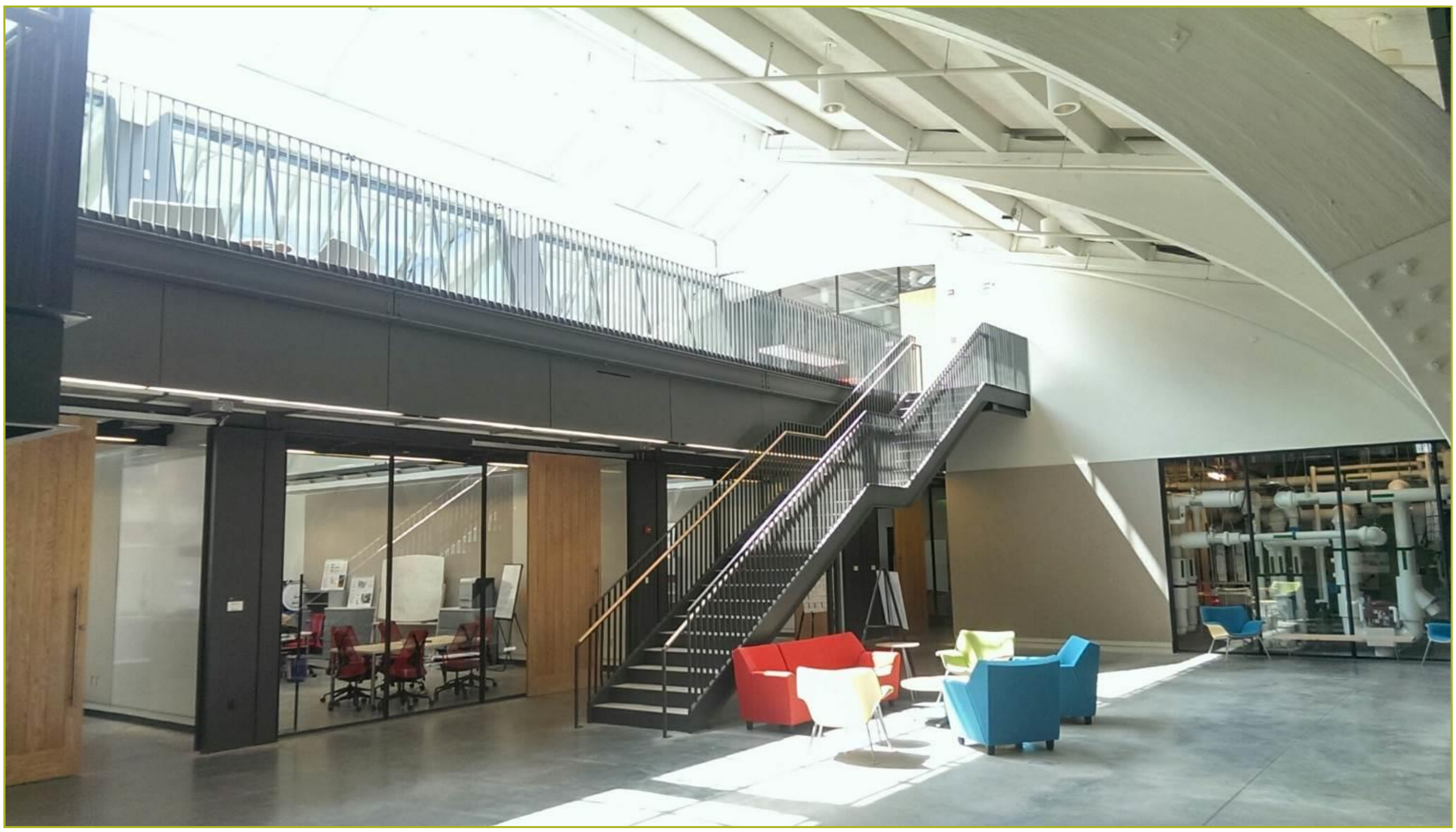
CONSORTIUM for BUILDING ENERGY INNOVATION

REPORT

Title:Targeting Rebate Program Customers with Benchmarking Data Analytics Methods

Report Date: April 28, 2016

Report Authors: Erica Cochran, Flore Marion, Scott Wagner, Ben Cohen, Leslie Billhymer



Funded by U.S. DOE

REPORT

Report Abstract

CBEI supported the Philadelphia region as it implemented its ordinance in 2012, acting as a neutral party to convene stakeholder meetings and providing technical expertise on the value of understanding how building energy performance compares regionally. CBEI used this experience to help develop solutions for regions to make use of benchmarking data. CBEI also collaborated with utility incentive program administrators to develop approaches for using benchmarking data to more specifically target buildings for incentives, making it easier for program administrators to reach the best candidates and therefore reduce the overhead cost of an incentive program.

Utilities expressed the need to utilize existing and new resources to help them strategically focus their rebate program target areas and enrich EE program initiatives. Therefore developed new analysis for DSM programs (energy efficiency and demand response) based on benchmarking data to show the usability of benchmarking data for utilities. This report compile the different materials used to present this methods to various stakeholder and demonstration of the applicability of benchmarking data to improve current and develop new rebates programs in a more targeted way.

Contact Information for lead researcher

Name: Erica Cochran, PhD;
Institution: Carnegie Mellon University
Email address: EricaC@andrew.cmu.edu
Phone number: 4122687145

Contributors

Flore Marion , Co-PI, Carnegie Mellon university

Scott Wagner, Pennsylvania State University

Ben Cohen, Pennsylvania State University

Leslie Billhymer, University of Pennsylvania

Valerie Patrick, Fulcrum Connection

Timothy Spencer, formerly of Carnegie Mellon University

Zoe Kaufman, formerly of Carnegie Mellon University

Juan Castellanos, Carnegie Mellon University

Prachi Renee Sharma, Carnegie Mellon University

Noreen Saeed, Carnegie Mellon University

Nilesh Bansal, Carnegie Mellon University

Project Milestones and Deliverables

Final BP5 Deliverables:

- D6.4.1 - Documentation of approach and recommendations for improving utility DSM programs through use of benchmarking
- D6.4.2 - Quantification of the value of benchmarking to DSM programs.
- D6.4.3 - Intermediate and advanced benchmarking data analytics guides describing methodologies for identifying buildings for potential utility rebates.
- "D6.4.4 - Package means and methods findings for new users and recorded webinar to share with and new and existing project partners

Milestones and Go/No-Gos:

M6.4.a	Engage one Utility partner to collaborate.	Commitment from one Utility.	✓
GN6.4.1	Detailed data plan	1) Description of data required, 2) List of data sets currently available, 3) Plan for data acquisition with roles and responsibilities of CBEI and utility program administrator	2 ✓
M6.4.b	Work with committed Utility partner/s to identify their needs and goals.	Report of needs and goals from utility partners	✓
M6.4.c	Preliminary analysis results quantifying value proposition for Utilities to support benchmarking programs.	Report on preliminary data analysis delivered to CBEI, DOE, and Utility. Report describes (A) methods and algorithms to utilize benchmarking data to improve DSM program outcomes and (B) quantitative analysis of value to utilities of building benchmarking based on ways in which utility has used benchmarking program.	✓
GN6.4.2	Preliminary data analysis	Analysis utilizing a sample data set to show that the proposed algorithms identify candidate buildings and provide recommended retrofits	7 ✓
M6.4.d	Final intermediate and advanced benchmarking data analytics guides	Report delivered to CBEI and DOE	✓
M6.4.d	Completion of final analysis. Package means and methods findings for new users and host webinar to share with new and existing project partners.	Delivery of final report describing (A) methods and algorithms to utilize benchmarking data to improve DSM program outcomes and (B) quantitative analysis of value to utilities of building benchmarking. Present findings in at least 2 public forums and document holding at least 4 additional knowledge transfer sessions that share these approaches with at least 5 additional utilities, at least 5 additional program implementers, and at least 5 municipalities.Submit packages and outcome report on webinar to CBEI.	✓

Targeting Rebate Program Customers with Benchmarking Data Analytics Methods

Table of Contents

LIST OF FIGURES

LIST OF TABLES

1 Introduction

The growing field of building energy benchmarking has opened the door to understand, on a city scale, how buildings use energy. Through an interdisciplinary collaboration between multiple Universities, multiple disciplines (Architecture, Engineering, Statistics, and Computer Science), and Industry, the CBEI project 6.4 team investigates new data mining and analysis techniques to identify statistically significant correlations between buildings attributes and energy consumptions and ENERGYSTAR scores of existing buildings. Such techniques are designed to lead towards refining design guidelines and recommendations for energy efficient retrofits and new construction. Additionally, the findings are to be used to identify facility management recommendations that can reduce a building's energy consumption.

This report presents the results of a sample size analysis of 115 buildings as a preliminary step toward large scale analysis of benchmarking. For each of the buildings within the dataset, data was gathered via building visits and internet mapping sources on 44 different physical building attributes, constituting of a total over 5,500 data points. In order to eliminate variability due to use type, this study was limited to office buildings in the Mid-Atlantic region, specifically buildings in greater Philadelphia, PA, and Washington, D.C. Variability in energy use due to occupancy was controlled for to the greatest degree possible, but was impossible to control for in all analyses. Variability for weather was controlled for utilizing weather data from the closest major airport and normalizing for the heating and cooling degree-days of each date.

2 Project Outline

The research outlined in this report analyzes energy use data at various intervals to understand the levels of correlation between certain physical attributes and energy use given available benchmarking data and data sets from existing DOE tools and other open-source tools, as well as potential future benchmarking portfolios.

The result is an analysis of how building attributes contribute to energy use in the region studied, which can be used to target retrofits of existing buildings. Energy use data has become available at several different levels: "annual", "monthly", and "interval" energy use. While benchmarking data sets utilize annual data, monthly utility data is available through utility bills, and interval (sub-hourly) data has recently become available through smart metering. All three levels of energy use data will be addressed and used in this research to demonstrate the inferences that can be made from the varying data types and sources.

The methods set out in the report constitute a replicable framework and can be applied to any region where benchmarking data is available. When analyses are performed, the data source and means of data acquisition are noted to later provide a report describing the types of conclusions made available by utilizing different data sources and by analyzing various intervals of energy data.

The diagram below describes the organizational roles in obtaining, analyzing, and applying building information for the purpose of augmenting utility rebate programs.

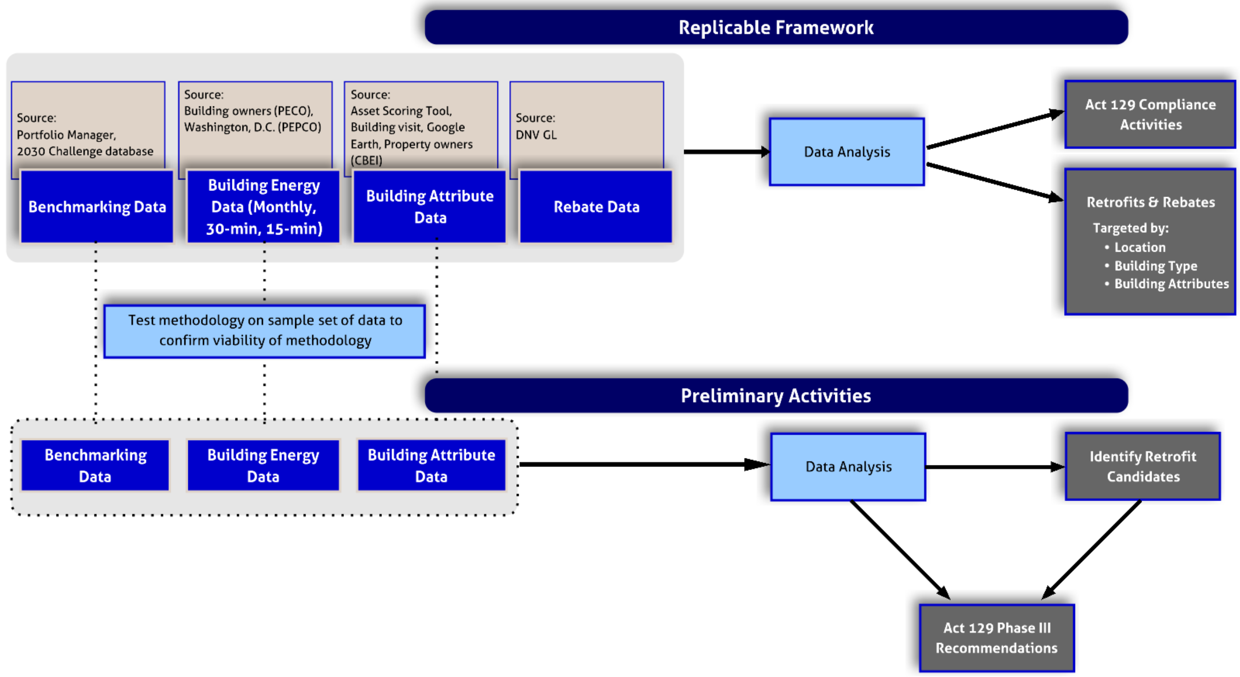


Figure 1 - Replicable framework for rebate analysis

The replicable framework laid out at the top shows that four types of data will be combined for analysis and eventually applied to helping utilities comply with Act 129 and informing utilities as to the most worthwhile retrofits and rebates for their particular building stock through building-specific targeted retrofits. Beneath this, we show that, in order to test the methodology on a smaller scale to verify its viability, a sample set of benchmarking data, building attribute data, and building energy data (annual, monthly, and interval) is collected and analyzed. This preliminary analysis will not only inform the methodology, but will also identify retrofit candidates and lend insight into shaping goals for the structuring of Act 129 Phase III and similar legislation nationwide.

3 Methods of Analysis of Energy data

3.1 Building Sample Size

The team secured over 144 buildings with available annual data and 58 with interval data. Post data cleaning, and removing faulty data sets, left us with 76 and 51 buildings with annual and interval data respectively (see Figure 2). The large database was the result of a successful outreach plan and a strong partnership with the City of Philadelphia which granted Portfolio Manager access for over 900 buildings and the partnership with BOMA Philadelphia. The team selected a large sample of office building to apply their methods.

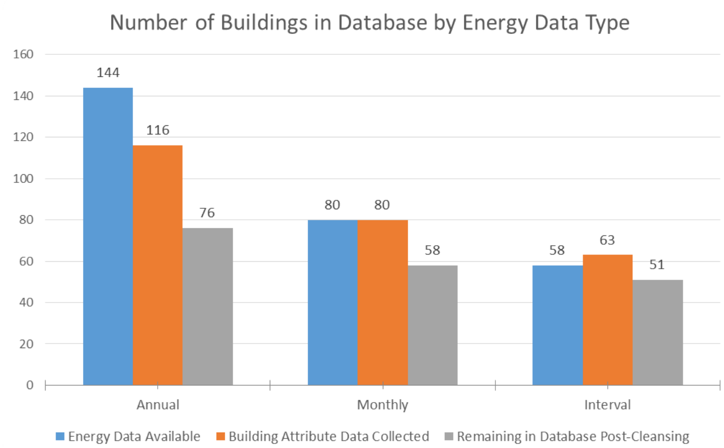


Figure 2 - Building energy and attribute information applicable to this project's subset of buildings

The "Geographic Distribution of Building Study Sites (2015)" map also shows the three neighborhood concentrations of the sample data set; Philadelphia, King of Prussia, and Berwyn.

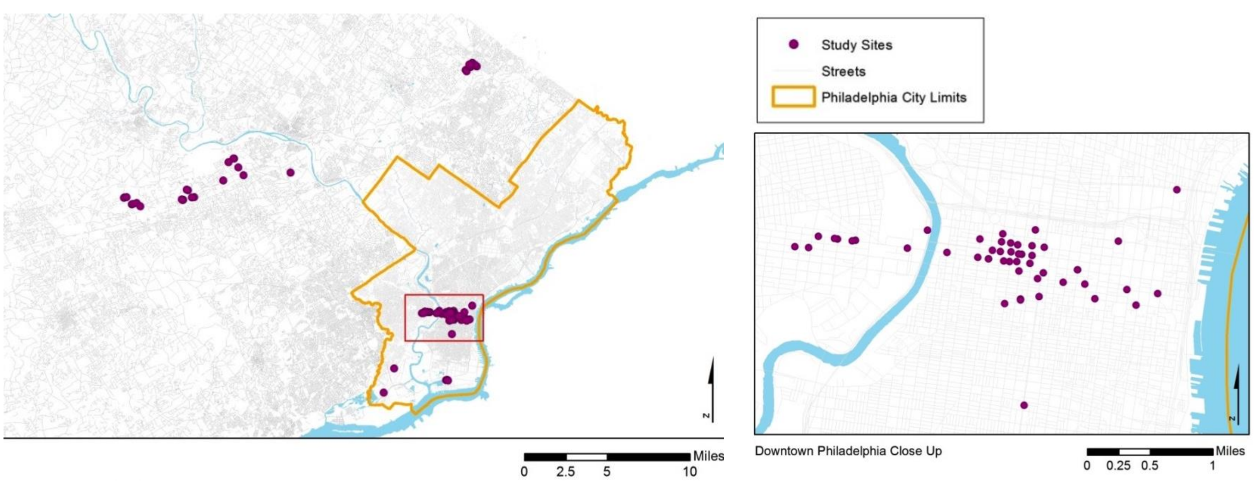


Figure 3 – Geographic Distribution of Building Study Sites (2015): (The Pennsylvania Geospatial Data Clearinghouse, 2008)

3.2 Energy Star Score

Energy Star Score is the most frequently significant dependent variable in annual data analysis. For several reasons Energy Star Score may be a better metric of building energy use than EUI. Although Energy Star Score is calculated using EUI, it also controls for many factors input into Portfolio Manager that are not retrievable in benchmarking data. Although EUI, as used here, could be controlled for building age, use type, and location, Energy Star Score additionally controls for occupancy, number of desktop computers, fuel mix, weather, and operational attributes like schedules and setbacks - a major confounding factor found in CBECS. Portfolio Manager does not release the information about number of occupants or number of computers for privacy reasons. Since factors like occupancy and plug loads can mask the impact of specific building attributes on energy consumption, therefore comparison of two buildings based solely on EUI becomes impossible. This makes Energy Star Score a better metric for determining the influence on energy use of specific building attributes.

While the team was able to collect data from cities with and without benchmarking ordinances in the same region, EUI were not available for the same years and only one full year of data was available for the towns without ordinance, preventing calculation of improved performances.

3.3 Monthly Data Analysis

For this part of the analysis, monthly energy data was disaggregated into discrete energy end uses. From the large pool of data sets available, only those buildings were selected for analysis, which were either all electric, or for which both monthly electrical and monthly gas usage data was available. Moreover, buildings with anomalous energy data, or occupancy patterns were also removed, leaving a total of 38 buildings in the first phase of analysis after the data cleansing process.

Next, each month's energy use was normalized for the number of days in the month, by dividing it by the number of days, yielding a dependent variable of energy use per day, which was then divided by the floor area to become EUI per day, or energy use per day per square foot.

Because monthly energy data is normalized by the number of days in the given month, heating and cooling degree days in the month must also be divided by the same number of days. Once this is accomplished, two independent variables exist for each dependent variable of energy use: heating degree days/day and cooling degree days/day. For this analysis, months with at least thrice as many heating degree-days than cooling degree-days, use the heating degree-days as the independent variable, and vice versa. For months whose heating and cooling degree days were almost equal, the month's data was removed from the analysis. This enabled a continuous degree-day x-axis and hence LEAN-Monthly regressions and load disaggregation could be performed.

To disaggregate particular loads, LEAN analysis as demonstrated by Kissock & Seryak was used with some modifications [10]. LEAN is most commonly used with a portfolio of buildings to distinguish between energy use patterns across the limited number of buildings in the portfolio, pointing to buildings that need the most attention and areas in which each building can improve, as done for Johnson Controls' building portfolio [11]. For this analysis, an inflection point between heatingdominated and cooling-dominated seasonal loads was used instead of separate breakeven temperatures to acknowledge residual heating and cooling that occur in reality, often simultaneously, as illustrated in figure 1 a &b. Another difference is that regression curves, rather than linear change point regression models, were used because individual heating and cooling season data points aligned best with quadratic equations rather than linear equations. For an easier visual comparison, all site energy was plotted on the same axis rather than separating heating and cooling regressions for mixed-fuel buildings based on the fuel type used. Henceforth this type of analysis will be called “LEAN Monthly” and will reference monthly analysis.

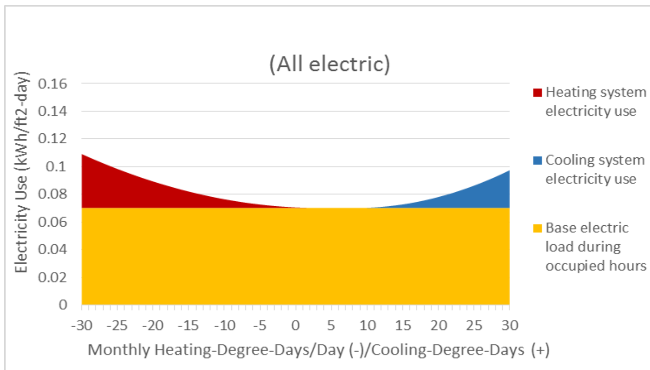


Figure 4: LEAN analysis of an all-electric building with a high baseload

3.4 Interval Data Analysis

The availability of interval data provides added value to building owners and municipalities. Current benchmarking policy in Philadelphia does not require interval data. Inclusion of interval data in any ordinance could require additional work from the building owners. Efforts to improve the task of sorting, organizing, and obtaining interval data in a seamless method could employ the SEED tool and the new features being developed by CBEI in project 4.4.

LEAN regressions applied to interval data can enhance understanding of energy use patterns and resulting costs when combined with building attribute data and building management data. LEAN analysis, relating to monthly data analysis, can be applied to interval data by using additional time points, resulting in further precision in curve modeling. Using this technique, temperature-related electric use at hourly intervals (as opposed to monthly intervals), separating occupied hours from unoccupied hours and weekends / holidays can be accomplished. Plotting energy use versus temperature during occupied or unoccupied hours parses heating and cooling specific occupied loads as well as occupied and unoccupied base-loads.

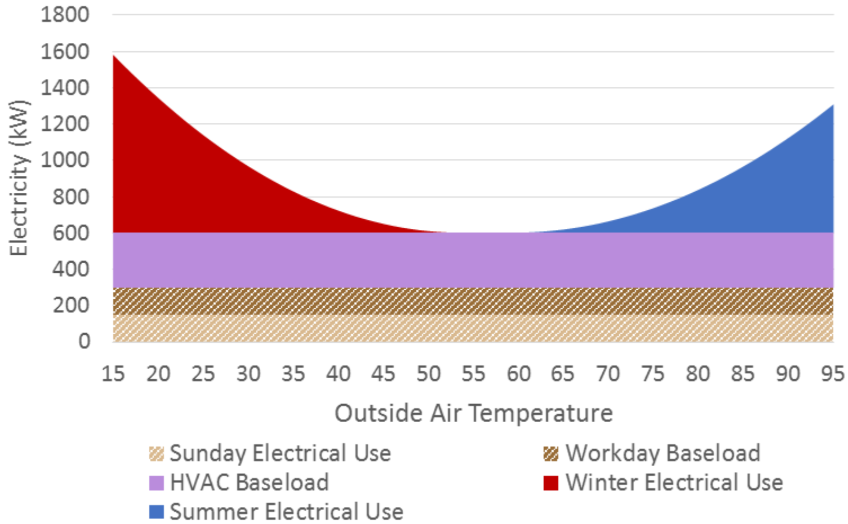


Figure 5: LEAN occupied hour methods with interval data

In addition to LEAN-derived techniques, Pacific Northwest National Laboratory (PNNL)'s add-in for Excel called ECAM was used to analyze interval data. ECAM not only uses LEAN to analyze interval data, but also graphs a building's typical daily energy use pattern averaged throughout a selected data range, separated by weekdays or holidays. Figure 6 and Figure 7 demonstrate how an average weekday differs from an average holiday, and how both differ between seasons. However, ECAM does not map occupied / unoccupied time and hence a new technique was developed to synthesize LEAN and ECAM. The technique labelled by this research as LEAN-Occupied-Hours for interval analysis uses regression curves instead of 3-, 4-, or 5-change-point models, separating curves for the heating and cooling seasons at the lowest point in the overall curve.

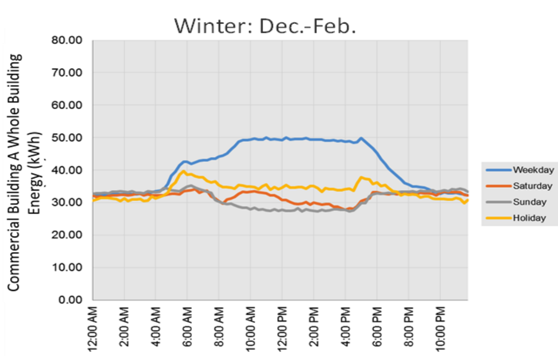


Figure 6 – Average Load Profile by Day Type in Winter (PNNL)

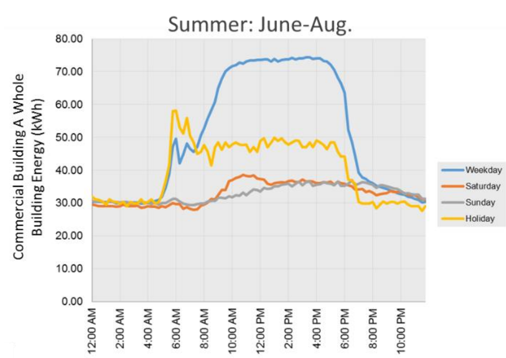


Figure 7 – Average Load Profile by Day Type in Summer

3.5 Summary of Metrics developed with these methods

This table summarized the metrics now available at different level if using the methods presented previously.

Table 1: Metrics available at various level of details

Annual	Monthly	Interval
Energy Star Score	Energy Star Score	Energy Star Score
Site EUI	Site EUI	Site EUI
Source EUI	Source EUI	Source EUI
Electricity EUI	Electricity EUI	Electricity EUI
Fuels EUI	Fuels EUI	Fuels EUI
	Peak heating load	Peak heating load
	Peak cooling load	Peak cooling load
	Overall inflection point	Overall inflection point
	Heating inflection point	Heating inflection point
	Cooling inflection point	Cooling inflection point
	Base energy use	Heating seasonal energy use
	Heating seasonal energy use	Cooling seasonal energy use
	Cooling seasonal energy use	Base electric load during occupied hours
		Base electric load during unoccupied daytime hours
		Heating seasonal energy use during occupied hours
		Cooling seasonal energy use during occupied hours
		Peak heating/cooling load during occupied hours
		Inflection point during occupied hours

4 Using these Energy Metrics to target buildings for retrofit

4.1 Using Municipal Benchmarking Data to Identify and Target Energy Inefficient Buildings for Utility Incentives

One of the value propositions examined in this project is the use benchmarking data to drive an increase in energy efficiency retrofits and associated rebates that would not have otherwise occurred for a utility. In cases where benchmarking data is disclosed to the public, utilities can utilize readily available benchmarking data to identify buildings deemed energy inefficient ("energy hogs") for targeted outreach as part of their commercial building rebate programs. The team developed a methodology for identifying subsets of inefficient properties based on ESPM data typically collected through a benchmarking program and applies to properties that received an ESPM score. (Based on conversations with DVN-GL staff in Fall 2015, the methodology described here was not used by their energy engineers when they assessed the value of benchmarking data to drive rebates.) The methodology uses certain selection metrics which guard against selection-bias such as selecting only very large properties or only properties with high total annual energy costs. The general goal of this methodology is to identify a subset of properties that have the following characteristics:

1. Low ESPM score
2. High EUI
3. High energy cost per square foot

The first step in identifying inefficient properties is to parse properties having ESPM scores of 74 or less. Properties with scores of 75 or higher are potentially qualifiers for Energy Star recognition, hence scores less than 75 are considered inefficient. The next step is to calculate the median site EUI for this group of properties. The EUI is an appropriate selection metric since it represents total annual energy use normalized by square footage, meaning any size property can be selected. Finally, using the subset of office properties selected in Step 2, the median cost per square foot for each set of properties can be calculated. If actual fuel costs are not available, average fuel costs can be used. Properties with high fuel costs typically offer shorter payback periods for retrofits, making the investment in a retrofit more attractive to the property owner.

4.2 Using Monthly and Interval Energy Data to Identify and Target Buildings for Incentives:

Analysis involving monthly and interval data holds the most value for targeting utility incentives. Regression analysis of monthly or interval data provides weather-driven heating, weather-driven cooling, and baseload load EUIs that when combined with LEAN analysis can associate certain building characteristics with high heating, cooling or baseload EUIs. Once a building has been determined to have a high heating, cooling or baseline EUI, loadshape analysis can help determine if the high EUIs are the result of unoccupied building operation, which can possibly represent both energy waste and the potential for certain retrofits and related rebates.

The following table and chart shows 14 office properties, with accurate regressions of daily energy use (weekdays) regressed against average daily outside air temperature, and their associated heating, cooling and baseload EUIs. This provides a reference set for comparing other buildings to determine if their heating, cooling or baseload EUIs are high.

Table 2 – List of EUI characteristics for 14 office buildings.

Site	Size (sq ft)	Annual Electric Use (kWh)	Annual EUI (kBTUs/sq ft/yr)	Base Load EUI (kBTUs/sq ft/yr)	Weather-Driven Heating EUI (kBTUs/sq ft/yr)	Weather-Driven Cooling EUI (kBTUs/sq ft/yr)
Bldg A	16,307	612,869	128.2	53.3	34.8	4.6
Bldg B	23,932	667,100	95.1	41.4	25.3	1.8
Bldg C	29,289	926,445	107.9	55.1	23.0	1.4
Bldg D	30,565	407,035	45.4	13.4	21.7	2.2
Bldg E	41,459	968,849	79.7	46.4	10.2	0.8
Bldg F	53,110	1,172,964	75.4	48.0	6.4	1.2
Bldg G	57,145	875,544	52.3	38.9	3.4	0.6
Bldg H	61,319	1,680,496	93.5	54.8	6.7	3.4
Bldg I	62,383	897,280	49.1	29.3	6.1	1.0
Bldg J	63,170	1,951,333	105.4	69.7	3.8	2.2
Bldg K	69,338	856,378	42.1	26.0	2.8	2.8
Bldg L	80,649	2,868,916	121.4	86.9	8.7	4.2
Bldg M	84,409	1,318,613	53.3	36.1	0.9	3.3
Bldg N	95,261	1,273,653	45.6	30.2	4.4	1.9

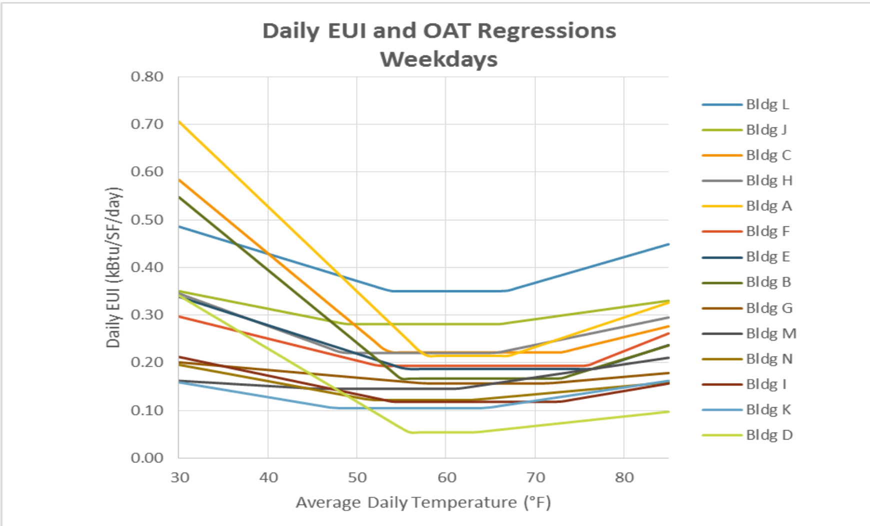


Figure 8 – 5-parameter energy/temperature regressions for 14 office buildings.

4.3 Ranking of Heating, Cooling and Baseload EUIs:

The 14 office properties are shown in the following table and chart sorted largest to smallest in three ways: baseload EUI, weather-driven cooling EUI and weather-driven heating EUI. The definition of a "high" EUI is an EUI equal to or greater than the median value. Bldg L, J and C have high baseload EUIs, Bldg A, L, H and M have high weather-driven cooling EUIs and Bldg A, B and C have high weather-driven heating EUIs. The graph clearly shows baseload EUIs as the dominate area of energy use for all 14 buildings, with weather-driven cooling EUIs having the least amount of energy consumption.

Table 3 – Sorted heating, cooling and baseload EUIs with median values for 14 office buildings.

High Baseload: Equal to or Greater Than **43.9** kBtu/sf/yr (Median Value))		High Weather-Driven Cooling Load: Equal to or Greater Than **2.0** kBtu/sf/yr (Median Value)		High Weather-Driven Heating Load: Equal to or Greater Than **6.5** kBtu/sf/yr (Median Value)	
Site	Base Load EUI (kBTUs/sf/yr)	Site	Weather-Driven Cooling EUI (kBTUs/sf/yr)	Site	Weather-Driven Heating EUI (kBTUs/sf/yr)
Bldg L	**86.9**	Bldg A	**4.6**	Bldg A	**34.8**
Bldg J	**69.7**	Bldg L	**4.2**	Bldg B	**25.3**
Bldg C	**55.1**	Bldg H	**3.4**	Bldg C	**23.0**
Bldg H	**54.8**	Bldg M	**3.3**	Bldg D	**21.7**
Bldg A	**53.3**	Bldg K	**2.8**	Bldg E	**10.2**
Bldg F	**48.0**	Bldg D	**2.2**	Bldg L	**8.7**
Bldg E	**46.4**	Bldg J	**2.2**	Bldg H	**6.7**
Bldg B	41.4	Bldg N	1.9	Bldg F	6.4
Bldg G	38.9	Bldg B	1.8	Bldg I	6.1
Bldg M	36.1	Bldg C	1.4	Bldg N	4.4
Bldg N	30.2	Bldg F	1.2	Bldg J	3.8
Bldg I	29.3	Bldg I	1.0	Bldg G	3.4
Bldg K	26.0	Bldg E	0.8	Bldg K	2.8
Bldg D	13.4	Bldg G	0.6	Bldg M	0.9

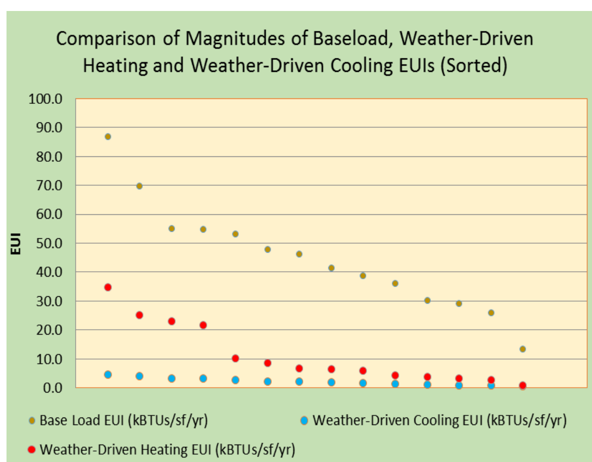


Figure 9 – Sorted Magnitudes of heating, cooling and baseload EUIs for 14 office buildings.

Other buildings can be analyzed and compared to these distributions to see if they have relatively high heating, cooling or baseload EUIs. If a building is found to have, for instance, both a high baseload EUI and interior lighting that was on at night (determined by a drive-by assessment), then the LEAN analysis done for buildings with these characteristics indicates a strong potential for energy savings. To further assess if a rebate is warranted or not for the building, loadshape analysis can be done to verify if, in fact, significant nighttime energy use is occurring. If the loadshape analysis reveals loadshape characteristics commensurate with interior lighting on during unoccupied hours, there is a high potential the building could benefit from a lighting control system or occupancy sensors, both of which can receive incentives.

5 Analyzing Building Attribute Data to Target Buildings for Retrofits

Several methodologies were explored for comparing measured energy data to building attributes. These methods included ANOVA, Regression, LEAN Monthly, LEAN occupancy, and five Machine Learning techniques. Analyses were conducted with interval, monthly, and annual energy data in order to understand which attributes could be found to be significant when different levels of data were available). This was done in order to establish a replicable model for both future research as well as benchmarking programs around the country that may be interested in analyzing energy data relative to building attributes.

We hypothesized that measured energy use data, including benchmarking data, in combination with data on building attributes, can be used to identify and prioritize targeted energy efficiency measures specific to stakeholders within a particular city or area.

In order to test this hypothesis, a large dataset was created. This dataset included 44 distinct building attributes (independent variables), which together amounted to over 5,500 data points. ANOVA statistical methods were used to analyze this data and discern statistical significance of relationships between building attributes and measured energy use. The results of this analysis were a series of significant impacts of building attributes on energy use, ranging from thermostat setbacks to WWR. These findings should be considered to be specific to the dataset, and while similar findings may be found in analysis of other data sets, variables such as climate zone, microclimate, and use type make these findings difficult to generalize. The methodology presented should primarily be considered useful to specific stakeholders within the region analyzed, in this case the greater Philadelphia and Washington, D.C. areas. After running over 240 statistical analyses, 34 significant relationships were found.

GIS mapping of attributes vs EUI showed no significant spatial distribution of attributes that could justify selecting a specific block or neighborhood base on relationship between energy metrics and attributes.

5.1 Data Collection

In order to gather data on physical building attributes this research explored a variety of methods. After an initial list of sub-hypotheses was created, a list of data points necessary to test each hypothesis was written. Each of these points was categorized into one of four groups: information available from publicly available geodatabases such as Google Maps and municipal GIS catalogues, information available by driving by the building, information available by an up-close inspection of the building, and information that would only be available by contacting building managers or owners. Ease of information gathering was prioritized, and it was found that a large number of data points needed for each building were able to be obtained through Google Maps alone. A cross-section of these data points is shown in table 1.

Following this step, building visits were made to gather information such as the number of glazing layers which was not available using online tools. Two sets of building visits were done – one visit during the day when the building was in operation, and one during the night in order to find whether or not lighting was left on during the night or turned off, and what percentage of lighting. Annual energy use and Energy Star score were publicly available while monthly energy consumption were obtained from the partner utilities. Annual building energy data for Philadelphia gathered through the city benchmarking portfolio and annual energy data for Washington, D.C. gathered through the BuildSmart DC website provide the annual-related dependent variables of Energy Star Score, site EUI, source EUI, and energy use by fuel type [9].

Table 4 - Building attribute data availability by data source

	Data points available	Portfolio Manager	Google Maps	Building exterior visit	Building exterior inspection	Interval Data	Asset Scoring Tool	Occupant Input	Metric	Evaluation Input	Time Required per Building
Potential Retrofit Measures											
HVAC											
1	Rooftop Packaged AC Units	○	●	○	○	○	●	○	Total Number of units and fans	Number	2 min
2	Rooftop Chillers	○	●	○	○	○	●	○	Total Number of units and fans	Number	
3	Rooftop Cooling Towers	○	●	○	○	○	●	○	Total Number of units and fans	Number	
4	Rooftop Condenser AC Units	○	●	○	○	○	●	○	Total Number of units and fans	Number	
5	Window AC units	○	●	○	○	○	●	●	Total Number of units	Number	
6	Shading of AC units	○	●	○	○	○	○	○	Nearby shading	# shaded chillers or RTUs	
Glazing											
7	# Glazing layers	○	○	○	●	○	●	●	Number of glazing layers	Number	30 sec
8	Tinted Glass	○	●	●	○	○	○	●	Shade of glass	No tint, Slight tint, Dark Tint, Reflective/Mirror	30 sec
9	Window frame material	○	○	○	●	○	●	●	Material	Metal, Wood, Vinyl	30 sec
10	Operable windows	○	●	○	○	○	○	●	Percentage of windows that are operable, rounded to 5%	percent	1 min
Shading											
11	External shading (on each orientation)	○	●	○	○	○	●	●	Estimated depth of external shading from the glazing surface	Feet	1 min
12	External shading device type	○	●	●	○	○	●	●	Type of external shading device used, if any	Horizontal, Vertical fin, Eggcrate	30 sec
13	Internal shading	○	●	○	○	○	○	●	Type of internal shading used, if any	Horizontal blind, Vertical blind, Curtain, Roller shade, None	1 min
Other											
14	Roof reflectivity	○	●	○	○	○	●	○	Color of the roof surface	Black, White, Gray, Brown, Glass	30 sec
15	Lighting fixture design	○	○	●	●	○	●	●	Type of ceiling mounted lighting fixtures	Pendant, Recessed, Parabolic louver	30 sec
16	All electric building	●	○	○	○	●	●	○	Gas bills less than 500 CCF/yr or steep electric heating curve	Yes/No	2 min

	Data points available	Portfolio Manager	Google Maps	Building exterior visit	Building exterior inspection	Interval Data	Asset Scoring Tool	Occupant Input	Metric	Evaluation Input	Time Required per Building
New Building Design											
Building Plan											
1	Gross Floor Area	●	○	○	○	○	●	○	Reported Square footage	square feet	0 min
2	Floor Plate Size	●	●	○	○	○	●	○	(Gross floor area)/(number of floors)	square feet	calculated
3	Orientation	○	●	○	○	○	●	○	Facades with greatest surface area	N/S, E/W, NE/SW, NW/SE, Equal faces	30 sec
4	Number of floors	○	●	○	○	○	●	●	Number of floors in building	number	1 min
5	Building Depth	○	●	○	○	○	●	○	Depth of building, in feet	feet	1 min
6	Building Shape	○	●	○	○	○	●	○	Categories of building layouts such as "L shaped" or "E shaped" (Cochran, 2014).	B1, B2, D, O1, X2, A1, A2, C1, C2, E, F, H1, H2, T1, T2, L	30 sec
7	Façade area/floor area ratio	○	●	○	○	○	●	○	Calculated wall area/reported floor area	square feet per square foot	2 min
8	Proximity to other buildings at each orientation	○	●	○	○	○	○	○	Measured distance to nearby buildings	feet	2 min
Glazing											
9	Overall WWR	○	●	●	○	○	●	○	Weighted average WWR of all facades	percent	calculated
10	WWR by facade	○	●	●	○	○	●	●	WWR at each façade estimated to 10%	percent	5 min
Other											
11	Envelope materials	○	○	●	●	○	●	●	Primary façade materials	Brick, Stone, Concrete, Steel, Spandrel panel cladding, Metal cladding, Glass	30 sec
12	Roof pitch and orientation	○	●	○	○	○	●	○	Cardinal direction and degree of roof tilt	N, S, E, W; Degrees	1 min

	Data points available	Portfolio Manager	Google Maps	Building exterior visit	Building exterior inspection	Interval Data	Asset Scoring Tool	Occupant Input	Metric	Evaluation Input	Time Required per Building
Building Management											
Lighting											
1	Lighting use at night	○	○	●	○	○	○	○	Percentage of lights on at night, rounded to 25%	percent	1 min
HVAC											
2	Thermostat Heating Setbacks	○	○	○	○	●	●	○	Setbacks visible in interval electrical data	Yes/No	5 min
3	Thermostat Cooling Setbacks	○	○	○	○	●	●	○	Setbacks visible in interval electrical data	Yes/No	

5.2 Research Findings

In this project, methodologies were presented for the development of tools to harness energy benchmarking data in order to target utility rebates and retrofit incentives for existing buildings within a particular region. One of the major findings of this project was the invention of a new LEAN analysis method that we have titled “LEAN Occupied Hours,” which represents in detail how energy is used during a building’s occupied hours. This method adds the measurements of occupied base load, unoccupied base load, and, in some cases, HVAC base load to the list of available disaggregated loads already including weather-driven loads. So far, we have shown this method to be accurate within 2%.

This research has thus shown that impacts of building attributes on energy use is capable of discerning the impact of building attributes on energy use, confirming the initial hypothesis:

5.2.1 Annual Data

Several conclusions can be drawn from table 5 : In general, annual data was not as useful in analyzing building attributes when compared with monthly or interval energy data. Site EUI is a very limited metric for understanding energy use patterns and not an adequate means to compare buildings to one another, as too much of a building’s deviation from the average EUI is due simply to occupancy, hours of operation, weather, location, or anomalous meter readings, none of which are actionable drivers of energy use in terms of building retrofits, and will inevitably obscure the analysis of any other driver of energy use. However, many areas with benchmarking ordinances only require the release of annual energy data through Portfolio Manager. When only annual data is available, analysis of some building attributes may still be fruitful by comparing them with Energy Star Score. Energy Star Score was found to be a consistently useful metric by which to analyze building attributes, and certain effects of building attributes were found to be visible through Energy Star Score that were not relevant to interval data. This is due to the many controls for building operation that are built into Energy Star Score.

5.2.2 Conclusions on LEAN Methodologies

Multiple levels of energy analysis were conducted in parallel: results that can be found in interval, monthly and annual data. Each method was used to analyze the same building attributes. It was shown that certain building attributes could be consistently found in one type of data but not another type. Separate conclusions can be drawn from each level of energy analysis. The type of data for which each attribute was found to display significant results is listed in

Table 5 below.

Table 5 - Type of Energy Data in Which Specific Building Attributes Can be Found

	Interval	Monthly	Annual
Lights on at night	●	○	○
Thermostat setbacks	○	●	○
Orientation	○	●	○
Building shape	●	●	○
Proximity	○	○	○
Façade-area/floor-area ratio	○	●	○
Number of floors	○	○	○
Floor plate size	○	○	○
Building depth	○	○	○
Dark glass	○	○	●
Operable windows	○	○	○
Overall window/wall ratio	●	○	● (ES score)
WWR by façade	●	●	● (ES score)
Number of glazing layers	○	○	○
Window frame material	○	○	○
External shading depth	○	○	○
External shading device type	○	○	○
Internal shading	○	○	○
Roof reflectivity	○	○	○
Envelope materials	○	○	○
Lighting fixture design	○	○	○
Rooftop cooling towers	○	○	○

5.2.3 Monthly versus Interval Data

In comparing two methods of LEAN analysis (LEAN-Monthly and LEAN-Occupied-Hours), both methods were shown to be useful in finding trends correlating energy use with building attributes. However, interval data provides the opportunity for disaggregating baseload energy use into more discrete packages, allowing for more detailed analysis of building attributes. Interval energy was also particularly useful in discerning building management practices. Although only two management practices were analyzed here, with more detailed data on specific building attributes more could be analyzed, and these may represent the highest return on investment of all energy conservation options. Interval data is extremely valuable for use in targeting utility rebates, and should be made a priority in data collection. Claims about interval energy data not yielding applicable information relating to tangible energy billing results are no longer founded, given the findings in this study.

LEAN-Monthly analysis was also valuable. The main advantage of monthly data over the interval dataset analyzed here was the availability of gas data. This was not available for many of the buildings in the interval data set. The combination of interval electrical data with monthly gas data is the most likely to yield actionable findings.

5.3 Statistical findings

By utilizing multiple investigative techniques (ie ANOVA, Regression, LEAN monthly, LEAN occupancy, machine learning), the team was able to verify the validity and accuracy of the methodologies.

Most commonly, statistical correlations were found in some of these metrics but not in all of them, demonstrating the need to quantify energy consumption in various ways in order to understand how energy is affected by building attributes. In most cases, the results of this analysis corroborate existing knowledge of building design and management, as in the case of buildings that leave the majority of lights on at night compared with those that do not. This research shows that based on the average 100k ft^2 office building and the cost of electricity in the Mid-Atlantic region, buildings that turn off their lights at night can save an average of $32,500 per year in utility costs. These savings could be gained either by installing lighting controls such as vacancy sensors, or programming building automation systems to turn out lights during unoccupied hours (both existing rebate options). The broader adoption of vacancy sensors or the adjustment of the Building Automation System (BAS) could be facilitated by targeted rebates. Additionally, buildings with heating setbacks have lower weather-driven heating loads and begin heating at lower temperatures on average. Appropriate setback implementation can also be supported by a targeted rebate program.

The most significant results of ANOVA statistical analysis conducted in the course of this research are listed below. The preliminary findings were confirmed when final calculations were conducted.

1. Buildings with Dark glass (T-Vis <0.5) have a Higher Electric consumption than other buildings ($p=0.014$)

2. Buildings with cooling towers use more electricity (p=0.022)
3. External shading on both the south and west facades is correlated with higher ES Score (p=0.084)
4. Buildings with only 0-25% of their lights on at night can save 32% of the average total lighting load, as they show a 45% reduction in unoccupied baseload energy use (p=0.027).
5. Buildings with heating setbacks use 60% less seasonal heating energy across all temperatures than do buildings lacking setbacks (p=0.009), as found using LEAN Monthly analysis.

5.4 Automating the analysis

The analysis of our dataset was conducted manually and was time consuming. Therefore, the team decided to investigate beyond the scope of the research and create solution that would automate the analyses. This study developed 2 applications: one to estimate static building electric usage (i.e. annual electric usage) and another to forecast dynamic building electric usage (i.e. hourly electric usage of an future hour), together realizing a comprehensive prediction of building electric usage.

We first compared most popular machine learning algorithms applied in related existing studies including Artificial Neural Network (ANN), Support Vector Machine (SVM) and Decision Tree Regression. The most suitable modeling algorithm was retained and then optimized based on our dataset. Feature Selection methods with single-score metrics such as Mutual Information and Pearson Correlation are used to target the feature that has the most significant and direct correlation with energy consumption.

Ultimately, this tool would allow program manager to apply our research findings without having to replicate our methods.

6 Targeting Rebates Customers with Energy Data Analytics

6.1 Incentive Mapping:

The following table shows an incentive mapping schema based on the commercial incentives and rebates currently offered by PECO Energy, under their Smart Ideas program, for electric fuel reductions. Based on the statistical findings, CBEI prioritized the current Smart Ideas rebates according to the data available. Specifically, if interval data is available and supporting data corroborate the suggested incentive, the opportunity for rebate is high. Similarly if only monthly data is available the opportunity that rebate would be applicable is more difficult to establish and the ranking is low. Based on the interval data and drive-by assessments, the Incentive Mapping Table shows the opportunity for targeting a building for a particular rebate, as an example, an interior lighting incentive (such as installation of occupancy sensors) is ranked as a 1 (highest) for this building. For a one-time cost of adding advanced lighting sensors and controls (ex. Utilize a rebate to install occupancy and/or daylight sensors), a typical office building of 100,000 SqFt could save ~$32,000 per year.

Table 6 – Incentive Mapping table

Incentive Category	Assessment from Data Analysis:	Applicable Type of Prescriptive Incentives:		Applicable Type of Custom Incentives:	Energy Data Granularity: Monthly	Energy Data Granularity: Interval	Significant Energy Usage During Unoccupied Hours	Supporting Data: Satellite Imagery Analysis or Drive-By Assessment	Targeted Rebate Opportunity (Ranked Within Incentive Category; 1 being the best)
HVAC	High Electric Cooling or Heating EUI (kBtu/sf/yr)	Packaged or Split-System Air-Source AC or Heat Pump	Water-Source Heat Pump	Chiller		O		O	1
						O	O	O	2
						O			3
						O	O		4
					O			O	5
					O				6
		Energy Management System				O	O		7
Variable Frequency Drives	High Heating Season Electric EUI (kBtu/sf/yr) for Gas-Heated Building	VFD on HVAC Fan	VFD on Heating Hot Water Pumps			O	O	O	1
						O		O	2
						O	O		5
						O			6
					O			O	11
					O				12
	High Baseload Electric EUI (kBtu/sf/yr)	VFD on HVAC Fan	VFD on Chilled Water or Heating Hot Water Pumps			O	O	O	3
						O		O	4
						O	O		7
						O			8
					O			O	13
					O				14
	High Electric Cooling or Heating EUI (kBtu/sf/yr)	VFD on HVAC Fan	VFD on Chilled Water or Heating Hot Water Pumps			O	O	O	9
						O		O	10
						O	O		15
						O			16
Interior Lighting	High Baseload Electric EUI (kBtu/sf/yr)	CFL, T12 to T8, T8 to Reduced Wattage T8, Ceramic Metal Halide Lamp, Daylight Sensor Controls, Occupancy Sensors, LED Exist Signs and Interior LED lighting				O	O	O	1
						O		O	2
						O	O		3
						O			4
					O			O	5
					O				6
Exterior Lighting	High Exterior Lighting Load	Exterior LED, CFL or Metal Halides				O	O	O	1
						O	O		2
								O	3

Incentive Category: Installation of specific measures related to HVAC, variable frequency drives, interior lighting and exterior lighting can receive a rebate.

Assessment from Data Analysis: Indicates if high weather-driven heating, weather-driven cooling or baseload EUI was assessed.

Applicable Type of Prescriptive Incentive: Lists the different types of equipment that can receive a prescriptive type of rebate.

Applicable Type of Custom Incentive: Chillers must be rebated as a custom incentive.

Energy Data Granularity – Monthly or Interval: Indicates what type of data was analyzed.

Significant Energy Usage During Unoccupied Hours: Loadshape analysis of interval data showed significant energy use during weekday nights and/or weekends.

Supporting Data: Satellite Imagery Analysis or Drive-By Assessment: Information about certain building characteristics, such as HVAC system type, allows for specific targeting of rebates.

Targeted Rebate Opportunity (Ranked Within Incentive Category; 1 being the best): Based on granularity of energy data, amount of energy usage during unoccupied hours and availability of supporting data, opportunities for rebates are ranked (1 being the best) within each Incentive category.

6.2 Example of How Incentive Mapping Works:

The building located at [address withheld for privacy concerns] is a good example of how incentive mapping works. The assessment from the interval data analysis for building Gg (See Figure 10) indicated:

1) High baseload consumption

2) Significant energy usage during unoccupied hours

3) Supporting data from a drive-by assessment indicated significant lighting on at night.

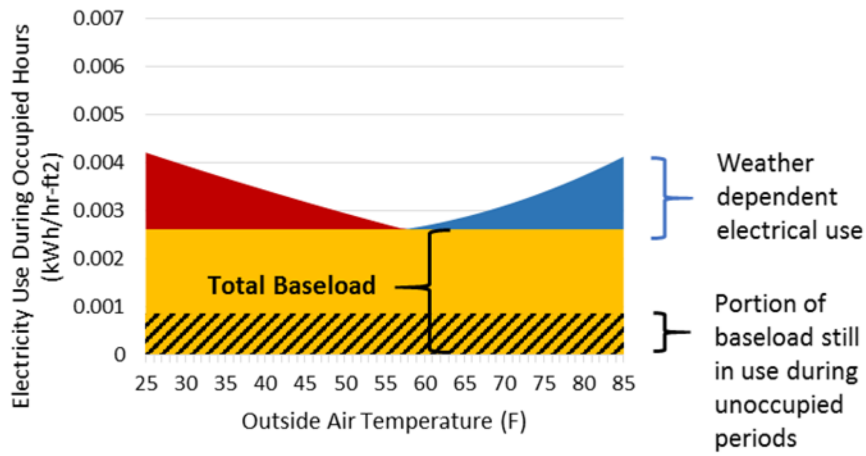


Figure 10 – Example of electricity load distribution and base load assessment during both occupied and unoccupied hours

Based on the interval data and drive-by assessments, the Incentive Mapping Table shows the opportunity for targeting an interior lighting incentive (such as installation of occupancy sensors) is ranked as a 1 (highest) for this building.

Table 7 – Incentive Mapping table

Incentive Category	Assessment from Data Analysis:	Applicable Type of Prescriptive Incentives:	Applicable Type of Custom Incentives:	Energy Data Granularity: Monthly	Interval	Significant Energy Usage During Unoccupied Hours	Supporting Data: Satellite Imagery Analysis or Drive-By Assessment	Targeted Rebate Opportunity (Ranked Within Incentive Category; 1 being the best)
Interior Lighting	High Baseload Electric EUI (kBtu/sf/yr)	CFL, T12 to T8, T8 to Reduced Wattage T8, Ceramic Metal Halide Lamp, Daylight Sensor Controls, Occupancy Sensors, LED Exist Signs			O	O	O	1
					O		O	2
					O	O		3
					O			4
				O			O	5
				O				6

Another example would be a building with a high weather-driven cooling EUI, no significant energy use during unoccupied hours and no supporting data. The Incentive Mapping Table below shows the opportunity for targeting an HVAC incentive would be ranked a 3.

Table 8 – Incentive Mapping table

Incentive Category	Assessment from Data Analysis:	Applicable Type of Prescriptive Incentives:		Applicable Type of Custom Incentives:	Energy Data Granularity: Monthly	Interval	Significant Energy Usage During Unoccupied Hours	Supporting Data: Satellite Imagery Analysis or Drive-By Assessment	Targeted Rebate Opportunity (Ranked Within Incentive Category; 1 being the best)
HVAC	High Electric Cooling or Heating EUI (kBtu/sf/yr)	Packaged or Split-System Air-Source AC or Heat Pump	Water-Source Heat Pump	Chiller		O		O	1
						O	O	O	2
						O			3
						O	O		4
					O			O	5
					O				6
		Energy Management System				O	O		7

In general, monthly energy data (i.e., typical monthly utility bill) can be used for regression, but will have energy values containing combined occupied and unoccupied energy use. Hence, the Incentive Mapping

Table ranks these assessments lower than interval data assessments, except in two cases relating to variable frequency drives (see main Incentive Mapping Table above).

6.3 Using benchmarking data to target buildings for retrofit based on building attributes.

The list below identifies the top 6 attributes that can be identified using only information available from Benchmarking ordinances. There are additional attributes, however these six provide the largest return on investments based on the rebate, cost of installation, and estimated energy savings. Based on our statistical analysis of benchmarking data we identify the following correlations that can help target buildings that have the most potential for energy saving with rebate programs:

1. **Solution 1: Replace glazing on buildings with dark glass (T-Vis<0.5) will lower electric consumption.**
   Finding1: Buildings with Dark glass (T-Vis <0.5) have a Higher Electric consumption than other buildings.
   Utilize PECO Glazing rebates or target "Energy Management System" Rebate to those building to reduce their energy use.
2. **Solution 2: Provide new lighting schedule controller to turn light OFF at night to reduce the building EUI.**
   Finding 2: Buildings with lights ON at night have a higher source EUI than others.
   Utilize PECO Custom Lighting Rebates to install a lighting schedule controller and program a night schedule for lighting systems.
3. **Solution 3: Provide new cooling tower to building owners of medium and large buildings to significantly reduce energy use.**
   Finding 3: Buildings with cooling tower use more electricity than others.
   Utilize PECO Custom incentive for chiller in their smart idea program.
4. **Solution 4: Provide upgraded cooling tower especially to small buildings will significantly reduce energy use.**
   Finding 4: Smaller building with cooling tower more energy.
   Utilize PECO Custom incentive for chiller in their smart idea program.
5. **Solution 5:Install shading device on south and west façade to increase energy star score**
   Findings 5: External shading on both the south and west facades is correlated with higher ES Score.
6. **Solution 6: Install new BAS and includes nighttime setback to lower the Building Site EUI.**
   Findings 6: Buildings with High EUI most likely do not have a proper setback schedule and would benefit from using a rebate to program night and week end set back in their Building automation system.
   Propose the PECO "Energy Management System" Rebate to those buildings to reduce their energy use.

6.4 Future Work

The team plans to continue to develop its database of buildings and extend its work to other cities incorporating different climatic zones. Future plans place a higher focus on automating various levels of analysis using machine learning techniques, which proved highly successful in our preliminary studies. The team also plans to find ways for automation of the collection of building attributes data by collecting them directly from Asset Score Tools when the option of reading from the tool is made available in the future.

Conclusion

This one year project developed new and improved methods of analyses of energy data that provide new metrics valuable to assist in targeting rebate customers for greater energy savings and recruitment of suitable buildings.

If program managers apply both of those methods to their territory they can improve their outreach effort by focusing on the buildings that will save the most energy and focus on 15% to 28% of their customers depending on the methods they select as detailed in deliverable D6.4.2.