to add an additional level of heating in the space if the thermostat controller in the space is calling for more heat than what is supplied by the cooling fan. Some of this air is then recirculated back to cooling fan, and some air is exhausted out of the building and instead replaced with fresh, outdoor air. A number of faults can arise in this simple type of configuration, such as dampers in the terminal getting stuck in one position; a heating or cooling valve getting stuck in one position; fans being overridden to run at 100% speed; or many other potential faults. With some commercial building having several fans, 50 or even 300 terminal boxes, and many other types of HVAC systems that are not included in the schematic, one could imagine the potential for thousands of faults to occur at any given time.

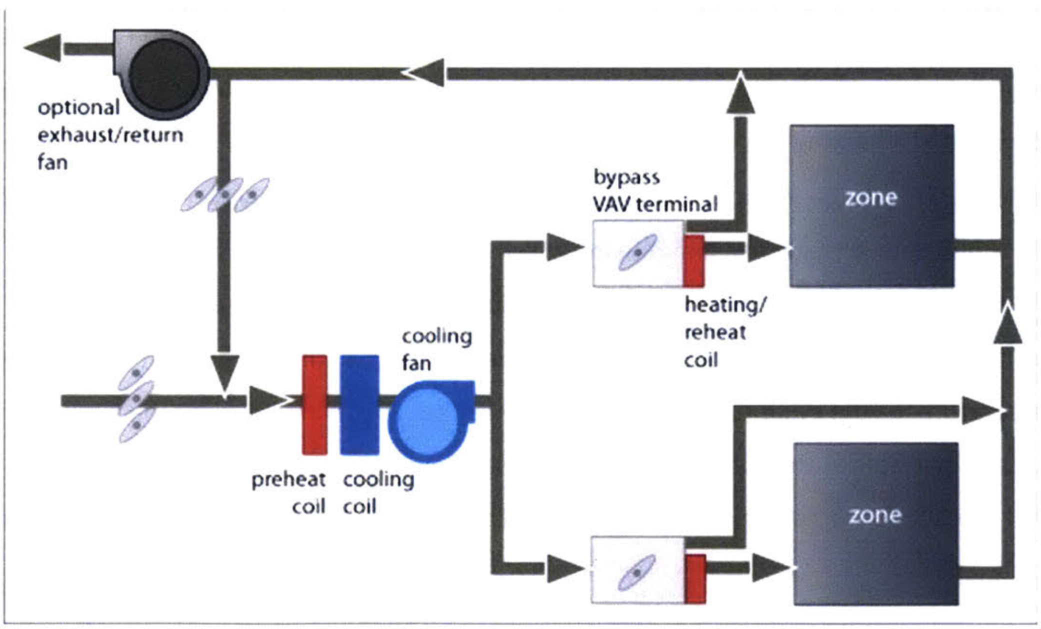


Figure 2.2 Basic schematic of a typical HVAC system. Source: Sathyananda (2017)

One such example of a fault mentioned above was identified by the Clockworks KGS fault detection and diagnostics system at MIT. Figure 2.3 is a schematic produced by KGS that shows the details of this particular fault with a visual of two data points that were pulled from the BAS and tracked over the course of 24 hours. In a typical configuration with a heat recovery coil in the VAV terminal box of a room, the thermostat controller in the room could be manually controlled by occupants or controlled at the BAS workstation. Whatever temperature set point the room is set to, the valve for the heating coil will open or close to some percentage point in order for air to flow over the coil and heat the room to the desired set point. In this particular fault in the figure below, the thermostat controller is commanding the heat recovery coil valve to be set at 0% for this 24-hour period. However, as seen in the figure, temperature is still rising significantly over this time period. In the absence of a fault, we would likely see the temperature stay constant once the valve is commanded to a 0% valve position. This particular fault could be caused by a number of issues. Even though the valve is commanded closed, it may be stuck open, which is a very common finding with valves, and thus the space could be overheating. Alternatively, there could be another HVAC heat source in a room that is controlled separately and causing the space to overheat. Ideally all systems would have the same controller so that a specified temperature set point can be maintained in the space. Regardless of what the actual fault is, FDD software is able to identify where problems may be occurring, and subsequently point to the building operator or technician to the site to evaluate what the problem is. Although FDD software cannot always diagnose the exact problem, it is an extremely useful tool for alerting building operators to potential issues. If a heating valve is stuck open and providing extra heat to a space, this issue could easily go unnoticed for years. Without fault detection & diagnostics, the only way to identify this issue would be to 1) pull data and analyze it manually (which mentioned above is a very tedious and expensive task with hundreds or thousands of data points to analyze), 2) to receive an occupant complaint of the space overheating (although sometimes occupants will simply open a window to control temperature, or the system itself will provide extra cooling to make up for the extra heating if a cooling source is available in the space), or 3) a technician could remove ceiling tiles and examine the VAV terminal box above the ceiling, which may occur if there is preventative maintenance to inspect every terminal box on a set schedule, or if an energy efficiency or retro-commissioning study is being conducted.

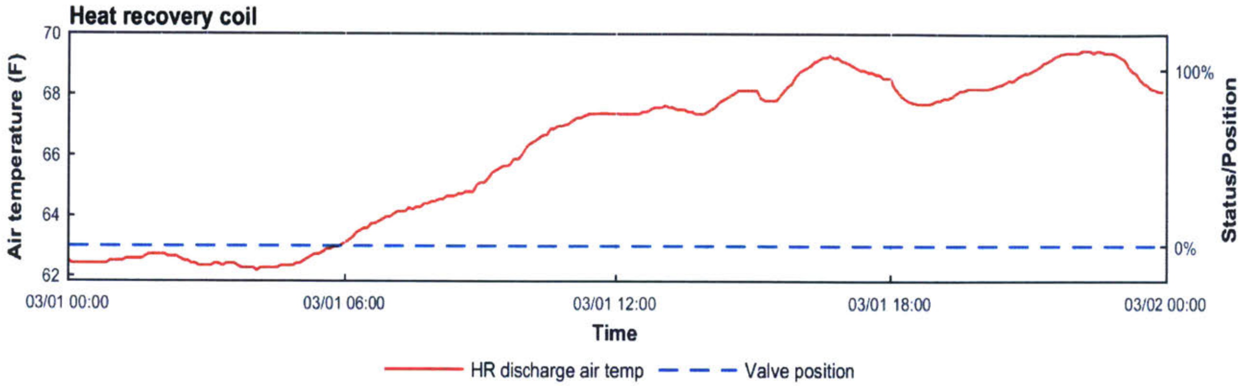


Figure 2.3 Example of a fault on the MIT FDD system. Source: KGS Clockworks

2.3 Market Snapshot

While the BAS industry has been around since the 1970s, it continues to grow and expand its market into new buildings in the US and many new buildings across the world. The birth of the FDD industry in recent years follows the expansion of the BAS industry, in addition to the expansion of big data analytic techniques and the expansion of general awareness of energy efficiency and climate mitigation techniques. The next two sections outline the market for the BAS industry as well as the FDD industry.

2.3.1 Building Automation System Market Snapshot

The BAS industry has expanded rapidly in the last decade, and the industry is expected to exceed $80 billion in revenue globally by 2020. Figure 2.4 shows the past and predicted total revenue each year from 2013 to 2023 across five regions.[32] As seen in the figure, the Asia-Pacific market is expected to grow the most, almost doubling its revenue over a decade. From an energy efficiency and climate mitigation perspective, it is crucial that these systems are optimized before they expand to other

[32] Navigant Research. 2014. "Commercial Building Automation System."

markets. Any energy shortcomings in existing systems will continue as the market expands if adjustments are not made.

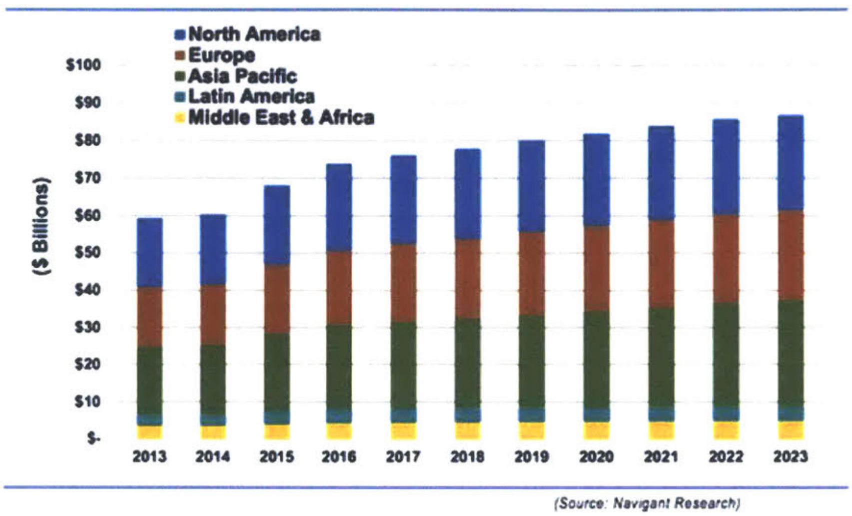


Figure 2.4 Commercial Building Automation Revenue by region. Source: Navigant Research

In addition to regional variations, BAS prevalence varies by building type and building size. As seen in Figure 2.5, educational buildings have the highest percentage of energy management and control systems, almost at 60% of floor space.[33] This may be due in part to the fact that educational institutions, including private colleges and universities; public colleges and universities; and private K-12 schools, often own their own buildings. Because energy management systems are long-term investments with a high capital cost and a long energy payback period, a control system is a better investment when the payback is realized by the owner. On the other end of the spectrum, the retail and service industry has only about 20% of floor space controlled with an EMCS.[34] This could be in part due to high turnover rates in ownership of retail and service industries as well as the relatively small floor space. As seen in Figure 2.6, there is a direct relationship between building size and percentage of

[33] Brambley, MR et al. 2005. "Pacific Northwest Laboratory: Advanced Sensors and Controls for Building Applications: Market Assessment and Potential R&D Pathways." *Pacific Northwest Laboratory.* <http://www.pnl.gov/main/publications/external/technical_reports/pnnl-15149.pdf>
[34] Ibid.

buildings with an EMCS. The larger the building, the more likely it is to have an EMCS. As buildings become larger, HVAC systems become more complex with more components and larger units. The need for an EMCS increases as the opportunity for energy efficient control strategies increases and the need for control and risk mitigation increases on larger assets.

Figure 2.5 Prevalence of energy management and control systems (EMCS) by building type (CBECS 1999)

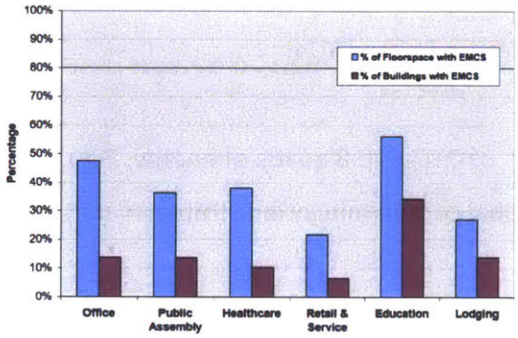
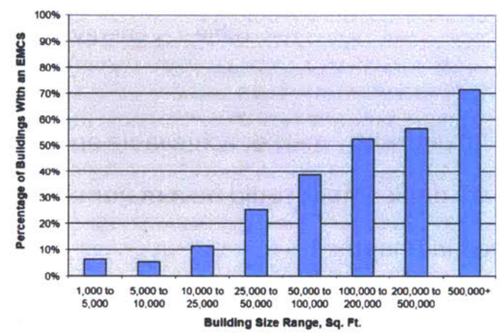


Figure 2.6 Prevalence of energy management and control systems (EMCS) by building size (CBECS 1999)



2.3.2 Fault Detection & Diagnostics Market Snapshot

While the FDD industry is only a few years old, it has expanded rapidly and is expected to continue to grow as the BAS industry and available data mining techniques grow. It is also possible that the industry will make drastic improvements in the software available, such as reducing the number of false positives and potentially automating the work order process when a fault is identified that needs to be repaired. There is a lot of opportunity to develop the industry, and a literature review of fault detection systems estimated the market potential for the industry. Table 2.1 identifies the market potential for fault detection identified in the literature review. The review estimates potential market penetration of 15 to 55% of buildings, resulting in 8 to 30 billion square feet of floor space.[35] Given this expected market growth for fault detection, it is imperative that new methodologies are developed and applied in order to evaluate the systems and continually improve performance.

[35] Bruton, Ken et al. 2014. "Review of Automated Fault Detection and Diagnostic Tools in Air Handling Units." Energy Efficiency 7.2: 335–351.

Attribute	Value	Notes
Technical Market Size	55 billion ft^2	Almost all commercial floor space
Relevant Annual Energy Consumption	9.8 quads	Heating, cooling, lighting, ventilation, and 50% of refrigeration energy
Energy Savings	5% - 15%	Similar to initial savings from commissioning
Technical Annual Energy Savings Potential	0.5 – 1.5 quads	
Commercial Building Peak Reduction	5-15%	Similar to initial savings from commissioning
Simple Payback Period	1 - 3 years	Very approximate; may be longer due to cost of implementing hardware fixes
Ultimate Market Penetration	15-55% / 6-24%	Lower values for continuous commissioning reflects limited market of buildings with EMCSs
Market Potential	8–30 / 3–13 billion ft^2	
Market-Achievable Annual Energy Savings	0.07–0.8 / 0.03–0.35 quads	Stand-alone and EMCS-based systems, respectively

Table 2.1 Energy savings of FDD. Source: Ken Bruton (2014)

Additionally, with "market achievable annual energy savings" of 0.07 to 0.8 quads of energy, FDD software deployment could be a major component of the global carbon mitigation strategies that were outlined in Chapter 1.

2.4 Energy Efficiency Gap for Building Automation Systems

As noted above, although a BAS has the potential to save significant amounts of energy, there is a lot of literature that shows that there exists an energy efficiency gap for building automation systems. This section provides an overview of the literature surrounding the energy efficiency gap as it applies to building automation systems. This section also provides a literature review of existing fault detection and diagnostics systems that have been evaluated in real buildings.

2.4.1 Evidence of the Energy Efficiency Gap for Building Automation Systems

The gap in energy savings between the potential that building automation systems can provide and what they actually provide in real buildings has been heavily studied. Building automation systems in existing buildings often do not meet their full potential and continue to operate with significant issues that go unnoticed. In fact, one paper on the topic found that 50% of buildings continue to operate with significant issues in their control systems.[36] If half of all buildings operate with significant issues, there is a large need for further analysis of the energy efficiency gap and the potential for fault detection to

[36] Ibid.

monitor systems. In a literature review on the topic the authors found that "ill-functioning building systems and equipment waste a significant quantity of energy, in some cases equal to up to 30% of the energy consumed by commercial buildings."[37] As noted in Section 2.2.3, there are hundreds or even thousands of components in building HVAC systems, so the potential failure rate of various components is high. When some components fail, they go into "fail-safe mode" to prevent further damage to the system, but this mode often increases energy usage. For example, a common fault that occurs as identified above includes heating coil valves failing. Over time the valve actuator that opens and closes the valve based on how much heating is needed begins to lose functionality. When the valve actuator fails, the valve is put in "fail-safe mode" and therefore stays stuck at 100% open. By staying at 100% open, the valve provides maximum heating to the coil and prevents the possibility of under heating the space. If a space is under heated, hot water coils could freeze and subsequently burst the pipe, causing significant damage. Failsafe mode prevents this damage from occurring, but it can take a serious toll on energy usage in a large building in the absence of any continuous monitoring. This example and many other examples of faults and other issues in HVAC systems are what cause some of the gap that is seen in the literature.

In another paper that reviewed the energy efficiency gap in building automation systems, the author found that for most control strategies, building automation systems were not being used to their full potential. The researchers provided a questionnaire survey to BAS users and conducted statistical analyses on the results. Most of the BAS users do not use the software as a full time job, but instead utilize the software an average of 6.5 hours per week and no more than 30 hours per week.[38] The BAS users were asked in the questionnaire about which control strategies were being utilized in their systems. As seen in Figure 2.7, the study found that building automation systems are significantly underutilized. Although this study came out before the development of fault detection systems, it was still able to identify significant shortcomings in HVAC control systems.

[37] Ibid.
[38] Lowry, G. 2002. "Technical Note: Factors Affecting the Success of Building Management System Installations." *Building Services Engineering Research & Technology* 23 (1) 57.

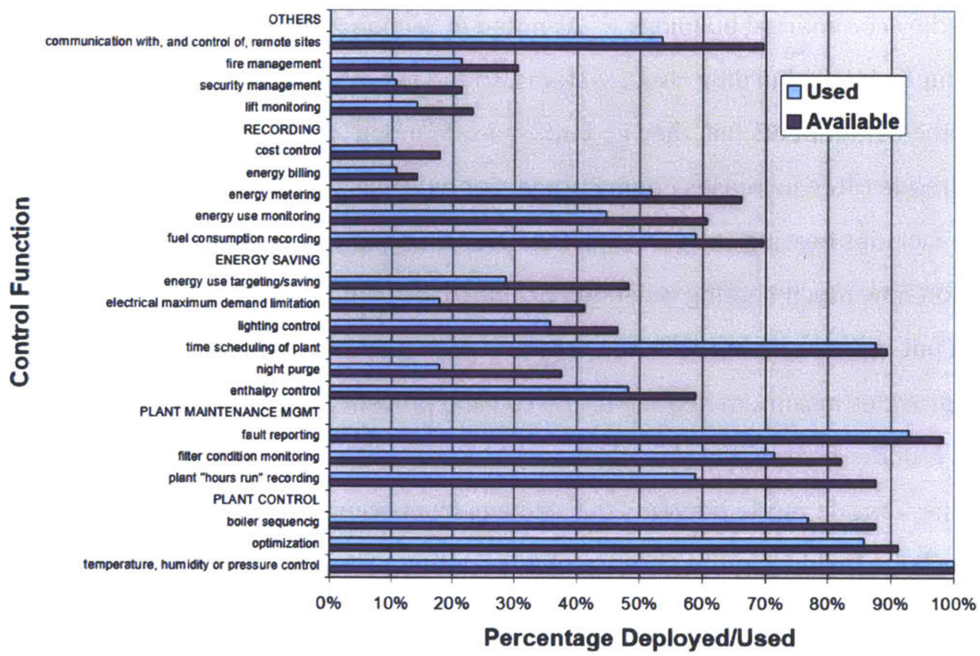


Figure 2.7 Utilization Rate of Building Automation Systems by control function. Source: Lowry (2002)

In terms of the exact energy savings that can be realized by utilizing building automation systems to their full potential, the literature has concluded with many different savings numbers across many different applications. In the 2014 literature review mentioned above, Ken Bruton reviewed many different studies in this domain and concluded an average savings of 5 to 15% savings that can be achieved by monitoring and addressing faults with fault detection systems. Of course, this number varies significantly building to building, as building age, complexity, size, and typical maintenance procedures vary substantially across buildings. Additionally, most of these studies are theoretical and do not evaluate existing fault detection systems in real buildings. In Table 2.2, Bruton provides a comparison of savings from FDD systems as compared with other control and monitoring approaches, such as lighting control and commissioning.

Control Technology	Technical Market Size [billions ft²]	Relevant Primary Energy [quads]	Energy Savings [%]	Technical Energy Savings Potential [quads]	Simple Payback Period [years]	Remaining Market Penetration	Market-Achievable Energy Savings [quads]
Energy Management and Control System (EMCS)	33	6.2	5-15%	0.3 – 0.9	8-10	5-10%	0.02-0.09
Commissioning	55	9.8[(a)]	5-15%	0.5 – 1.5	2-10	3-30%	0.015-0.5
Automatic Fault Detection and Diagnostics (AFDD)/Continuous Commissioning	55	9.8[(a)]	5-15%	0.5 – 1.5	1-3	15-55% / 6-24%	0.07 – 0.8 / 0.03-0.35
Occupancy Sensors for Lighting Control	50	3.5	20-28%	0.7 – 1.0	1 – 5	0-45%	0-0.45
Photosensor-Based Lighting Control	55	3.9	20-60%	0.8 – 2.3	1 – 7	8-55%	0.08-1.3
Demand Controlled Ventilation (DCV)	55	5.4	10-15%	0.5 – 0.8	2 – 3	15-30%	0.08-0.25
(a) This figure includes 0.5 quads for supermarket refrigeration systems and walk-in refrigeration.							

Table 2.2 Energy savings of various control approaches. Source: Ken Bruton (2014)

2.4.2 Literature Review for Fault Detection

In terms of literature evaluation of FDD systems in real buildings, there has been very little done on the topic, as most of these systems have been in place only for a few years. Much of the literature identifies the need for more research on the topic. For example, in his literature review of FDD software, Bruton argues in his conclusions that,

> *"finally, and possibly most importantly, it is critical that tools are tested on a large number of actual AHUs in HVAC systems in a variety of buildings and organizations, with validated energy savings results, in order to build a commercial case for this process."*[39]

Although there could be potential for FDD software to make a large impact on carbon mitigation techniques, it is imperative that we evaluate these systems in order to both build the case for these systems as well as help to improve the industry to optimize energy savings. Improved evaluations of these systems could identify how the system should operate and what the best methodological approaches are for identifying and flagging faults in different kinds of buildings. Although there has

[39] Bruton, Ken et al. 2014. "Review of Automated Fault Detection and Diagnostic Tools in Air Handling Units." *Energy Efficiency* 7.2: 335–351.

been limited literature evaluating these systems in real buildings and comparing it to energy usage, Table 2.3 outlines literature that has attempted to evaluate FDD software in real buildings.

The first study shown in the table by Mary Piette was conducted by the Lawrence Berkeley National Laboratory and examined a commercial office building in San Francisco.[40] The study did not compare the predictions to actual energy usage, and it did evaluate an existing HVAC monitoring system in a real building to estimate performance. The study took data from the BAS to evaluate the performance of the FDD software, and subsequently estimated energy savings based on the equipment in the building. The study found 20% of the energy cost of the building could be saved by implementing corrections to the faults that were identified by the monitoring system. The study further found $20,000 in operational savings reducing preventative maintenance tasks. Although this study did not use actual energy data to make these estimations, and it did not study more than one building, it still provides insight into the potential opportunities for continuous BAS monitoring systems. A second study outlined in the table examined 150 HVAC monitoring systems across the state of Washington, but the study only looked at packaged air conditioners and heat pumps.[41] Although the study had a larger sample than the Piette study, it only looked at two systems and left out many of the other HVAC system that can be monitored with FDD, such as chillers, boilers, VAV boxes, non-packaged air handling units, and many other systems. Additionally, the author did not publish the energy savings from the study. Finally, the third study in the table has both a large sample size and estimated energy savings.[42] This paper was able to find $0.027 to $0.071/KWH of energy cost savings available from faults identified by the FDD software across 20 small commercial buildings in California. Similar to the other two studies however, the study did not look at actual energy metered data, but simply estimated energy data. As we see in detail from Chapter 1 of this thesis, estimated energy savings often over predict the actual savings realized in a building. Thus, it is imperative that more research be done in existing buildings to evaluate fault detection and diagnostics system against actual energy metered data.

Research Study	**Site(s)**	**Estimated savings**	**HVAC units included**	**Data used**

[40] Piette, Mary Ann, Sat Kartar Kinney, and Philip Haves. 2001. "Analysis of an Information Monitoring and diagnostic system to improve building operations." *Energy and Buildings*.
[41] Katipamula, Srinivas. 2008. "Transforming the Practices of Building Operation and Maintenance Professionals: A Washington State Pilot Program." *ACEEE Summer Study on Energy Efficiency in Building.*
[42] Li, Haorong, and James E. Braun. 2007. "Economic Evaluation of Benefits Associated with Automated Fault Detection and Diagnosis in Rooftop Air Conditioners." *ASHRAE Transactions* 113.2: 200–210.

Piette, Mary Ann, Sat Kartar Kinney, and Philip Haves.	1 commercial office building; San Francisco	20% energy cost savings; $20,000 operations savings	Cooling system and air supply	No energy data; only BAS data
Katipamula, Srinivas.	8 million square feet of buildings in Washington state; 150 wireless monitoring systems	No verified savings yet. Study says it should be published at a later date.	Packaged air conditioners and heat pumps only	None. Plans to use energy data.
Li, Haorong, and James E. Braun	20 small commercial buildings in California	$42/KW per year of service cost savings and $0.027 to $0.071/KWH of utility cost savings	packaged rooftop units (only nine types of faults)	No energy data; estimated savings based on a model

Table 2.3 Summary of existing literature of FDD software

Finally, an Association of Heating Refrigeration and Air-Conditioning (ASHRAE) study, "Whole Building Commercial System HVAC Simulation for Use in Energy Consumption Fault Detection" was able to examine real energy data in a commercial, academic building on a college campus.[43] Rather than evaluate an existing FDD system, the research team built a model to evaluate energy usage that could self-identify faults, simply from finding regression discontinuities in the energy data alone. The research team looked at changes in daily energy usage, monthly energy usage, and continuous energy usage to identify and flag faults. Through this methodology, the team was able to identify several faults that were later verified in the field. For example, the model found a sharp increase in hot-water consumption from December 2001 to April 2002. A small increase and chilled water consumption followed the large increase in hot-water consumption. The research team flagged this change as a fault and followed up with facilities to discuss the findings. It was confirmed from discussions with facilities that there had been a hot water valve problem that was identified and repaired in April 2002, likely verifying the changes that the team found in the energy data. The team found that during this time hot-water usage increased by 150% during peak hours, and the cumulative total energy waste was 1200 MMBtu of hot-water, resulting in $12,000 of wasted energy usage. Figure 2.8 shows the sharp increase in hot water

[43] Seung Uk, Lee, et al. 2007. "Whole-Building Commercial HVAC System Simulation for Use in Energy Consumption Fault Detection." *ASHRAE Transactions* 113 (2) 52-61.

(HW) usage for this time period. The research team compiled the average daily hot water usage and calculated the percent change in consumption over this time period to identify this fault.[44]

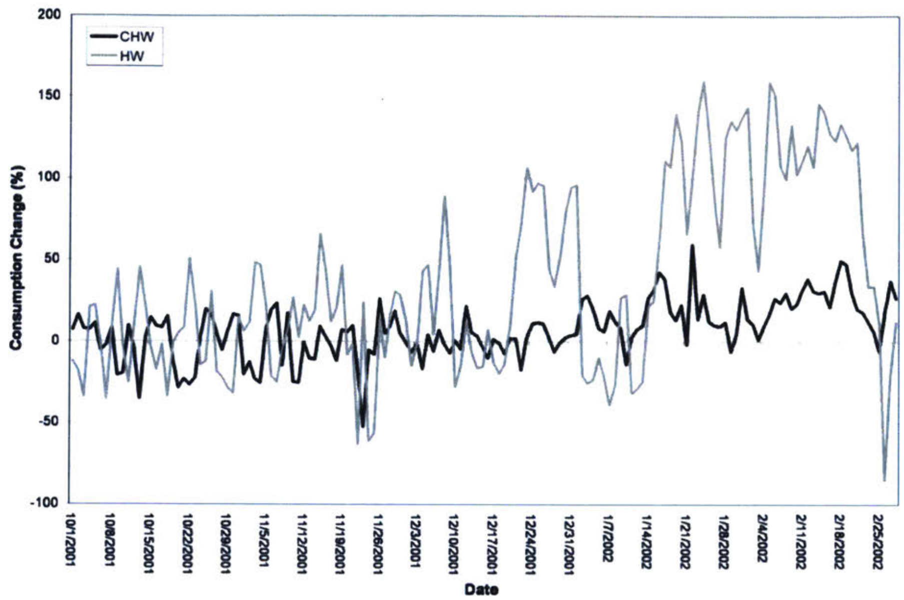


Figure 2.8 Change in average daily hot water (HW) and chilled water (CHW) usage in an academic building. Source: Lee (2007)

The team identified and reported in this research study several additional faults from the energy data analysis that were subsequently verified. Although this study was not evaluating an existing FDD system that analyzes BAS data, the methodological approach could easily be applied to verify the performance of an FDD system by comparing it to actual changes in energy consumption. In theory, this type of analysis, and the type of analysis outlined in this thesis, could be applied to a building on an ongoing basis in addition to a standard FDD system to optimize the results. An automatic FDD system could utilize both techniques, analyzing both BAS data and energy data to better identify faults with reduced false positives and actual energy savings predictions alongside the fault.

[44] Seung Uk, Lee, et al. 2007. "Whole-Building Commercial HVAC System Simulation for Use in Energy Consumption Fault Detection." *ASHRAE Transactions* 113 (2) 52-61.

2.5 Conclusion

This chapter builds off of Chapter 1 by providing more evidence for the energy efficiency gap, specifically as it applies to building automation systems. This chapter outlines the research need for evaluating fault detection diagnostics systems, and chapters 3 and 4 outline the methodologies and results developed for this thesis to fill in these research gaps.

3 Building Energy Use Modeling: Methods & Results

3.1 Introduction

This chapter outlines my methodological approach and results for developing counterfactual building energy use models for large commercial buildings. Chapter 4 then outlines the methodology and results for applying these models to time periods in which major faults in the buildings occurred, with the goal of estimating the energy increase that resulted from the faults.

3.2 Research Questions

The goal of the research outlined in this chapter is to determine if more advanced machine learning algorithms for modeling building energy usage can outperform multiple linear regression models for the given applications. In this chapter I compare the results from four separate modeling techniques to determine which modeling technique is the most accurate for each building application, particularly when the model is applied to a new set of data. As outlined in Chapters 1 and 2, there is a strong need for improved modeling techniques to assess baseline building energy consumption. I compare the performance of each modeling technique across building type and fuel source analyzed (electricity, chilled water, or steam).

3.3 Data Collection

In order to build these models, I used data from two main sources. First, I collected weather data from a weather station that sits on the roof of Building 54 at MIT in Cambridge Massachusetts, also known as the "Green Building". This data was provided to me from the MIT Earth, Atmosphere, and Planetary Sciences "EAPS" department at MIT. I utilized the 10-minute interval temperature and humidity data from 2010 to 2017 provided by this station. Rain rate and other measurements from the data set were not useful to the model and did not improve the accuracy of the models. Several data points with missing temperature or humidity values were removed from the data set. This resulted in a total of only 6.5% of the data being removed from the data set. Thus, the missing data points will only have a negligible impact on the model development.

Figure 3.1 Weather station used for data collection. Source: MIT Synoptic Laboratory

The second data set I used came from MIT Facilities. I collected energy data for each building from two main sources: 1) directly from Facilities, and 2) from "Energize_MIT," a relatively new energy data source available to students and faculty at MIT. Energize_MIT is a portal with data sets provided by the Office of Sustainability. The data from both sources comes in 15-minute intervals directly from campus sub-meters owned and operated by MIT. The available data varies by building, but the majority of buildings that I included in the study have 15-minute interval energy data from 2010 to 2017 for all three energy sources: electricity, steam, and chilled water.

As seen in Figure 2.2, there is a campus central utilities plant that provides electricity, chilled water, and steam to some buildings on campus. Most buildings in this study receive district heating and cooling. The steam and chilled water that are supplied to these buildings are typically sub-metered, providing a valuable data source for modeling applications. Chilled water is a cooling source used in many commercial buildings, in which approximately 44°F water is circulated through the building. The chilled water is passed through chilled water coils in air handling units or other HVAC systems, and air is typically passed over the coils and distributed directly to spaces or via ductwork throughout the building. Steam is used as the main heating source for the buildings in this study, and the steam is either passed directly through steam coils in a similar fashion to chilled water, or its passed through a heat exchanger and hot water is circulated through the building for heating.

In addition to energy data, I also collected information through interviews with facility staff to better understand additional information on the buildings. I collected information on the accuracy of the meters for the buildings they serve; any changes to the buildings that have occurred during the time period investigated; and any issues of concern for these buildings.

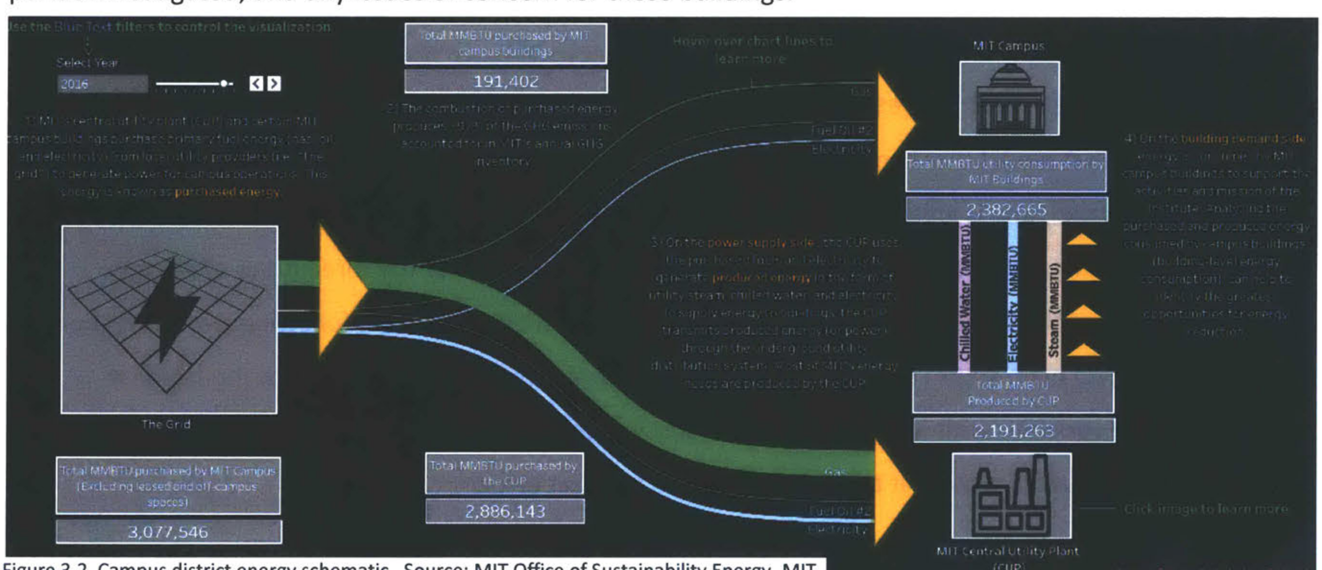


Figure 3.2. Campus district energy schematic. Source: MIT Office of Sustainability Energy_MIT

3.4 Methodology

I used four methods to build counterfactual energy consumption models for each building. I then compared the results for each. For each model I interacted the following variables: temperature, humidity, year, month, day of the week, hour of the day, minute of the day, and the full date and time. The first model is a multiple linear regression model in which I interacted the variables for each model in 20-30 different ways to find the most accurate model. I built the three additional models using three separate machine learning algorithms that drastically improve the accuracy of the models with novel approaches. With machine learning, these approaches are able to optimize the coefficients and selection of integrated variables that yields the most accurate model. These three approaches, Ridge regression, the Lasso model, and elastic net regression model each use different methodologies to select a model with improved accuracy. For each model I broke the data set into two separate data sets: a data set with several years of "training data" to build the model, and a completely new set of "test data" to test the accuracy of the model against. The size of the test set was approximately a quarter of the size of the training data for each use case. By building a model with only the training data and subsequently testing it on brand new test data, I am able to determine how accurate the models are for each building. For each application, I compare the variance forecast error between the model and test data, as defined in Equation 4, across all four models. Below is a summary of the methods used for each model. I utilized R and R Studio to build each model.

3.4.1 Multiple Linear Regression Model

The first model that I analyzed was the multiple linear regression model. In a multiple linear regression model, one is able to build a linear approximation model based on several different variables. For each building, I built a multiple linear regression model for each of the three types of energy consumption at the building level: electricity, steam, and chilled water. The following two subsections outline the two main steps I took for the multiple linear regression models. For each model I first selected the variables with which to test, and I subsequently interacted the variables in a number of different ways until I arrived at the lowest variance forecast error.

3.4.1.1 Variable Selection

The following equations (1-3) summarize the multiple linear regression models and the variables used to define them.

$$E_{b,dt} = T_{dt} + H_{dt} + h + D + M + dt + \epsilon_b \quad (1)$$

$$S_{b,dt} = t_i T_{dt} + H_{dt} + h + D + M + dt + \epsilon_b \quad (2)$$

$$C_{b,dt} = t_i T_{dt} + H_{dt} + h + D + M + dt + \epsilon_b \quad (3)$$

where:

$E_{b,dt}$ = natural log of kilowatt hours (kWh) of electricity consumed by each building at a point in time *dt*

$S_{b,dt}$ = natural log of million British Thermal Units (mmbtu) of steam consumed by each building at a point in time for heating purposes

$C_{b,dt}$ = natural log of tonnage (tons) of chilled water consumed by each building at a point in time for cooling purposes

b = building

dt = date time, in the format of (yyyy/mm/dd) h:mm in fifteen minute intervals

T_{dt} = Outdoor air temperature at a point in time *dt*

H_{dt} = Outdoor air humidity at a point in time *dt*

h = hour of the day from 0 to 24

D = day of the week, Sunday to Saturday

M = month, January to December

ϵ_b = error by building

t_i = temperature coefficient, *i* = 1…6

As noted in the equations, I took the natural log of electricity, steam, and chilled water usage in order to remove negative predictions of energy usage from the models. Because these models are linear approximations, they will yield negative predictions in the absence of this step. However, in real life energy usage cannot be negative, and by estimating the natural log of energy usage, the model's predictive power increased significantly.

The steam and chilled water calculations include a coefficient t_i to account for temperature dependency. Because there is a theoretical limit to the maximum steam that can be used in the building, based on the maximum capacity of the heating systems, a simple linear regression with continuous temperature dependency will over predict steam consumption during low temperatures. For example, Figure 3.3, shows a model of steam consumption in which a continuous temperature variable is used. As seen in the figure, the linear model grossly over predicts steam consumption at low temperatures, creating an almost linear relationship between temperature and steam. However, the actual steam consumption shows that steam usage maxes out at approximately 13000 pounds per hour, regardless of further temperature drop. To address this issue, I created a spline function for temperature, similar to a methodology introduced in a paper by the Lawrensce Berkeley National Lab.[45] I first broke the temperature data into six separate bands of temperature from the minimum outdoor air temperature of -10°F to the maximum outdoor air temperature of 110°F. Thus, I grouped each set of data into a temperature interval, either -10°F to 10°F, 10°F to 30°F, 30°F to 50°F, 50°F to 70°F, 70°F to 90°F, or 90°F to 110°F. I then created six variables, t_1 to t_6, in which t_i is equal to zero when the temperature T_{dt} is outside of the boundaries of the temperature band. t_i is equal to a value between zero and 20, depending on where it sits within the temperature band.

	t_1	t_2	t_3	t_4	t_5	t_6
5°F	5	0	0	0	0	0
17°F	0	7	0	0	0	0
42°F	0	0	12	0	0	0
53°F	0	0	0	3	0	0
89°F	0	0	0	0	19	0
107°F	0	0	0	0	0	17

Table 3.1. Example of variable selection for creating a piecewise linear temperature dependency

By adding these variables, I was able to create spline function for temperature such that the model does not over predict energy consumption at the maximum or minimum loads available to the system. Figure 3.4 shows the same metered steam data as figure 3.3, but the model in figure 3.4 instead utilizes a spline function. As seen in the two graphs, there are significantly less predictions in figure 3.4 that exceed the metered data at low temperatures. Although a different number of temperature intervals can be used for different models, I selected six intervals for this particular data, because the metered data for this building shows a very large maximum steam capacity been with from approximately -10°F

[45] Mathieu, Johanna L. et al. 2010. "Quantifying Changes in Building Electricity Use, with Application to Demand Response." *IEEE Transactions on Smart Grid.*

to 25°F. A larger number of temperature intervals would add more data points at lower temperatures, and a smaller number of temperature intervals would provide too little information to the model.

Figure 3.3 Metered vs counterfactual steam consumption for Building 13 as a function of outdoor air temperature, with temperature as a continuous variable

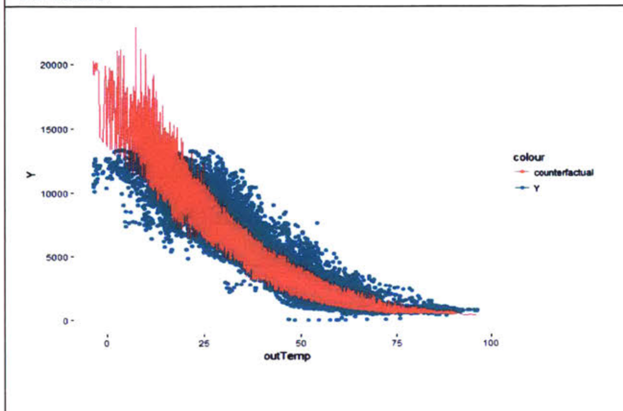
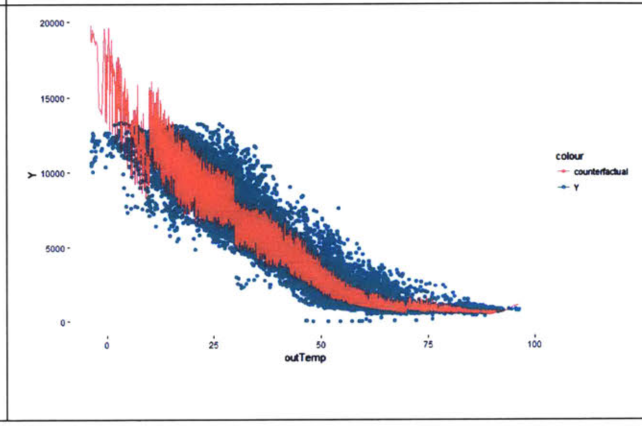


Figure 3.4 Metered vs counterfactual steam consumption for Building 13 as a function of outdoor air temperature using a Spine function



Further, I reviewed and analyzed the models to improve accuracy both at the large-scale through variance and bias estimates, but also on smaller scales at the monthly and daily levels. A model could have high accuracy at predicting energy consumption throughout the year, but it may not yield appropriate results to account for characteristic fluctuations at the daily level. For example, for each of the models I built, I included a categorical variable for "hour" instead of continuous. A continuous variable for hour allows the model to better understand energy consumption hour by hour throughout the day. However, a continuous variable will yield results such as those in Figure 3.5 in which energy consumption is predicted in a continuous manner from hour 0 to 24, but there is a huge gap between hour 24 and hour zero, which is impossible in real life. By instead using a categorical value for hour, as seen in Figure 3.6, I was able to eliminate this issue and yield higher accuracy for both large and small fluctuations in the data.

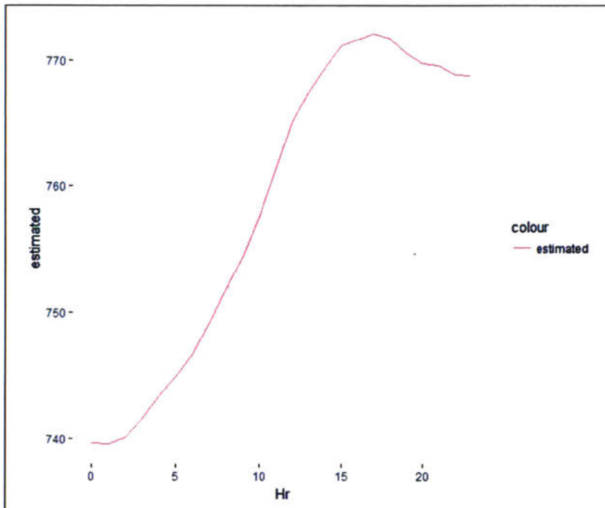


Figure 3.5 average daily load profile modeled using a continuous temperature variable

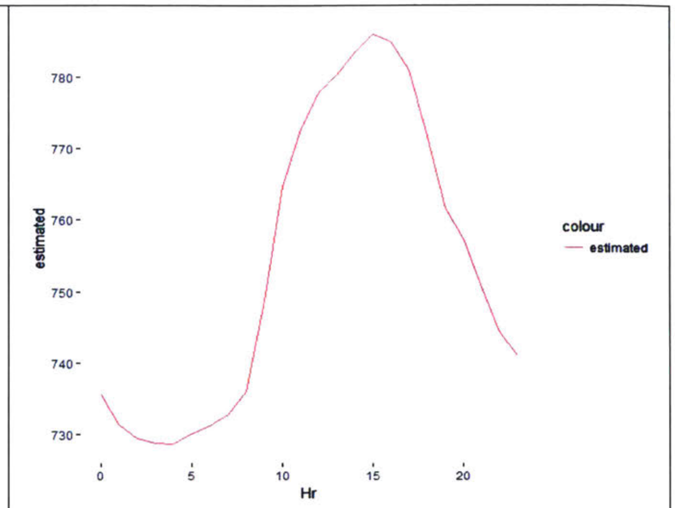


Figure 3.6 average daily load profile modeled using a categorical temperature variable

Finally, while many linear models simply correlate energy data with one time variable, there are a number of useful variables that I was able to extract from the time variable to provide more information to the model. For example, for many of these buildings, month is an important variable to add to the model as a categorical variable. While a continuous month variable would provide information on the continuous changes throughout the year, a categorical month variable shows the large changes that can occur month to month on a college campus. For example, Figure 3.6 shows average daily electricity usage each month, averaged over several years. The figure shows similar electricity usage month the month, with a large drop in usage each November, December, January and February. This building has very low occupancy during Thanksgiving break in November; winter break/the end of the semester in December; MIT's "Independent Activities Period" (IAP) in January and the often the first week of February. By adding month as a separate categorical variable, I was able to capture these month-to-month changes in energy consumption for certain buildings.

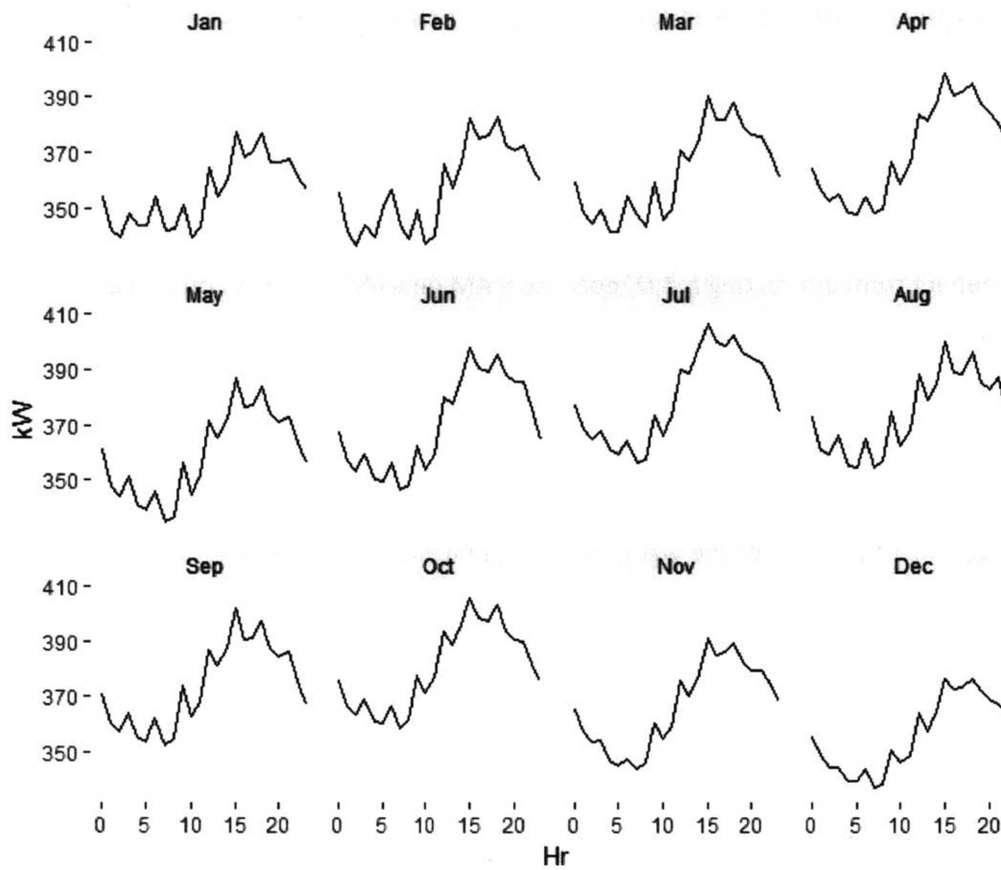


Figure 3.7 average daily load of power consumption in KW by month for Building 66

Each of these examples show the importance of carefully reviewing data and selecting optimal variables for modeling energy consumption. While many models simply correlate energy usage with a time variable and a temperature variable, there are a significant number of additional variables that provide considerably more information to a model.

3.4.1.2 Variable Interaction

After selecting a large sample of variables to use, I then interacted the variables for each model in a number of different ways to find the most accurate model for each energy source. Interacting the variables in different ways can yield very different results and provide even more useful information to the model. For example, by interacting $T_{dt} * H_{dt}$, the model is able to explain more information about chilled water usage and steam usage in some buildings. Temperature and humidity together determine dew point, and some components in HVAC systems are controlled based on enthalpy values, which is a combination of temperature and humidity, rather than just temperature. For these buildings, both temperature and humidity are controlled together to optimize comfort levels, rather than simply cooling a building to a comfortable temperature but an uncomfortable humidity level. Similarly, building models can typically benefit from interacting $h * D$, because 9 AM on a Monday in an office building will look very different than a 9 AM on a Sunday in terms of energy usage. For example, Figure 3.7 shows the average daily load profile for a research building at MIT, Building 76. These graphs show the average power demand for each day of the week from hour zero (midnight) to hour 24 (midnight). The graphs are based on 15-minute interval electric data from 2010 to 2017. As seen in the graphs, the daily load profiles vary significantly by day of the week, with a significantly higher peak load on weekdays than weekends. By interacting weekday with hour of the day, we can provide the model with more information on how energy use varies by weekday*hour.

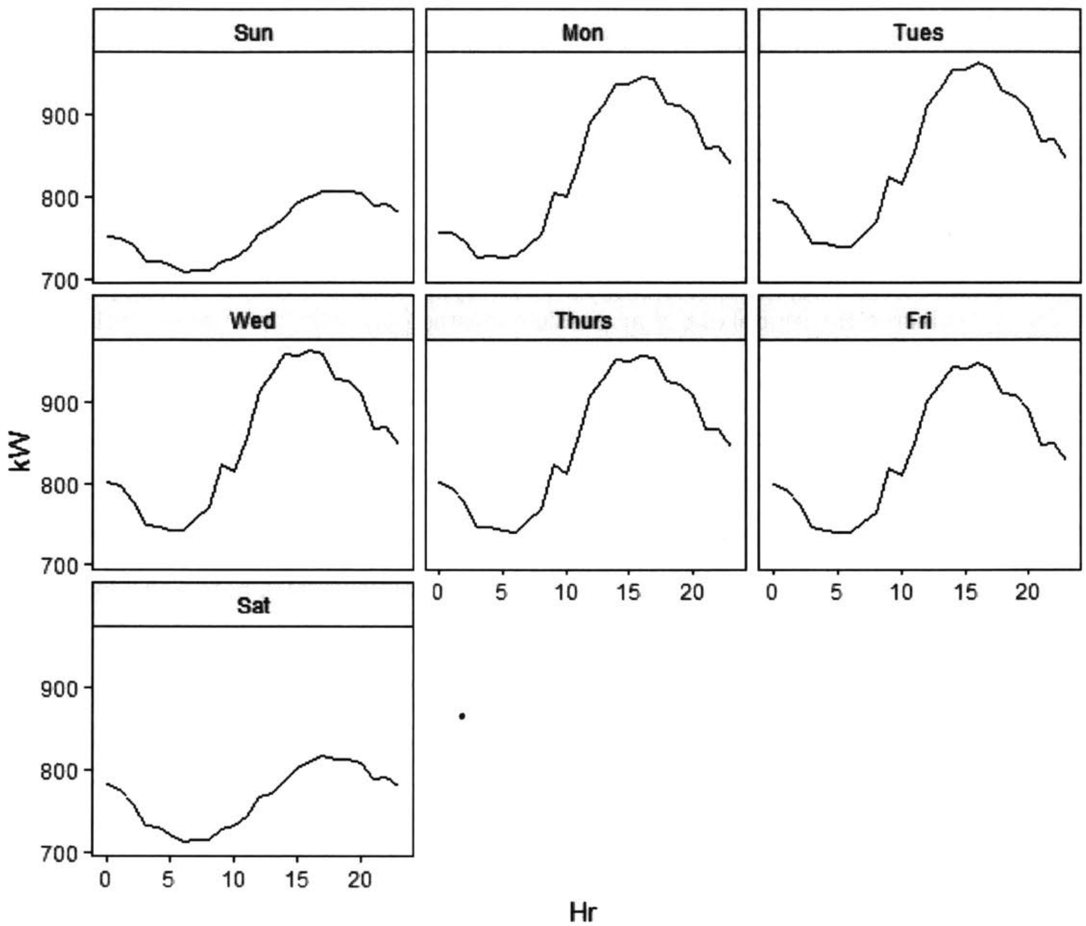


Figure 3.8 Average daily load profile of power consumption in Building 76 at MIT, "The David H. Koch Institute for Integrative Cancer Research." The data is based on 15-minute interval power data from 2010 to 2017.

If instead the model were based only on these variables independently without taking into account their interaction, as many models in the field and in literature do, the model would not capture this additional layer of information. Although these examples are readily understood, there are numerous other interactions available that could yield more accurate results, but are less clear in their meaning. In order to pick the best multiple linear regression model, I built each model option with training data; ran the model against test data; and compared the variance forecast errors, as defined in Equation 4 below.

$$variance\ forecast\ error = \frac{\sum(E_{b,dt,training} - E_{b,dt,test})^2}{n} \quad (4)$$

$$bias = \frac{\sum(E_{b,dt,training} - E_{b,dt,test})}{n} \quad (5)$$

where:

$$n = number\ of\ observations$$

Finally, I selected the model out of all possible interactions with the smallest variance forecast error. Please see the results section for more detail on the interactions selected and the results for each. Of course, this trial and error approach introduces bias in the selection process. It is well-known that this bias can yield very different results depending on the interaction selected by the researcher and by the method of selecting the model with the least variance on test data. Thus, machine learning algorithms that can select variables based on a set of conditions and optimization parameters will likely yield more accurate results. The next three sections walk through the three machine learning approaches that I utilized for this thesis.

3.4.2 Ridge Regression Model

The first machine learning model I used for analysis is the Ridge Regression Model. Ridge regression is able to improve prediction through penalization techniques. The Ridge regression model minimizes the residual sum of squares by restricting the L2 – norm of the coefficients. I developed the models using a built-in functionality in R for Ridge regression, and I fed the model a matrix of all possible interacted variables. I broke the data into the same data sets as the multiple linear regression model, with a training set and a test set.

The goal of this approach was to yield results in addition to those from the multiple linear regression model, but with the reduction of potential bias. However, unlike the next two model techniques used, Ridge regression does not have the ability to eliminate predictors that are not useful to the model. In Ridge regression, all variables are used, and the only adjustments are the sizes of the coefficients. Thus, I continued to compare these first two models with two additional methodologies to derive the most appropriate set of models per building and utility.

3.4.3 The Lasso Model

The next modeling technique that I use for the analysis is the Lasso model, the "least absolute shrinkage and selection operator." Similar to Ridge regression, the Lasso model is able to penalize coefficients, in this case with an L1-penalty. In addition, the model is able to simultaneously eliminate certain variables by bringing the coefficient to zero, and therefore select the most appropriate variable interactions. Robert Tibshirani developed the Lasso model approach in his 1995 paper "Regression Shrinkage and Selection Via the Lasso."[46] The technique has since been used in a variety of applications, but very little has been done in the energy space, as discussed in detail Chapter 1.

While the Ridge regression model uses L2 regularization to penalize coefficients based on on a square of weights, the Lasso model uses an L1 regularization to penalize coefficients based on an absolute value calculation. Thus, the Lasso model is able to bring coefficients to zero, whereas the Ridge regression can reduce coefficients to a very small number. For example, Table 3.2 shows the coefficients selected for Ridge regression first Lasso regression for a model I built for Building 13 steam usage. Often when the Lasso model chooses to bring a variable coefficient or variable interaction coefficient to zero, the Ridge regression model chooses to bring the coefficient to a small number.

variable	Lasso coefficient	Ridge coefficient	variable	Lasso coefficient	Ridge coefficient
(Intercept)	1.30E+01	6.06E+01	Hr29	4.03E-02	2.73E-02
(Intercept)	.	.	outTemp:Hr	-1.13E-05	-2.58E-04
outTemp	-3.65E-02	-1.99E-02	outTemp:outHumidity	-8.25E-05	-1.19E-04
Hr	.	3.44E-03	outTemp:WeekDay.L	-1.02E-04	-1.67E-04
outHumidity	4.34E-03	4.56E-03	outTemp:WeekDay.Q	.	1.84E-04
WeekDay.L	.	1.24E-03	outTemp:WeekDay.C	.	1.57E-04
WeekDay.Q	1.01E-02	1.55E-02	outTemp:WeekDay^4	.	-8.52E-05
WeekDay.C	.	2.18E-04	outTemp:WeekDay^5	.	-1.39E-04
WeekDay^4	-3.46E-02	-2.19E-02	outTemp:WeekDay^6	.	-3.30E-05
WeekDay^5	.	-6.83E-03	Hr:WeekDay.L	.	3.52E-04
WeekDay^6	.	-1.31E-03	Hr:WeekDay.Q	.	8.76E-04
Month.L	-7.04E-02	-1.67E-01	Hr:WeekDay.C	-8.68E-04	-1.04E-03
Month.Q	3.48E-01	6.07E-01	Hr:WeekDay^4	.	1.73E-04
Month.C	1.61E-01	2.34E-01	Hr:WeekDay^5	.	4.40E-04
Month^4	-2.10E-01	-2.84E-01	Hr:WeekDay^6	.	-5.13E-04
Month^5	-1.19E-01	-1.55E-01	Hr:outHumidity	3.36E-05	1.44E-04
Month^6	-7.73E-03	6.61E-03	outHumidity:WeekDay.L	.	5.54E-05
Month^7	-2.64E-02	5.72E-03	outHumidity:WeekDay.Q	-1.36E-04	-1.55E-04
Month^8	6.65E-02	7.03E-02	outHumidity:WeekDay.C	.	-2.96E-05
Month^9	3.97E-02	1.30E-02	outHumidity:WeekDay^4	.	-1.46E-04
Month^10	1.67E-02	2.06E-02	outHumidity:WeekDay^5	.	1.40E-05
Month^11	1.78E-02	2.88E-02	outHumidity:WeekDay^6	.	5.79E-05
Year	.	-2.47E-02	outTemp:T1	-4.35E-03	-8.70E-04
Time	.	-2.06E-05	outTemp:T2	1.88E-03	1.17E-04
dateTime	-2.30E-09	-1.33E-09	outTemp:T3	.	1.04E-04
T1	-1.68E-02	-2.29E-03	outTemp:T4	-6.38E-04	-1.62E-04
T2	-5.92E-02	1.32E-03	outTemp:T5	.	.
T3	8.49E-03	9.03E-03	outTemp:T6	.	2.92E-04
T4	2.50E-02	-6.98E-03	outTemp:Hr:WeekDay.L	.	-5.36E-06
T5	.	.	outTemp:Hr:WeekDay.Q	2.47E-05	1.83E-05
T6	1.23E-01	3.28E-02	outTemp:Hr:WeekDay.C	.	-3.01E-06
Hr21	-6.13E-03	-1.43E-02	outTemp:Hr:WeekDay^4	2.34E-05	1.66E-05
Hr210	4.28E-02	1.42E-02	outTemp:Hr:WeekDay^5	.	1.70E-06
Hr211	4.07E-02	5.48E-03	outTemp:Hr:WeekDay^6	.	-6.90E-06
Hr212	3.59E-02	-6.80E-03	outTemp:Hr:outHumidity	-2.58E-07	-1.25E-06
Hr213	2.73E-02	-1.22E-02	outTemp:outHumidity:We	.	-2.60E-06
Hr214	1.94E-02	-1.57E-02	outTemp:outHumidity:We	.	-1.75E-06
Hr215	1.09E-02	-1.57E-02	outTemp:outHumidity:We	.	-2.99E-07
Hr216	.	-1.23E-02	outTemp:outHumidity:We	.	5.34E-07
Hr217	.	-6.90E-03	outTemp:outHumidity:We	.	1.17E-06
Hr218	.	1.34E-02	outTemp:outHumidity:We	1.09E-06	2.48E-06
Hr219	.	2.63E-02	Hr:outHumidity:WeekDay	.	9.66E-06
Hr22	-8.81E-04	-1.96E-03	Hr:outHumidity:WeekDay	.	-1.39E-05
Hr220	.	3.74E-02	Hr:outHumidity:WeekDay	.	-7.91E-07
Hr221	-8.76E-03	4.09E-02	Hr:outHumidity:WeekDay	.	6.77E-06
Hr222	-1.71E-02	4.16E-02	Hr:outHumidity:WeekDay	.	4.65E-06
Hr223	-2.82E-02	3.99E-02	Hr:outHumidity:WeekDay	.	1.99E-06
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Hr28	4.07E-02	4.79E-02	outTemp:Hr:outHumidity	.	8.12E-08

Table 3.2 Lasso verse Ridge coefficient selection for a steam model

The Lasso model is useful in cases where there are a large number of unknown interactions between the predictors, and it is unclear which variables or how many variables to include in the model. This is the case for the case studies analyzed in this thesis, as there are likely interactions between the variables that are unknown. As noted in the literature, the ability to select key variables and interactions

---

46 Tibshirani, Robert. 1996. "Shrinkage and selection via the lasso." *Journal of the Royal Statistical Society.* 58 (1) 267-288

is an integral component to building an accurate model for prediction.[47] The Lasso model allows us to select these key variable interactions and is an improvement upon the trial and error approach of the multiple linear regression model.

Finally, the Lasso model is able to tradeoff between variance and bias. Variance measures the spread of data, whereas bias measures the average difference. Two models of power consumption for building 76 show the difference between a typical multiple linear regression model estimation and a Lasso model estimation. As seen in Figure 2.5, while actual energy consumption (shown in red) fluctuates from 700 to 1000, the predicted energy consumption (shown in green) hovers around 800. Although the bias appears to be very low, the model is not capturing the fluctuations in data and thus has a very large forecast error. However, Figure 2.6 shows a Lasso model built on the same data, with predicted energies consumption (in green) able to capture significantly more of the fluctuations in data than seen in Figure 2.5. The Lasso model is able to trade-off between these two measures of error to provide a more accurate model.

[47] Hsu, David. 2015. "Identifying key variables and interactions in statistical models of building energy consumption using regularization." *Energy*.

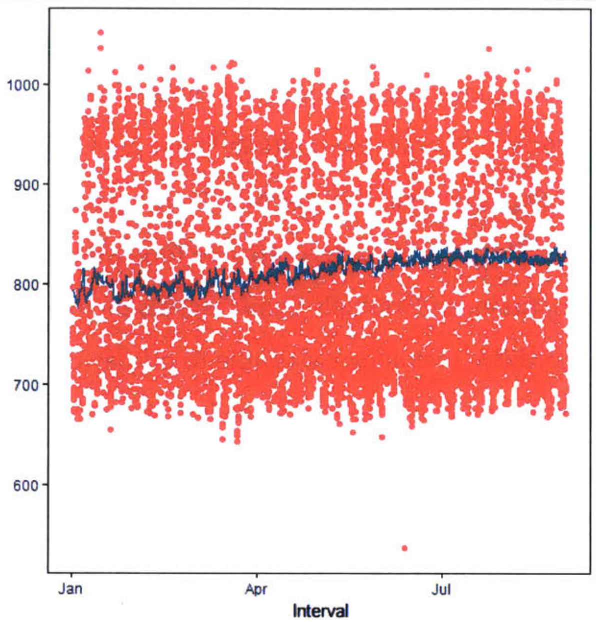


Figure 3.9 multiple linear regression model predictions (green) compared to actual energy consumption (red) for Building 76

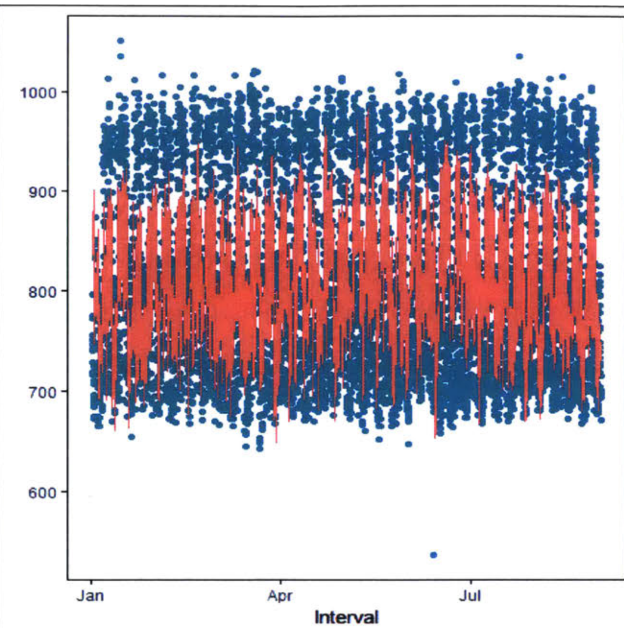


Figure 3.10 Lasso model predictions (red) compared to actual energy consumption (red) for Building 76

3.4.4 Elastic Net Model

Finally, the elastic net model is a combination of the Lasso and Ridge modeling techniques. With the elastic net technique, the model can use the L1 penalty in part and the L2 penalty in part by adjusting the value of α anywhere from 0 to 1. When the value for α is set to zero, the model is equivalent to a Ridge regression model, and when the value for α is equal to one, the model is equivalent to a Lasso regression model. The Elastic Net model was developed by Hui Zou and Trevor Hastie and published in 2005.[48] In certain situations, the elastic net model has proven to make more accurate predictions than the Lasso model. For example, this technique has the potential to more effectively address collinearity[49] For these reasons, I compared all four models for each application, and Section 4.5 outlines the results for each.

[48] Hui Zou and Trevor Hastie. 2005. "Regularization and Variable Selection via the Elastic Net." *Journal of the Royal Statistical Society*. 2 () 301.
[49] Ibid.

3.5 Results

This chapter outlines the results for two of the building energy use models by comparing the performance of each type of model for each building and utility. Overall, the multiple linear regression models occasionally outperform the other models in reducing bias, while the machine learning models consistently outperformed the multiple linear regression model in terms of the variance forecast error for all eight models. Between elastic net, Ridge regression, Lasso, and multiple linear regression models, there was at least one model application where one of the above techniques was able to outperform the others in terms of bias. However, as seen in the tables below, one of the machine learning models was always able to outperform the multiple linear regression model in terms of forecast error. Below are two examples of the multiple linear regression models compared to the machine learning models. Chapter 4 compares the three machine learning models for all 8 models.

3.5.1 Building 68 Electricity Modeling

For the building 68 electricity model, the Ridge regression had the lowest forecast error and the lowest bias. Because forecast error is an important factor for evaluating energy prediction models, due to the large variability across daily usage and monthly usage, the Ridge regression model was selected for this building for the analysis in Chapter 4. Table 3.3 shows the performance comparison across models for both variance and bias when the model was applied to test data. The bias error is a measurement of bias as a percentage of the average electricity usage over the course of the test data.

Method	variance forecast error	square root forecast error	bias	average Ytest	error (bias)
Elastic Net	3674	48	16	1165	1.3%
Ridge Regression	3674	48	16	1165	1.3%
Lasso Regression	3757	49	22	1165	1.9%
Multiple Linear Regression	5338	73	73	1165	4.4%

Table 3.3 performance comparison of all four models for building 68 electricity predictions

Further, Table 3.4 shows the variance forecast error between the model and the test data for a number of various variable selection and interactions. As seen in the table, although I was able to reduce the forecast errors significantly through this trial and error approach, the variance forecast error from the multiple linear regression model was never able to drop to the level of that produced by the Lasso model. However, it did outperform the Ridge regression model and the elastic net model.

variance forecast error	Temperature	Humidity	Month	Weekday	Hour	Interval	Interval*Temp	Temperature *Hr	weekday*hr	weekday*hr *Month	Year	Temperature*Humidity	Temperature*Humidity* Hr	temp*humidity*interval
11,249	X													
14,748		X												
12,704			X											
13,669				X										
14,225					X									
8,410						X								
5,481	X		X	X	X	X								
5,414			X	X	X		X							
5,407		X	X	X	X		X							
5,400			X	X	X		X					X		
9,689			X	X	X			X				X		
5,424			X	X			X						X	
5,458					X		X					X		
5,338			X				X		X			X		
5,341			X				X		X		X	X		
5,417			X				X		X	X	X	X		
5,371			X				X		X	X	X	X		X

Table 3.4 variance forecast error iterations of variable selection and interaction run through the multiple linear regression model for building 68

3.5.2 Building 76 Electricity Modeling

In this case, the elastic net outperformed each of the other three models in terms of variance forecast error, with the Lasso model very close. However, the multiple linear regression model outperformed each of the models in terms of bias error. Regardless, each of the models were able to perform well in terms of bias with percent bias equal to or below 1% of the average electricity usage. Similar to Table 3.3, Table 3.6 shows that the variance forecast error was able to drop with the trial and error approach of the multiple linear regression model, but it never dropped as low as the forecast developed by the three machine learning models.

Method	variance forecast error	square root forecast error	bias	average Ytest	error (bias)
Elastic Net	1,123	25	2.0	830	0.25%
Ridge Regression	1,271	28	2.8	830	0.33%
Lasso Regression	1,125	25	2.0	830	0.24%
Multiple Linear Regression	5,828	76	1.1	830	0.13%

Table 3.5 performance comparison of all four models for building 76 electricity predictions

variance forecast error	Temperature	Humidity	Month	Weekday	Hour	Interval	Interval*out Temp	Temperature *Hr	weekday*m onth	weekday*hr *Month	Year	Temperature *Humidity	Temperature* Humidity*Hr	weekday* hr*Month* year
9,992	X													
9,851			X											
8,103				X										
8,583					X									
9,860						X								
6,275	X		X	X	X	X								
6,284			X	X	X		X							
5,961		X	X	X	X		X							
5,968			X	X	X		X					X		
5,992			X	X	X			X				X		
5,860			X	X				X					X	
5,828			X	X				X			X		X	
5,831			X	X		X					X		X	
5,884						X			X		X		X	
5,863						X				X	X		X	
6,337						X							X	X

Table 3.6 Variance forecast error iterations of variable selection and interaction run through the multiple linear regression model for Building 76

3.6 Conclusion

In conclusion, each of these buildings were able to achieve very low bias and forecast errors when applied to test data, indicating that the models are able to achieve high accuracy predictions. The optimal model selection varied somewhat by building and by utility type, indicating that building multiple types of models per application allows for improved accuracy in building energy modeling predictions. With the limited number of applications, it was not possible to determine a statistically significant correlation between the optimal model and the building characteristics. However, I was able to take advantage of the optimal model selection to use in the analysis in Chapter 4. Chapter 4 applies these models to periods of time in which a fault occurred, in order to determine the energy increase associated with the fault.

4 Evaluation of Energy Impact of Faults: Methodology & Results

4.1 Introduction

This chapter outlines the methodology and results used to estimate the energy increased caused by a fault identified by the KGS system. While Chapter 3 focuses on building the counterfactual energy consumption models, this section applies those models to time periods in which a fault occurred to calculate the energy increase caused by the fault.

4.2 Research Question

There are two main research questions answered by the methodology and results of this chapter. First, the goal of this chapter is to determine whether the energy consumption of faults can be isolated by using 15-minute interval energy data and sophisticated statistical machine learning algorithms. Second, an additional question addressed by this chapter is whether the energy savings predictions made by the KGS system accurately predict energy impact.

4.3 Methodology

After building a baseline energy model for each building and validating it against test data, I then applied these models to a period of time in which a fault occurred to estimate the energy impact of the fault. The methodology for this section consisted of three main steps, 1) selecting the faults to test out of the thousands of faults identified on the system; 2) applying the model to the period of time in which a fault occurred to estimate the energy impact; and 3) comparing these energy savings numbers to the predicted energy savings identified by KGS. Figure 4.1 shows an example of faults that are provided by the KGS system. As seen from the figure, building 18 had over 500 faults identified by the system from July 1 to July 31 in 2017. The system provides additional detail on each fault, including some of the data used to identify the fault along with details on what was found.

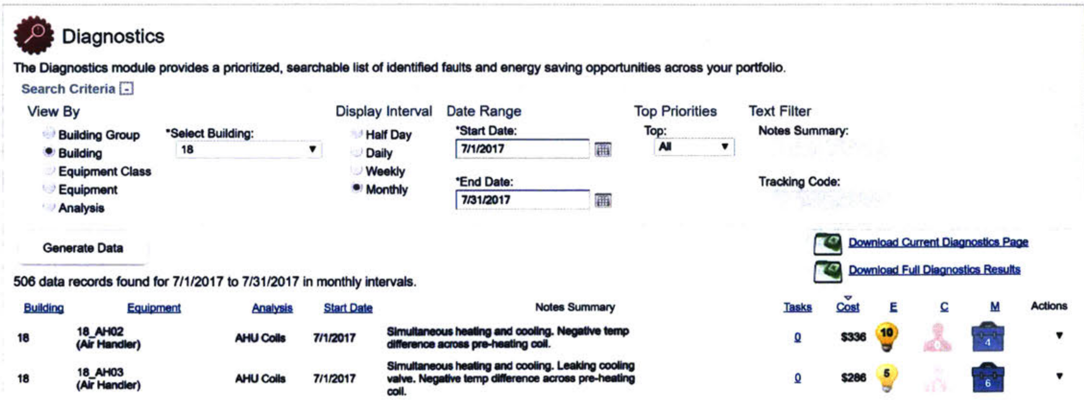


Figure 4.1 Example of fault data provided by KGS. Source: KGS Clockworks

4.3.1 Selecting the Faults to Analyze

Although there are thousands of faults identified by the system, we only tested a subset of the largest faults that were identified. Future work will require a simpler selection process of the faults to analyze. For several reasons, it would not currently be possible to analyze energy consumption of each fault. First, some buildings had gaps in energy data or captured additional buildings in their metered data, and thus these buildings were not good candidates to evaluate the FDD system against energy data. Additionally, some faults are too small to analyze via energy consumption. The margin of error for each model varied from 0.5% to 10%, and some of the smaller, less significant faults, such as a single room overheated by a few degrees in a 200,000 square-foot building, fall within the margin of error of the models. However, some of the larger faults that waste more energy and are also thus higher priority to analyze, are visible within the energy data. Finally, because of the high number of faults, many of them have not been addressed yet by MIT Facilities, and thus the fault does not change status from when the software was implemented until present. With these in mind, I developed the following set of criteria for selecting faults to analyze. These criteria should be met if a similar study is carried out on a fault detection system.

1) I analyzed faults in only those buildings that had a full set of electricity, chilled water, and steam data from approximately 2010 to 2017. I eliminated buildings that have utilities that serve other buildings, or are suspected of supplying energy to other buildings.
2) I selected only those faults with high estimated energy savings (>$600/month) estimated by KGS, or faults that are likely to yield significant energy waste by the

nature of the fault. For example, faults that affect major HVAC systems, such as air handlers, chillers, or boilers, are likely to yield higher energy waste than a sub-system, such as VAV terminal box in a classroom.

3) Next, I removed faults from the subset that did not change status during the time period in which faults were being analyzed by the KGS system. The fault detection system was installed in most buildings in this study in 2014, but many of the faults have not been fixed since they started occurring pre-installation of the system. I only analyze faults that either began occurring at some point between 2014 and 2017, or faults that were fixed during the time period. With this methodology, it is necessary to build counterfactual energy consumption in the absence of the fault in order to compare the model with the actual usage.
4) Finally, some buildings had too many faults overlapping that created noise in the data, making it impossible to analyze each fault based only on total metered data. In the future, issues such as these could be addressed in some circumstances with sub meters. Additionally, as fault detection and diagnostics become more widely used and streamlined within a building such that faults are fixed on a more frequent matter, it will likely be possible to capture more faults within the data.

4.3.2 Evaluation of Energy Impact

This part of the analysis builds off of the methodology outlined in Chapter 3. In chapter 3, we built baseline counterfactual energy consumption and compared the model's predictions to actual energy consumption with a new set of test data to ensure a highly predictive power. In this chapter, I then apply the model to periods in which faults occurred to see if there is a quantifiable increase in energy consumption and compare the increase to the estimated energy consumption from KGS, if this number is provided. We call the period in which the fault occurred the "treatment" group.

$$Kwh\ Increase\ Caused\ by\ the\ Fault = \left[\frac{\sum(E_{b,dt,treatment} - E_{b,dt,model})}{n}\right] * h_{fault} \qquad (6)$$

$$Steam\ mmbtu\ Increase\ Caused\ by\ the\ Fault = \left[\frac{\sum(S_{b,dt,treatment} - S_{b,dt,model})}{n}\right] * h_{fault} \qquad (7)$$

$$Cooling\ CHW\ Increase\ Caused\ by\ the\ Fault = \left[\frac{\Sigma(C_{b,dt,treatment} - C_{b,dt,model})}{n}\right] * h_{fault} \tag{8}$$

where:

$$n = number\ of\ observations$$

$$h_{fault} = number\ of\ hours\ the\ fault\ occured$$

$E_{b,dt}$ = *natural log of kilowatt hours (kWh) of electricity consumed by each building at a point in time dt*

$S_{b,dt}$ = *natural log of million British Thermal Units (mmbtu) of steam consumed by each building at a point in time for heating purposes*

$C_{b,dt}$ = *natural log of tonnage (tons) of chilled water consumed by each building at a point in time for cooling purposes*

b = *building*

dt = *date time, in the format of (yyyy/mm/dd) h:mm in fifteen minute intervals*

In order to verify that the increase in energy usage did not fall within the margin of error of the model and was indeed a result of the faults, I compared the variance and bias of the model on the treatment data with the various and bias of the model on the testing data. For example, I compared the bias for the test data as a percentage of the average value for test data with the bias for the treatment data as a percentage of the average value of the treatment data. To do this, I calculated these errors as defined in the following equations for electricity:

$$\epsilon_{bias,test} = \frac{\frac{\Sigma(E_{b,dt,test} - E_{b,dt,model})}{n}}{\frac{\Sigma(E_{b,dt,test})}{n}} \tag{9}$$

$$\epsilon_{bias,treatment} = \frac{\frac{\Sigma(E_{b,dt,treatment} - E_{b,dt,model})}{n}}{\frac{\Sigma(E_{b,dt,treatment})}{n}} \tag{10}$$

for steam usage:

$$\epsilon_{bias,test} = \frac{\frac{\Sigma(S_{b,dt,test} - S_{b,dt,model})}{n}}{\frac{\Sigma(S_{b,dt,test})}{n}} \tag{11}$$

$$\epsilon_{bias,treatment} = \frac{\frac{\Sigma(S_{b,dt,test} - S_{b,dt,model})}{n}}{\frac{\Sigma(S_{b,dt,treatment})}{n}} \tag{12}$$

and for chilled water usage:

$$\epsilon_{bias,tonnage} = \frac{\frac{\Sigma(S_{b,dt,test} - S_{b,dt,model})}{n}}{\frac{\Sigma(S_{b,dt,test})}{n}} \tag{13}$$

$$\epsilon_{bias,tonnage} = \frac{\frac{\Sigma(S_{b,dt,test} - S_{b,dt,model})}{n}}{\frac{\Sigma(S_{b,dt,test})}{n}} \tag{14}$$

Similarly, I compared the errors in terms of variance for both the model as it is performed on the test data and the model is performed on the treatment data. It is important to test both of these, because the variance and the bias provide very different sets of information on the model forecasting accuracy. For example, as mentioned in chapter 3, multiple linear regression models typically outperform lasso models in terms of predictions on bias, but lasso models outperform linear regression when variance is being evaluated. I compared the errors in terms of variance between the model's performance on test data in treatment data according to the following equations, for electricity:

$$\epsilon_{variance,test} = \frac{\sqrt{\frac{\Sigma(E_{b,dt,test} - E_{b,dt,model})^2}{n}}}{\frac{\Sigma(E_{b,dt,test})}{n}} \tag{15}$$

$$\epsilon_{variance,treatment} = \frac{\sqrt{\frac{\Sigma(E_{b,dt,treatment} - E_{b,dt,model})^2}{n}}}{\frac{\Sigma(E_{b,dt,treatment})}{n}} \tag{16}$$

for steam models:

$$\epsilon_{variance,test} = \frac{\sqrt{\frac{\Sigma(S_{b,dt,test} - S_{b,dt,model})^2}{n}}}{\frac{\Sigma(S_{b,dt,test})}{n}} \tag{17}$$

$$\epsilon_{variance,treatment} = \frac{\sqrt{\frac{\Sigma(S_{b,dt,treatment} - S_{b,dt,model})^2}{n}}}{\frac{\Sigma(S_{b,dt,treatment})}{n}} \tag{18}$$

and for chilled water models:

$$\epsilon_{variance,test} = \frac{\sqrt{\frac{\Sigma(C_{b,dt,test} - C_{b,dt,model})^2}{n}}}{\frac{\Sigma(C_{b,dt,test})}{n}} \tag{19}$$

$$\epsilon_{variance,treatment} = \frac{\sqrt{\frac{\Sigma(C_{b,dt,treatment} - C_{b,dt,model})^2}{n}}}{\frac{\Sigma(C_{b,dt,treatment})}{n}} \tag{20}$$

If the error for predicted energy usage is much higher on the treatment data set than the test data set, then the model is more likely to be accurately estimating the energy impact of the fault. I provide each of these values for each model in the results section.

4.4 Results

This section outlines the results of evaluating the energy impact of eight faults across four buildings. As seen in the table below, all eight of these faults showed an increase in energy usage after the fault occurred, and each model had error rates of 12% or less. However, some of the models had higher predictive power than others. While the third model had an error of 11% on the test data, most of these models had errors of 4% or less, with one model having only 0.25% error. Further, the savings predicted by these models were close to the savings predicted by KGS, with some overestimating and some underestimating the savings. Finally, it is important to note that all three modeling techniques, Ridge, lasso and elastic net, proved to be most effective in certain applications and were thus each use for the evaluation below.

building	fault	utility	model methodology	Model error against test data	Percent increase in energy after fault	Total increase in energy due to fault (one month)	Total increase in energy due to fault (one month)	Total increase in energy cost due to fault (one month)	KGS estimate of increase in energy due to fault (one month)
13	Supply fan at constant speed	Electricity	Ridge	4.3%	10.2%	255527	145267	$ 15,979	$ 10,131
13	Exhaust fans at constant speed	Steam	Ridge	3.0%	9.9%	247640	136107	$ 1,061	NA
13	Exhaust fans a constant speed	Chilled water	Lasso	11.3%	37.8%	57863	43624	$ 3,054	NA
76	three supply fans running full speed	Electricity	Elastic Net	0.25%	3.5%	21396	19930	$ 2,192	NA
68	exhaust fans running constant speed	Electricity	Ridge	1.3%	4.5%	37004	25772	$ 2,835	$ 2,498
68	simultaneous heating cooling	Chilled water	Elastic Net	8.5%	11.2%	43246	17929	$ 1,255	$ 1,078
18	fans running constant speed	Electricity	Lasso	3.4%	4.5%	36	10	$ 792	$ 1,851

Table 4.1 Summary table of results. (note energy use is measured in KWh for electricity, pounds for steam, and tonnes for chilled water.)

It is important to note that while the above table shows the overall average increase associated with each fault, the graphs shown in the following sections provide significant additional insight into the faults by showing the change over time as well as the daily load profile shifts by either average day, by average week day, hour by average day each month. The next two sections outline further detail on each fault and the findings identified by this methodology.

4.4.1 Building 13 Results

Building 13, the Bush building, houses offices, labs, and classrooms. It has a typical 9-to-5 Monday through Friday schedule, with some usage on nights and weekends. Figure 4.2 shows a picture of the building and the location of the building on a campus map. For this building, I analyzed one fault against its electricity usage (4.4.1.1), and the second false against its steam usage (4.4.1.2) and chilled water usage (4.4.1.3).

Figure 4.2 Building 13 (left) and the location of the building on the campus map (right) courtesy of MIT

4.4.1.1 Impact of a Fault on Electricity Usage

I used my model of electricity usage in Building 13 to analyze a fault in which an air handler was running at a constant speed in Building 13. Air handlers supply conditioned air throughout a building, and the fan speed modulates in a variable air volume system such that spaces only receive the amount of air needed at any given time. However, the fault detection software identified a fault in which the fan for an air handler in this building was running at a constant speed from January 2015 to March 2016. Table 4.1 shows the dates I used to build the model with training data, the dates I use to test the model on, and the dates I used to evaluate the model after the "treatment". This fault could have occurred for a number of reasons. The building operator, technician, or other staff member may have manually overridden the fan to run at full speed for a specific need and forgot to remove the override. There could also be sensor issues, control issues, or stuck dampers in the terminal boxes. No matter what the exact reasoning is, this fault is likely wasting energy, as verified in the results below.

training data		test data		treatment data	
start	end	start	end	start	end
1/1/2010	5/31/2014	6/1/2014	12/31/2014	1/1/2015	3/1/2016

Table 4.2 breakdown of dates used to build the model, test the model, and analyze the fault

As seen in Table 4.2, the error for the model against the test data was less than 0.5%, while the error for the model on the treatment data was 4%. Thus, the data after the fault occurred shows an increase in electricity usage.

Method	variance forecast error	square root forecast error	bias	average Ytest	error (bias)	bias	average Ytreatment	error (bias)	mean difference (Ymodel - Ytreatment data)
Elastic Net	522245	479	153	3537	4%	355	3478	10%	355
Ridge	522245	479	153	3537	4%	355	3478	10%	355
Lasso	663976	485	105	3537	3%	345	3478	10%	345

Table 4.3 Results from testing the model on test data set and the treatment data

Figures 4.3 and 4.4 show the average daily profiles for the training data and the treatment data. Figure 3 shows the counterfactual electricity consumption very closely tied to the actual metered electricity consumption for data during the test period. This close correlation in the visual, in addition to the average bias error of 0.3%, suggests that the model is highly accurate in its ability to predict electricity consumption for this building. Figure 4 shows the counterfactual electricity consumption produced by the model compared to actual electricity consumption from the metered data. As seen in figure 4, the metered data is significantly higher than the counterfactual data. In fact, the graph shows that the metered data exceeds the counterfactual data approximately before 10 AM and after 2 PM. This could be explained by the fan typically running at full speed between the hours of 10 AM and 3 PM to meet peak loads. Thus, the difference between the counterfactual and metered data is much more significant when the time is well outside of business hours and the fan would typically run at very low speeds.

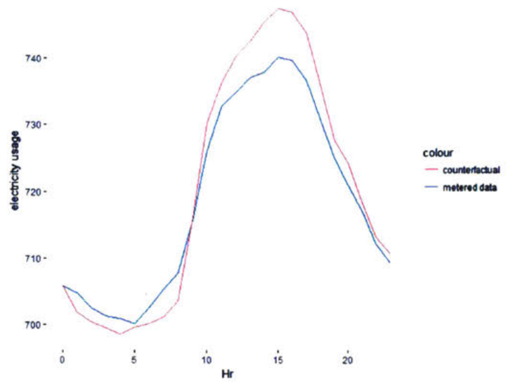


Figure 4.3 Average daily load profile comparison of counterfactual electricity data from the model and actual electricity data from the meter over the test period

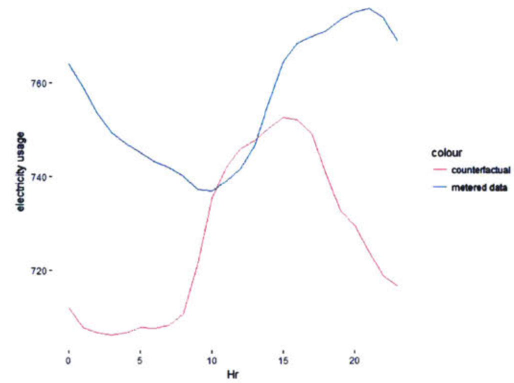


Figure 4.4 Average daily load profile comparison of counterfactual electricity data from the model and actual electricity data from the meter over the treatment period

Figure 4.5 shows the daily load profiles by day of the week, which tell even more information about this fault and provide compelling evidence for the energy loss estimates. Similar to how we would expect nonbusiness hours to have a larger discrepancy between counterfactual and metered data, we would expect the same on Saturdays and Sundays when the classrooms and offices are not typically in use. As seen in Figure 4.5, there is a much larger gap between metered and counterfactual data on Saturdays and Sundays than any of the weekdays.

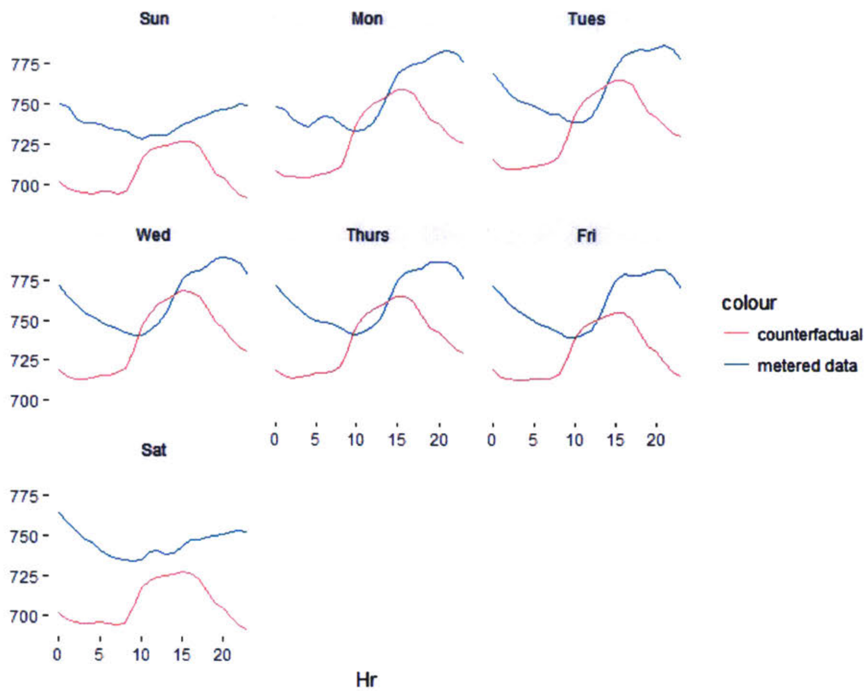


Figure 4.5 Average daily load profile comparison of counterfactual and metered electricity data over the treatment period

Finally, Figure 4.6 shows when the fault began occurring, designated by a black line, overlaid with the counterfactual and metered energy data over the test period (before the fault) and the treatment period (after the fault). As seen in the figure, the metered data exceeds the counterfactual data after the fault occurred.

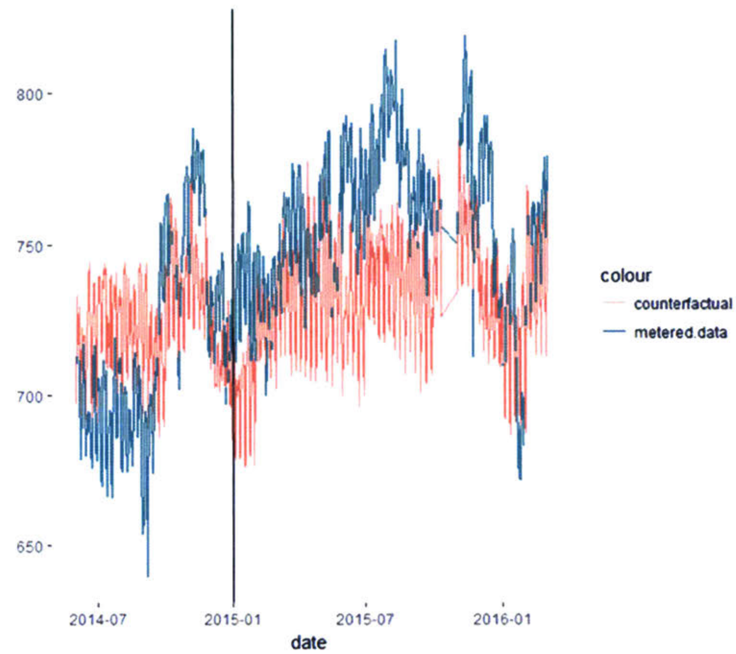


Figure 4.6 Metered Kw data compared to counterfactual data with the black line designating the fault, the data proceeding the black line designating the test period, and the data following the black line designating the treatment period.

4.4.1.2 *Impact of a Fault on Steam Usage*

Next, I analyzed a second fault in Building 13 against steam and chilled water usage in the building. This section outlines the results from steam usage. This particular fault began occurring in January 2015, and table 4.3 shows the breakdown of data I used to build, test, and analyze the model. According to KGS, five exhaust fans in the building were running at constant speed as of January 2015. Because these fans were running at constant speed and likely exhausting more air than necessary, new air needed to be brought into the building and heated and cooled. Thus, this fault appears to have an impact on both steam and chilled water usage in the building.

training		test		treatment	
start	end	start	end	start	end
1/1/2010	12/31/2014	1/1/2014	12/31/2014	1/1/2015	7/1/2017

Table 4.4 Breakdown of dates used to build the model, test the model, and analyze the fault

As mentioned previously, the Ridge regression model had the lowest variance forecast error when the model was applied to test data, so I used the Ridge regression model for the following analysis. While the test data shows an error in terms of bias of 4%, the treatment data had an error of 10%, yielding an overall steam usage increase as compared with the model. Although these numbers are somewhat close when looking at the total average difference in the data, the following graphs show more evidence of the fault at a granular level of daily and monthly usage.

	test data					treatment data			
Method	variance forecast error	square root forecast error	bias	average Ytest	error (bias)	bias	average Ytreatment	error (bias)	mean difference (Ymodel - Ytreatment data)
Elastic Net	522245	479	153	3537	4%	355	3478	10%	355
Ridge	522245	479	153	3537	4%	355	3478	10%	355
Lasso	663976	485	105	3537	3%	345	3478	10%	345

Table 4.5 Results from testing the model on test data set and the treatment data

In terms of daily load profiles, figures 4.7 and 4.8 compare the counterfactual and metered data during the test period with counterfactual and metered data during the treatment period. As shown in the figures, the metered data only slightly exceeds the counterfactual data for most hours during the test period, but the metered data significantly exceeds the counterfactual data in the treatment period during business hours. This discrepancy suggests that during business hours, when the building is maintaining a high temperature set point in the winter, the building uses significantly more steam is the exhaust fans are exhausting excessive amounts of air. However, even if the fans are exhausting a lot of air that needs to be reheated at night, the building likely has a temperature set back set point such that the air does not need to be heated to the same degree as during the day. This difference between daytime and nighttime space temperature requirements may help to explain the large peak seen in the counterfactual usage in Figure 4.8.

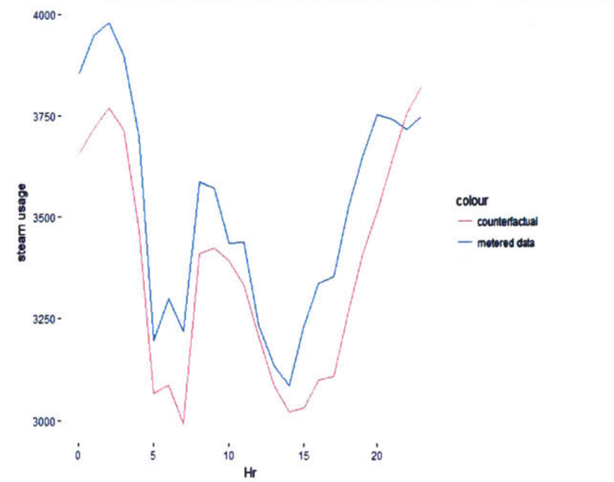


Figure 4.7 Average daily load profile comparison of counterfactual steam data from the model and actual electricity data from the meter over the test period

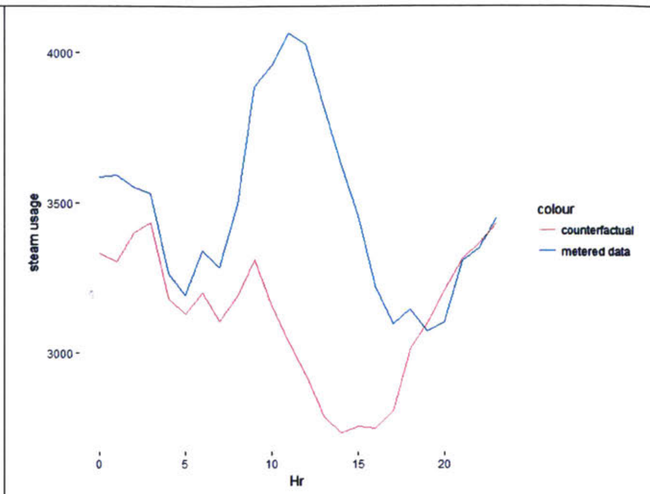


Figure 4.8 Average daily load profile comparison of counterfactual steam data from the model and actual electricity data from the meter over the treatment period

Another useful tool to see the change in energy use as a result of this fault is with a lowess graph. Figure 4.9 shows a lowess graph in which the electricity data is smoothed out with a separate function for the counterfactual test data, the metered test data, the counterfactual treatment data in the counterfactual treatment data. Before the fault occurred, as designated by the black line, the training counterfactual and metered data are very closely aligned. However, after the fault, the metered data is shifted up from the counterfactual data, suggesting of faults did in fact occur in January of 2015.

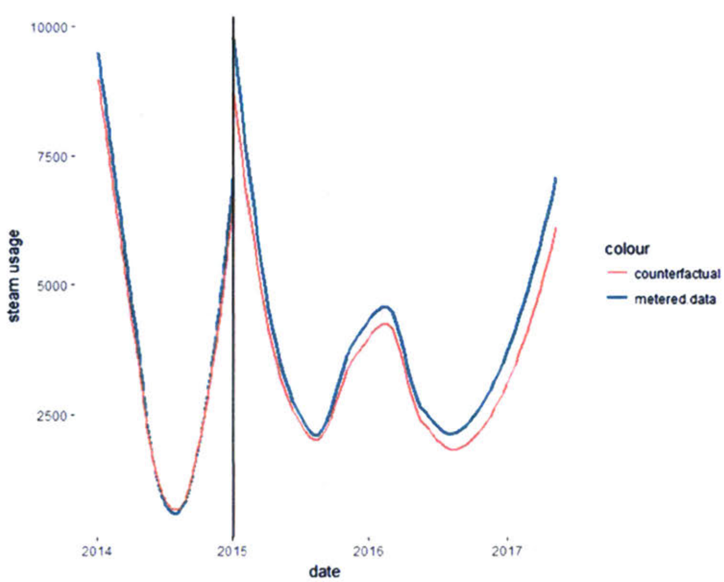


Figure 4.9 Lowess Graph comparing metered and counterfactual data before and after the fault

4.4.1.3 Impact of a Fault on Chilled Water Usage

Next I analyzed the same fault as in section 4.4.1.2 against chilled water usage instead of steam usage. This fault was defined as exhaust fans running at constant speed starting January 1, 2015. The tables below shows the data sets used by time period in addition to the forecast error and bias results. Although there was a large error on the model itself, with an error of 11.3%, there was still a large increase of chilled water usage during the treatment period. As seen in the table, the error on the treatment is more than three times that of the error on the test data. This difference suggests that there was likely an increase an chilled water usage on January 1, however this is a precautionary conclusion to take given the high error on the model itself.

training		test		prediction	
start	end	start	end	start	end
1/1/2010	12/31/2014	1/1/2014	12/31/2014	1/1/2015	12/1/2015

Table 4.6 Breakdown of dates used to build the model, test the model, and analyze the fault

	test data					treatment data			
Method	variance forecast error	square root forecast error	bias	average Ytest	error (bias)	bias	average Ytreatment	error (bias)	mean difference (Ymodel - Ytreatment data)
Elastic Net	5625	50	20	174	11.3%	80	213	37.8%	80
Ridge	6253	53	15	174	8.5%	76	213	35.6%	76
Lasso	5625	50	20	174	11.3%	80	213	37.8%	80

Table 4.7 Results from testing the model on test data set and the treatment data

Further, as seen in the following figures, there was a much larger increase in chilled water usage compared to the counterfactual model during the treatment period than the test period. This difference is highlighted in the summer months when the cooling load is highest. The difference between counterfactual and metered data from the test period.

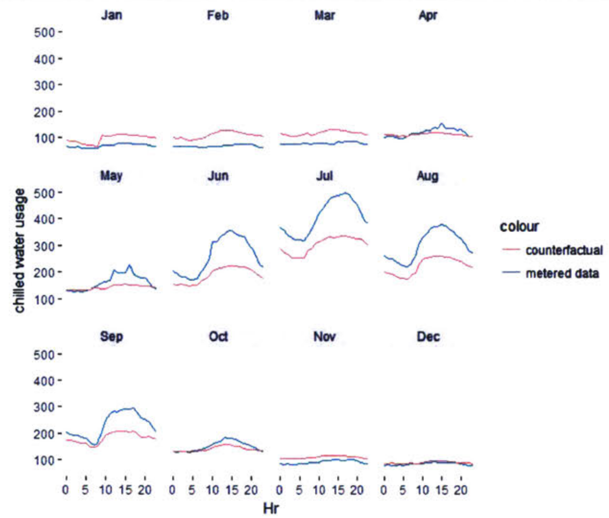


Figure 4.10 Average daily load profile comparison of counterfactual chilled water data from the model and actual chilled water data from the meter over the test period, by month

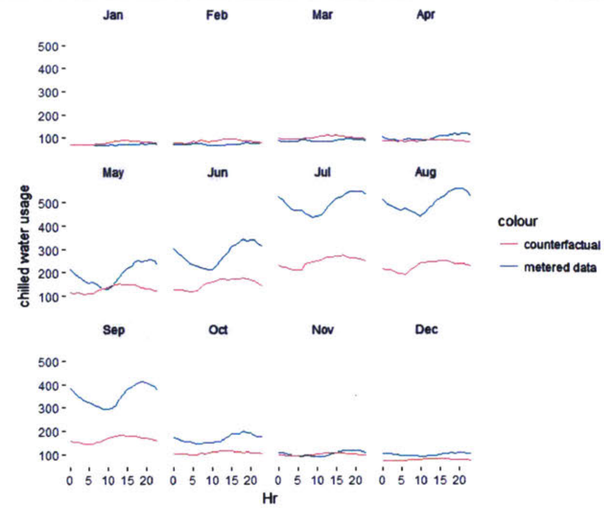


Figure 4.11 Average daily load profile comparison of counterfactual chilled water data from the model and actual chilled water data from the meter over the treatment period, by month

Finally, the following figure shows the shift in energy use with the lowess graph. As mentioned above, the test data showed higher chilled water usage compared to the counterfactual data, but the metered data after the fault was more than three times higher on average than the counterfactual data. The large bias of the model against test data could either signify that the model is a poor predictor of usage, or there was a significant increase in chilled water usage during the test data set that was not uncovered in this study. Although there were no significant faults that were identified during that time period, and facilities did not indicate a change in system usage, there could have been other reasons that usage went up. For example, a fault may have occurred but not been identified by the system, or there was simply a change in occupancy usage of the building.

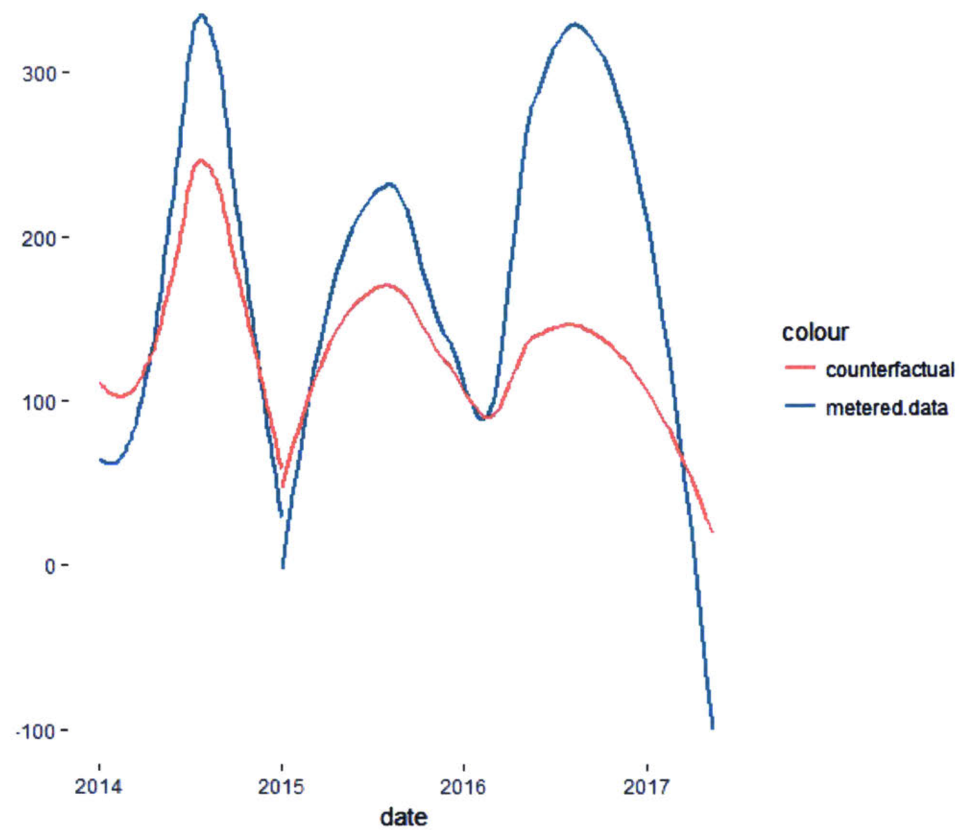


Figure 4.12 Lowess Graph comparing metered and counterfactual data before and after the fault

4.4.2 Building 68 Results

For building 68, I analyzed two faults one for electricity data, and one for steam data. The details and results of these analyses are outlined below.

4.4.2.1 Impact of a Fault on Electricity Usage

For this fault, I used the model I built with electricity training data in Chapter 3 and applied it to the time period in which the fault occurred. Table 4.5 shows the breakdown of training, test, and treatment data. Starting on January 1 of 2015, the fault detection software identified a fault in which the exhaust fans in Building 68 were running at constant speed and no longer modulating to meet demand.

training		test		treatment	
start	end	start	end	start	end
4/1/2011	12/31/2013	1/1/2014	12/31/2014	1/1/2015	11/6/2016

Table 4.8 Breakdown of dates used to build the model, test the model, and analyze the fault

As mentioned in Chapter 3, the Ridge regression model had the least variance forecast error and bias, and was thus used for the analysis on the treatment data. Table 4.6 shows the results from each model, with the Ridge regression model showing an error of only 1.3% for the test data, and an increase of 4.5% after applying the model to the treatment data. Thus, when I applied the model to the treatment data, it showed a significant increase in electricity usage after the fault occurred.

	test data					treatment data			
Method	variance forecast error	square root forecast error	bias	average Ytest	error (bias)	bias	average Ytreatment	error (bias)	mean difference (Ymodel - Ytreatment data)
Elastic Net	3674	48	16	1165	1.3%	51	1149	4.5%	51
Ridge Regression	3674	48	16	1165	1.3%	51	1149	4.5%	51
Lasso Regression	3757	49	22	1165	1.9%	60	1149	5.2%	60

Table 4.9 Results from testing the model on test data set and the treatment data

Figures 4.10 and 4.11 show the average daily load profiles for the test data and treatment data. These results are very similar to the electric daily load profiles in Figures 4.3 and 4.4, in which a very similar fault, fans running at constant speed, occurred in another building. As seen below, there is only a slight change in the average daily electricity usage during the test data period. However, in the treatment data, there is a much larger shift upwards and to the right. Unlike Figure 4.10, the baseload during nonbusiness hours is shifted up significantly, most likely due to the fans running at constant speed and failing to reduce load when the building is unoccupied.

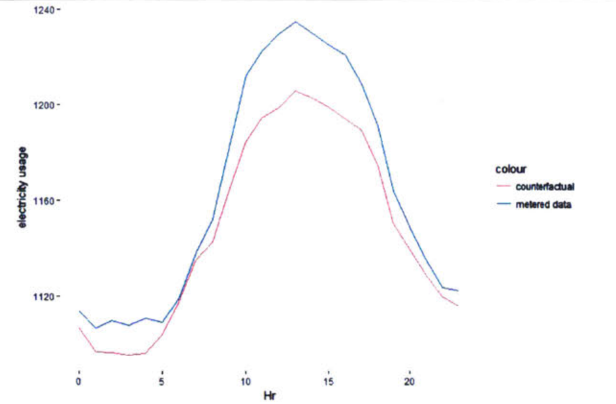


Figure 4.13 Average daily load profile comparison of counterfactual electric data from the model and actual electricity data from the meter over the test period

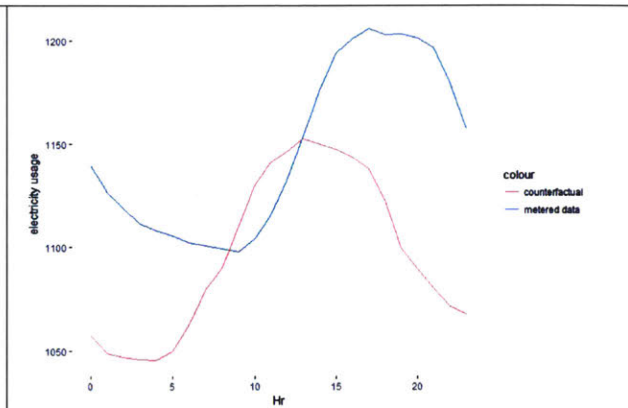


Figure 4.14 Average daily load profile comparison of counterfactual electric data from the model and actual electricity data from the meter over the treatment period

In terms of the change of electricity usage over time, the lowess graph in figure 4.12 shows the change in energy use that occurred as a result of the fault. As seen in the figure, after the fault occurred in January 2015, the metered data was significantly higher than the counterfactual energy data, clearly showing the change in energy use. Unlike the data after the fault, the counterfactual and metered data are very similar to each other prior to the fault occurring.

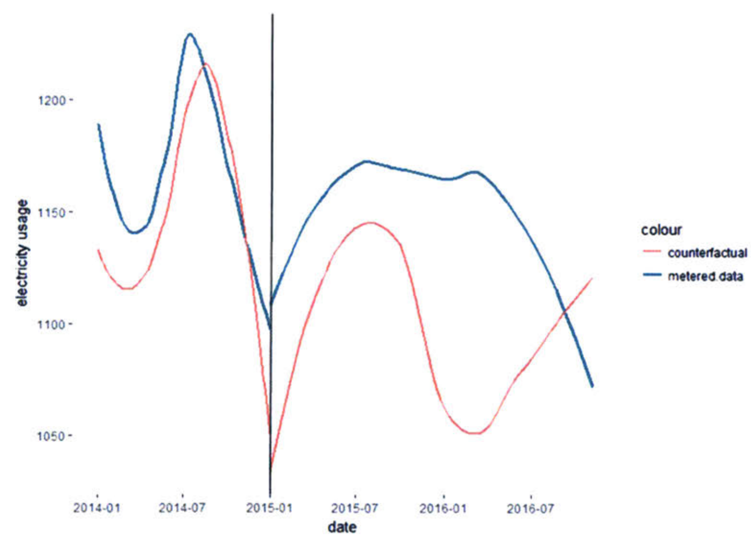


Figure 4.15 Lowess Graph comparing metered and counterfactual data before and after the fault

4.4.2.2 Impact of a Fault on Chilled Water Usage

The next fault I analyzed included two faults occurring at the same time. The first fault found suggested that units were simultaneously heating and cooling. Units will often heat and cold the same time if there is another problem system, such as leaking valves or sensor problems. If a unit is either overheating or over cooling, steam or chilled water will be used to overcompensate in order to achieve a desired temperature set point. This wastes both heating and cooling energy. This fault began occurring in November 2014 for one unit and December 2014 for another unit in the building. The second fault is the same fault identified above in section 4.4.2.1 in which exhaust fans were running at constant speed. Because these faults overlapped at different times, I analyzed only the time period in which each fault was occurring simultaneously, starting January 2015. Table 4.7 shows the breakdown of data sets used for this analysis. Although these faults are expected to also impact steam in a similar manner, a significant amount of data was unfortunately missing for the steam consumption for this building in 2014, not rendering it a potential candidate for analysis.

training		test		prediction	
start	end	start	end	start	end
4/1/2010	12/31/2013	1/1/2014	10/1/2014	1/1/2015	12/31/2015

Table 4.10 Breakdown of dates used to build the model, test the model, and analyze the fault

The results from the model are summarized in Table 4.8. Unfortunately, the model had a high error of 8.5% when compared to the test data, making it difficult to evaluate discrepancies between the model and the treatment data. As seen in the table, while the error for the treatment data is higher than that for the test data, it is only a few percentage points higher at 11.2%. Thus, the difference between the model and actual data is not a good measurement of the energy impact of this fault. However, the daily load profiles below do show evidence that a fault has in fact occurred.

	test data					treatment data			
Method	variance forecast error	square root forecast error	bias	average Ytest	error (bias)	bias	average Ytreatment	error (bias)	mean difference (Ymodel - Ytreatment data)
Elastic Net	8578	61	26	313	8.5%	33	292	11.2%	33
Ridge	10753	67	31	313	10.1%	46	292	15.7%	46
Lasso	8530	61	27	313	8.5%	33	292	11.2%	33

Table 4.11 Results from testing the model on test data set and the treatment data

As seen in figures 4.13 and 4.14, after the fault occurred, the average daily chilled water usage increased substantially during nonbusiness hours as compared with the baseline counterfactual usage. While the chilled water usage closely matches that of the counterfactual modeled data during business hours, figure 4.14 shows the sharp increase during nonbusiness hours. We do not see the same sharp increase in the test data set, suggesting that the increase in chilled water usage during nonbusiness hours is unique to the treatment data after the fault began occurring. This large increase at night could be due to a leaking cooling valve. While the cooling valve might be fully open during business hours, it is likely that the building set back temperatures reduce the need for cooling at night, as seen by the low cooling usage during nonbusiness hours in the metered data over the test period, the counterfactual data over the test period, and the counterfactual data over the treatment period. The significant increase in chilled water usage at night in figure 4.14 suggests that a cooling valve was leaking by or another fault was occurring that prevented the building from reducing cooling usage at night. This discrepancy during nonbusiness hours can be seen more clearly when separating the daily load profiles by month.

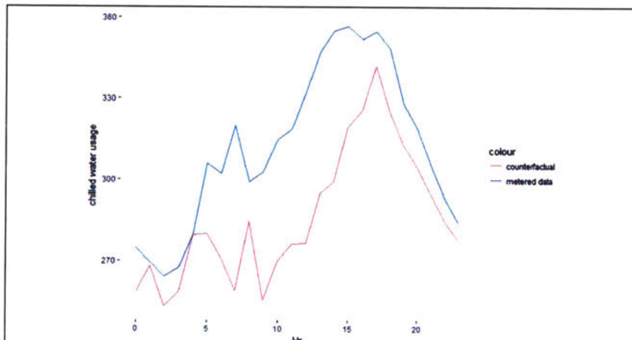


Figure 4.16 Average daily load profile comparison of counterfactual chilled water data from the model and actual chilled water data from the meter over the test period

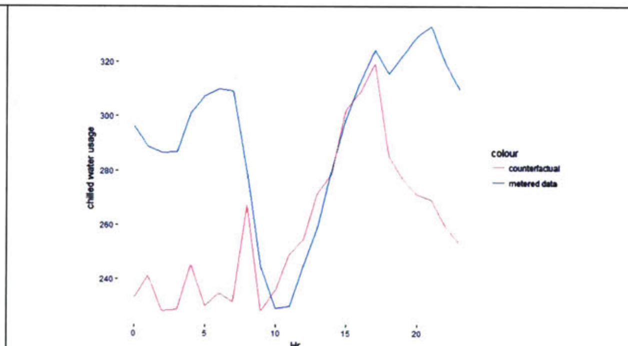


Figure 4.17 Average daily load profile comparison of counterfactual chilled water data from the model and actual chilled water data from the meter over the treatment period

Figures 4.15 and 4.16 show the daily load profiles by month, which clearly separates the daily loads by the season. As seen in figure 4.16, the daytime cooling usage closely matches that of the counterfactual estimates, while the nighttime consumption typically exceeds that of the daytime consumption. Further, the actual metered consumption is flatter than the counterfactual data or the test data, suggesting that a fault was occurring that prevented modulating cooling usage based on demand. Although figure 4.15 also suggests something occurred during this time period during the summer to increase usage, the particular daily load profile discrepancy mentioned above is unique to the data after the fault occurred.

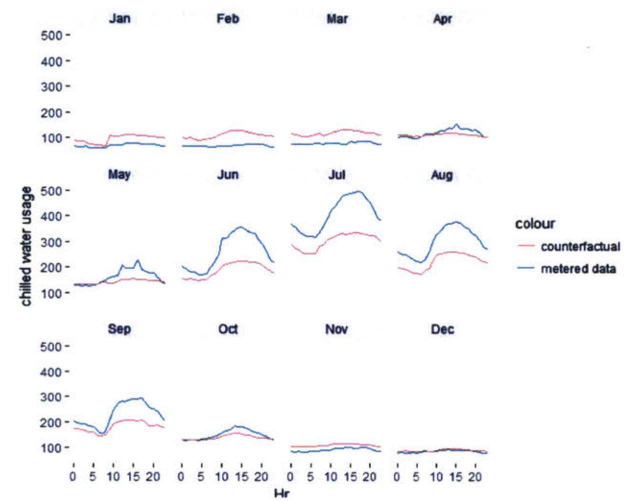


Figure 4.18 Average daily load profile comparison of counterfactual chilled water data from the model and actual chilled water data from the meter over the test period by month

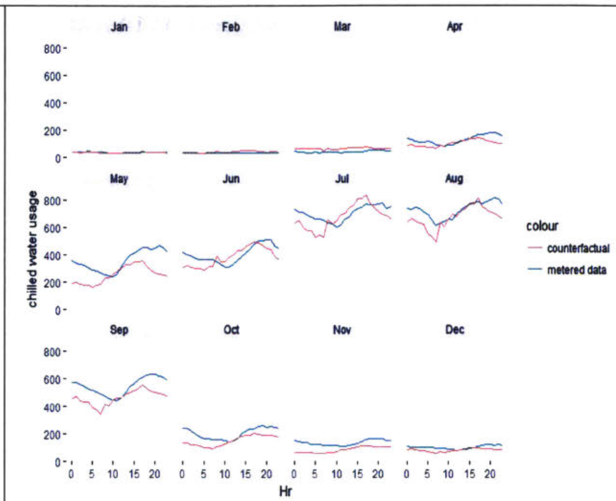


Figure 4.19 Average daily load profile comparison of counterfactual chilled water data from the model and actual chilled water data from the meter over the treatment period by month

Finally, figure 4.16 shows the metered and counterfactual data before and after the fault. The gap in between is data not analyzed because only one of two faults was occurring. As seen in the graphs, both the test data and treatment data exceeded the counterfactual data for most of the data points, making it difficult to quantify the increase in chilled water consumption associated with the fault.

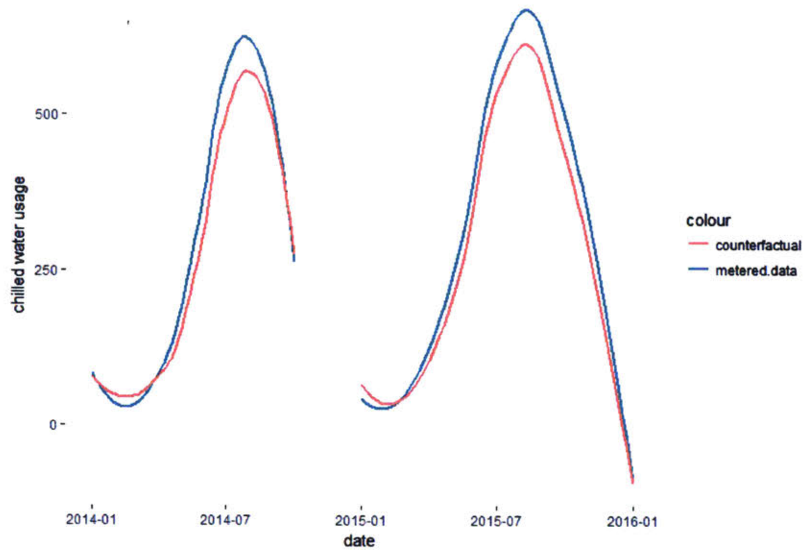


Figure 4.20 Lowess graph of test data and treatment data, with a gap in the middle of data not analyzed

4.4.3 Building 76 Results

For building 76, I analyzed one fault on electricity usage, as outlined in the following section.

4.4.3.1 Impact of a Fault on Electricity Usage

In this section I analyze a fault in which all three fans in three air handling units were running at a constant full speed. This could have occurred from a manual override, sensor issues, damper issues, or other problems in the system. Regardless of what cause the faults, fans stuck running at full speed when they could have been modulating can waste significant amounts of energy.

training		test		prediction	
start	end	start	end	start	end
4/1/2011	12/31/2013	1/1/2014	9/1/2014	1/1/2015	9/1/2015

Table 4.12 Breakdown of dates used to build the model, test the model, and analyze the fault

As seen in table 4.10, the model performed really well against the test data with an error of only 0.2%. After the fault occurred, the model showed a 4% increase in electricity usage. This increase suggests a quantifiable increase in energy consumption as a result of the fault.

	test data					treatment data			
Method	variance forecast error	square root forecast error	bias	average Ytest	error (bias)	bias	average Ytreatment	error (bias)	mean difference (Ymodel - Ytreatment data)
Elastic Net	1123	25	2.03495	830.120558	0.2%	30	844	4%	30
Ridge Regression	1271	28	2.77132	830.120558	0.3%	31	844	4%	31
Lasso Regression	1125	25	2.02017	830.120558	0.2%	30	844	4%	30

Table 4.13 Results from testing the model on the test data set and the treatment data

As seen in the figures below, the daily load profile of counterfactual and metered data is almost identical over the test data. However, over the treatment data, the metered data is shifted right and has a higher base load. These results are very similar to those found in Figures 4.4 and 4.11, in which both show the treatment data after a fault occurred in which a fan was running at full speed. Because the fans are likely running full speed at night, the base load is higher than the counterfactual usage during nonbusiness hours.

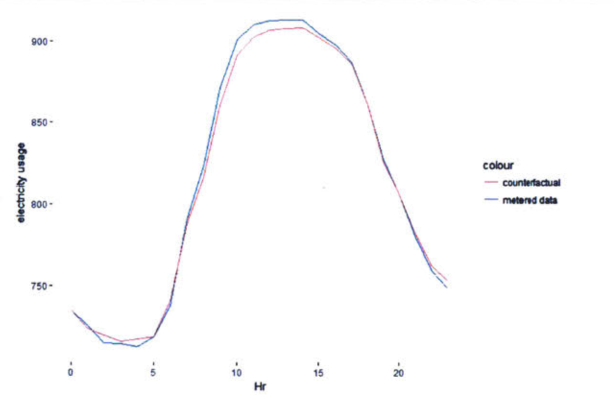


Figure 4.21 Average daily load profile comparison of counterfactual electricity data from the model and actual electricity data from the meter over the test period

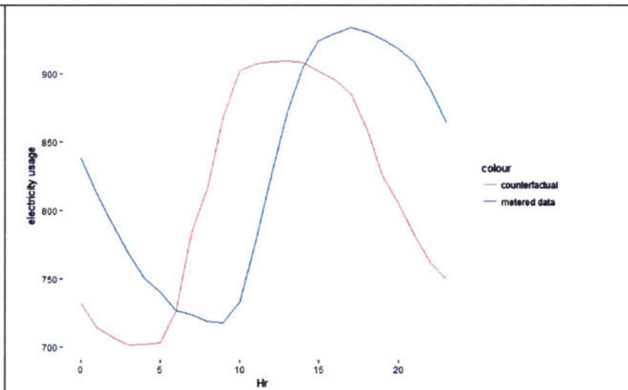


Figure 4.22 Average daily load profile comparison of counterfactual electricity data from the model and actual electricity data from the meter over the treatment period

Further, figure 4.18 shows a lowess graph of the counterfactual and metered data over time. Although there is a gap in the data due to some data not well-equipped to be analyzed, there is a clear shift in the metered data compared to the counterfactual data after the fault occurred in January 2015, as compared to the test data.

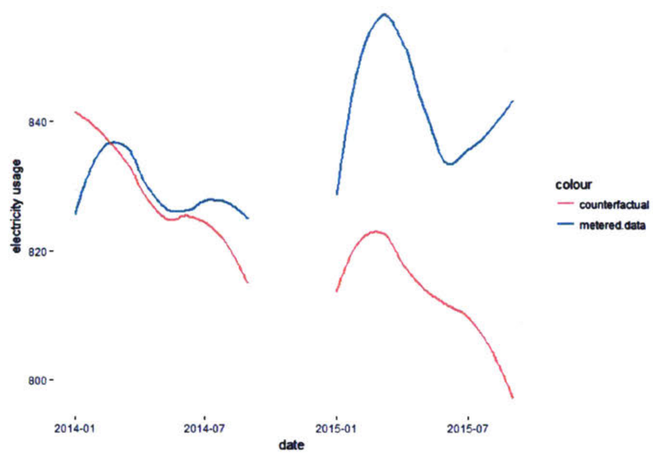


Figure 4.23 Lowess graph of test data and treatment data, with a gap in the middle of data not analyzed

4.4.4 Building 18 Results

Finally, for building 18, I analyzed a fault on electricity usage.

4.4.4.1 Impact of a Fault on Electricity Usage

The following tables provide the results from a fault that occurred on January 1 of 2015 when fans began running a constant speed, similar to many of the faults found above. As seen in the results table, while the error went up from the test data to the treatment data, these numbers are still very close, with 3.2% and 4.5%. Although these numbers are too close to prove a fault occurred with high confidence, the daily load profiles do suggest that a fault in fact occurred.

	test data					treatment data			
Method	variance forecast error	square root forecast error	bias	average Ytest	error (bias)	bias	average Ytreatment	error (bias)	mean difference (Ymodel - Ytreatment data)
elnet	1993	34	24	766	3.2%	35	765	4.5%	35
ridge	1993	34	24	766	3.2%	35	765	4.5%	35
lasso	2113	35	26	766	3.4%	36	765	4.7%	36

Table 4.14 Results from testing the model on test data set and the treatment data

As seen in the daily load profiles below, similar to all previous faults discussed above in which a fan or fans were running a constant speed, daily load profile electricity usage shifts up and to right, with an increase in the baseload during nonbusiness hours. This is consistent with what we would expect to see with a fan running at full speed, because the largest difference should occur at night when fans would be expected to shift to a lower speed when occupancy is reduced.

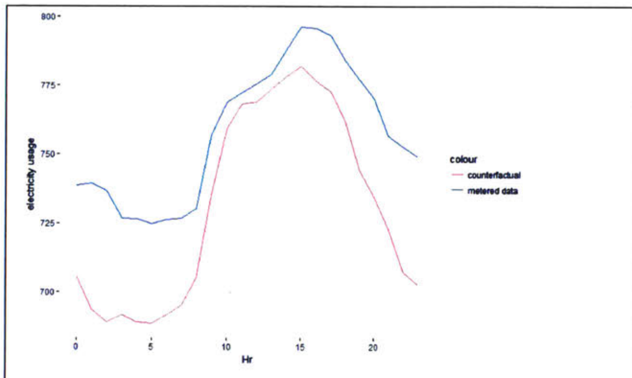


Figure 4.24 Average daily load profile comparison of counterfactual electricity data from the model and actual electricity data from the meter over the test period

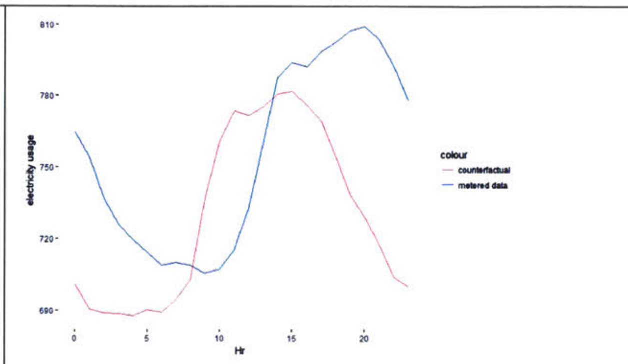


Figure 4.25 Average daily load profile comparison of counterfactual electricity data from the model and actual electricity data from the meter over the treatment period

Finally, the change in energy use can be seen in the lowess graphs below. Although it is unclear what caused the electricity usage to increase only in winter months during the test data period, Figure 4.22 shows a full shift upward in electricity usage from counterfactual data to metered data.

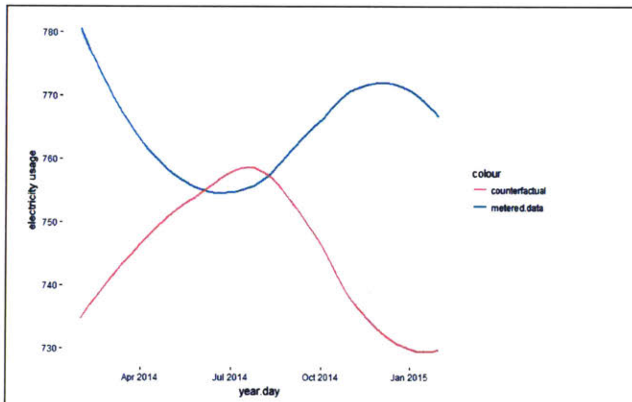


Figure 4.26 Lowess graph comparison of counterfactual electricity data from the model and actual electricity data from the meter over the test period

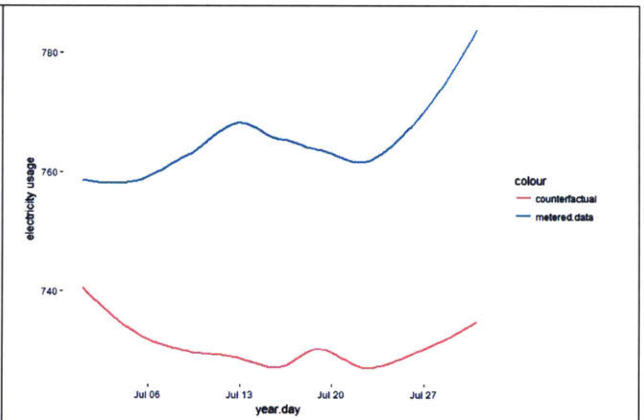


Figure 4.27 Lowess graph of counterfactual electricity data from the model and actual electricity data from the meter over the treatment period

4.5 Conclusion

The results from this chapter highlight several key findings. Sections 4.5.1 to 4.5.4 discuss some of the related to this chapter. These results show that it is likely possible to identify faults in energy data;

highlight faults that are easier to identify; show the level of accuracy of the estimates produced by the fault detection software; and show the need for further research.

4.5.1 It is likely possible to identify some faults in the energy data

The results from this section show that it is likely possible to identify faults in energy data. While some are more convincing than others, each of these faults reflected a corresponding energy increase in the energy data. Several these models had errors of 5% or less, with the highest error at 11%. The buildings studied are very complex and have very variable energy usage based on variable occupancy loads and complex HVAC equipment. Because of this, these error numbers likely show highly accurate models. However, there were several instances within the energy data were energy dropped or increased where we were unable to identify the cause. These unexplained fluctuations made it difficult to have a high confidence level of the modeled predictions compared to raw energy data, because in theory another unexplained fluctuation may have been occurring instead. Further, due to the large number of faults reported in each building (hundreds to thousands), there was some level of bias introduced in terms of selecting which faults to analyze. Although I used the criteria outlined in this chapter to select faults, there was some level of discretion in terms of determining which faults are the most problematic and whether or not another coexisting fault should have an impact on the model. Thus, this methodology will prove to be more accurate when more faults are resolved and buildings will have fewer faults that can be isolated more clearly and with less bias.

4.5.2 Faults related to fan speed were easiest to identify

I found faults related to fan speed to be easier to identify in the energy data then other types of faults for several reasons. Fan speed changes are more abrupt compared to faults that change over time or based on whether. For example, a leaking valve might slowly start to fail over time, and it may prove better or worse depending on the weather. However, fan speed changes show a concrete change in the data that can be more easily isolated. Additionally, fans running a constant speed is also a typical fault to find in a building if a building operator or technician overrides the fans. Further, fans have a very characteristic load shape and make up a large percentage electricity data, making it easier to identify this fault with the energy data alone. Finally, there were simply more instances of fans shifting from modulating flow to running constant speed. A lot of the faults associated with simultaneously heating and cooling or leaking valves have been occurring since the software came online, whereas there were several instances of fans shifting from fluctuating to constant speed due to the nature of the fault.

4.5.3 KGS estimates were close to actual energy savings

Overall, the KGS estimates for energy savings associated with fixing each fault were close to the changes found in the energy data, with some faults having been over predicted and some faults having been under predicted. Because this analysis only examined eight faults out of thousands reported by the system, it is by no means a full evaluation of the system as a whole at estimating energy savings. However, I was able to analyze some of the higher energy consuming faults that were good candidates for this type of analysis, and most of them are off by 10 to 30%. Further, while some of the faults included details on the energy savings calculations with errors; others had very simple dollar savings; while still others had no savings reported at all. The software can be improved by simultaneously running models of the energy data alongside the fault detection data to better estimate energy impact and help facilities groups prioritize faults based on energy savings.

4.5.4 Supplemental analysis is necessary to further test the hypothesis of this thesis

Finally, it is important to note that further analysis is necessary to test fault detection systems and determine their effectiveness at identifying faults and evaluating energy savings. There were many overlapping faults in the data, rendering it difficult to determine which faults could be analyzed in which had too much noise around them. Further research is needed in systems with fewer faults or faults can be better isolated in the energy data. Studying more faults would also help to answer questions on which faults yield the highest accuracy in energy estimates; which faults are easiest to identify in the data; and what the characteristic load shape shifts are like on a fault by fault basis.

5 Conclusion & Discussion

5.1 Introduction

As discussed in Chapters 1 and 2, there is a strong need for this type of research and there are several applications to apply this research to. Because of the large gap identified in energy efficiency, there is a strong need for improved methodologies to evaluate counterfactual energy data. This thesis compares several different novel machine learning techniques which were each able to show significantly improved counterfactual predictions. While there is a lot of literature evaluating different measured machine learning techniques and statistical methods for making predictions, there is very limited research in applying these methods to energy usage. Further, there is very little research in this domain on commercial and industrial facilities, and instead most literature evaluates residential facilities. Additionally, these studies focus mostly on electricity data, with little to no research done on high-frequency steam and chilled water metered energy usage. Finally, there is very limited research evaluating fault detection diagnostics software in terms of energy savings, and I have not found any research combining these machine learning techniques for evaluating energy usage identified by an existing fault detection and diagnostics system in a real building. This thesis addresses each of these gaps and provides an evaluation of a real world example of a fault detection system and a novel methodology to evaluate the energy savings associated with the system. By filling these gaps, the results from this thesis, in addition to further similar research needed in this domain, will help to both inform policy and improve existing software.

5.2 Informing Policy

The methodology outlined in this thesis can help inform policy on how and when to specify FDD requirements by providing an evaluation tool for energy savings. The table below summarizes these opportunities.

Summary of opportunities for improving policy

- Inform building code developers on the most appropriate applications for FDD requirements
- Inform building code developers on a methodology to evaluate the energy impact of FDD software

Currently, the majority of building code requirements for commercial buildings are focused on the as–built construction requirements, with limited focus on software or other ongoing requirements for the post–occupancy period. Occupancy permits are not issued until building code requirements are met, and after these permits are issued, the city does not have much jurisdiction on any ongoing maintenance requirements. However, fault detection and diagnostics provide an opportunity for ongoing commissioning to ensure buildings continue to operate properly post occupancy. There has been discussion amongst the Association of Heating Refrigeration and Air-Conditioning (ASHRAE) on whether fault detection diagnostics should be included in any code requirements, and Title 24 in California has already included code to require fault detection in specific applications for air handling units. According to the California Energy Commission,

> *"Title 24, Part 6 Section 120.2(i) requires that economizer fault detection and diagnostic functions (FDD) be installed on air-cooled unitary air conditioning systems over 4.5 tons cooling capacity, with the ability to detect the faults specified in Section 120.2(i). Each air conditioning system manufacturer, controls supplier, or FDD supplier wishing to certify that their FDD analytics conform to the FDD requirements of Title 24, Part 6 may do so in a written declaration."*[50]

ASHRAE writes the building code that is used for many jurisdictions in the United States, and Title 24 writes the building code used in all jurisdictions in California. We are starting to see a trend in interest in requiring some level of fault detection in building code requirements for certain applications, and this trend is likely to increase as the software continues to spread to the market and improve in

[50] California Energy Commission. 2016. "2016 Manufacturer Certification for Equipment, Products and Devices." <http://www.energy.ca.gov/title24/equipment_cert/fdd/>

performance. With this increase in interest, it is imperative that we develop improved methodologies for evaluating these systems and ensuring proper functionality. There is currently a very limited research on evaluating these systems in existing buildings and outside of the laboratory, making it very difficult to determine exactly how to specify these systems. Further, limited research makes it difficult to determine which types of systems are most effective and should be highlighted. Improved research also allows us to trade off the benefits and costs of these systems in order to specify the most effective applications for this type of software. This thesis proves that the methodology outlined in Chapters 3 and 4 are likely to be effective at helping to start to answering some of these questions by testing many more buildings, fault detection systems, and faults.

5.3 Improving Existing Software

In addition to informing policy, the results from this thesis could help to improve existing fault detection and diagnostics software on the market. The table below highlights the opportunities this methodology provides for improving the software.

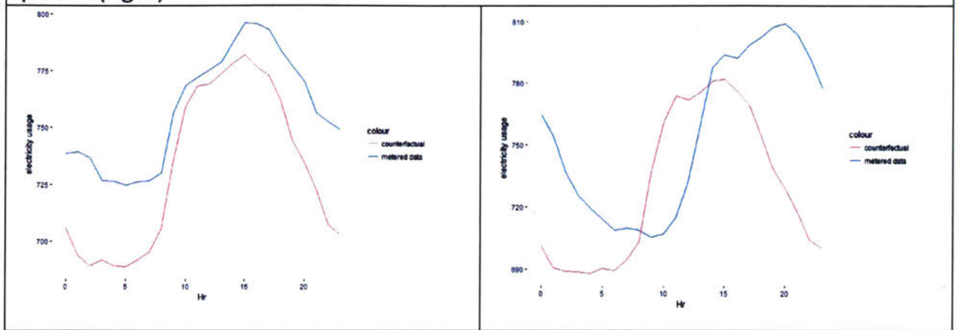
Opportunities summary for improving FDD software with energy modeling methodology
• **FDD manufacturers** can add a component to existing FDD software to continually monitor energy simultaneously • **Building owners** can use this methodology to prioritize faults • **FDD manufacturers** can offer a standalone software to identify faults only with energy data

First, a fault detection and diagnostics system could continually run a second software that analyzes energy data similar to the methods outlined in this thesis. The software could identify shifts in energy usage and assign energy savings associated with the faults that are identified by the fault detection diagnostics software. Because existing fault detection diagnostics software systems tend to identify many false positives in the data, this approach could help reduce the number of false positives by testing the energy data to see if there was in fact an increase in usage when the fault began occurring. This methodology has the potential to reduce the number of false positives, which in turn would save costs for facility owners. A reduction in false positives would reduce the wasted cost of sending maintenance professionals to diagnose false positives. A reduction in false positives would also reduce wasted cost by immediately addressing only those faults that are accurate, and preventing the faults from continuing to waste energy.

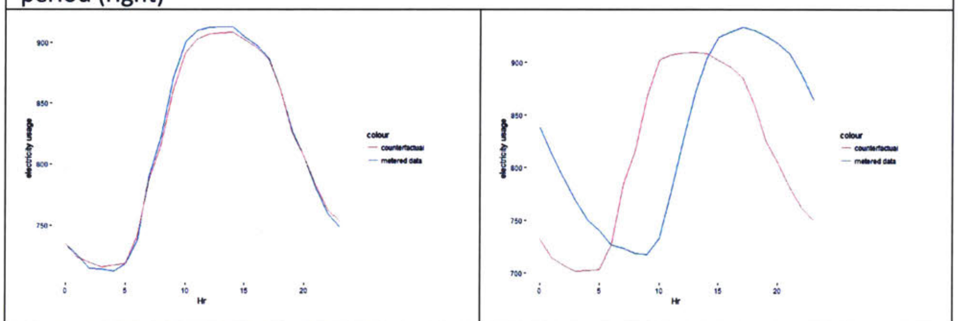
Second, a methodology such as the approach outlined in this thesis could help facilities to better prioritize which faults to analyze and address. If a facilities staff person is reviewing thousands of faults, it can be time-consuming and cumbersome to prioritize and diagnose faults. This methodology can help to more accurately assign energy savings to each fault so that facilities personnel can better target the highest energy users and the most cost-effective faults.

Finally, if a building does not have a sophisticated building automation system, or a building owner is not confident in the accuracy of the existing data from the building automation system (such as failed sensors, poorly placed sensors, or points mapped incorrectly), a methodology similar to that outlined in this thesis could be used in the absence of a fault detection and diagnostics system. Rather, a building owner could simply use a software with these algorithms to quickly identify a statistically significant shift in energy usage. Surprisingly, I found that the faults I analyzed in this thesis to not only show an average shift upward of energy usage, but there were consistent shifts in daily load profiles after certain faults occurred. For example, Figure 5.1 shows that there was a very characteristic shift in the daily load profile for all four faults in which fans began running a constant speed in four different buildings. In all cases below, the counterfactual and metered data were very similar during the test period, but the daily load profile of the metered data was shifted up and to the right with a high baseload during the treatment period. This is likely to have occurred because the fans were running a constant speed at night, resulting in a high baseload. Because the fans were running a constant speed at night, it likely did not take as long for the building to heat up or cool down in the first few hours of the day, causing the shift right of the data. These hypotheses are supported by the graphs shown below.

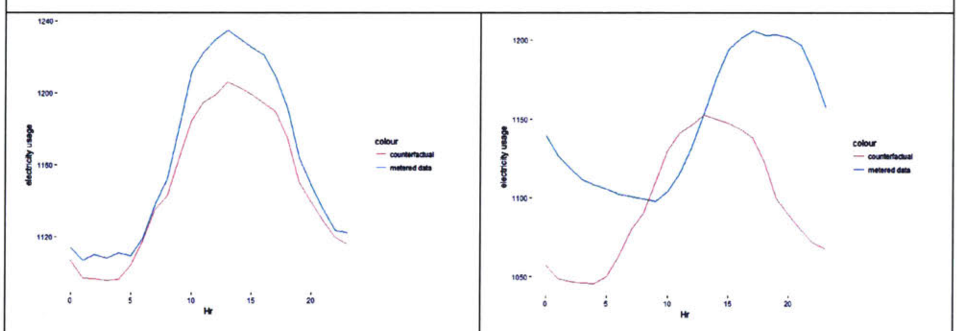
Building 18 metered vs counterfactual data over test period (left) and treatment period (right)



Building 76 metered vs counterfactual data over test period (left) and treatment period (right)



Building 68 metered vs counterfactual data over test period (left) and treatment period (right)



Building 13 metered vs counterfactual data over test period (left) and treatment period (right)

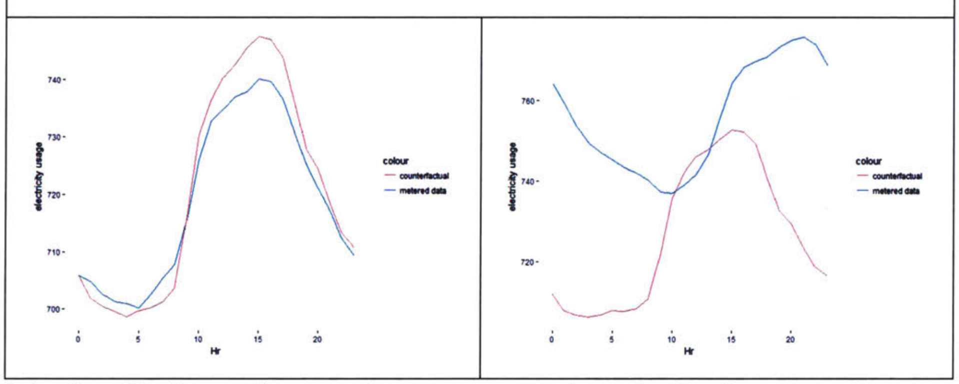


Figure 5.1 Daily load profile changes from test data to treatment data after fans began running a constant speed in four different buildings

Thus, a fault detection diagnostics system that relies only on energy data could not only look at the energy data shifts over time, but it can look at average daily load profile shifts and average daily load profile shifts by day of the week. These multiple sources of data could help to provide additional information in order to identify a likely fault in a building in the absence of a BAS.

Further, the results from Figure 5 shows that this type of analysis can help facilities groups better understand the economic consequences of certain faults if there exists time of use pricing. For example, the graphs above show that running the fans at night increases the overall electricity costs, but it likely reduces electricity costs for a certain period of the day. For example, each graph during the treatment period shows that the counterfactual data is higher than the metered data for a small period of time in the morning, likely due to the fans having a "head start" on conditioning the building if they were running at night. If there is time of use pricing, and the price during that time period is such that the increase in electricity costs exceeds the cost savings from turning the fans off at night, the fault may actually be saving overall electricity costs. Thus, this finding is significant in applications in which there is time of use pricing, such as MIT. Thus, future studies on the energy savings of fault detection and diagnostics should take into account time of use pricing when making estimates and pay particular attention to these shifts in daily load profiles.

5.4 Opportunities for Further Research

As mentioned previously, this thesis proves that there is a strong need for additional research in this domain. While this research shows that it is likely possible to identify faults in energy data, policy proposals and the industry could greatly improve from significantly more studies done on more buildings, more systems, and on more faults. This research could help provide a stronger case for how to include fault detection and diagnostics in existing policy and how to better improve these systems. Further research would also help to identify which modeling technique is best under which circumstances. Although I found that different modeling techniques were more accurate under certain applications, it could be possible to determine which techniques are best suited for exactly which applications.

5.5 Conclusion

The findings in this thesis provide useful insight into applying machine learning algorithms to energy prediction and in estimating the energy impact of various faults. The findings from this study will be provided to MIT Facilities and the MIT Office of Sustainability. The methodology may be continued across additional buildings and across additional faults to help to improve their system. Additional future work outside of the MIT with different systems and different faults could help to make a significant improvement in policy and the industry as a whole.