Delivering Value to Commercial Customers with Load Shape Segmentation

Efficiency Vermont R&D Project: Greenhouse Gas Reduction

December 2022

Abigail Hotaling

Efficiency Vermont

20 Winooski Falls Way
Winooski, VT 05404

Contents

Executive Summary

Vermont's push to reduce greenhouse gas emissions through beneficial electrification and increased renewable energy production necessitates improvements in energy load management. Using advanced metering infrastructure (AMI) data, Efficiency Vermont can identify and prioritize projects to optimize program goals.

Using a random sample of 10,000 commercial customer meters, this study demonstrates how load clustering, can be used to identify typical load patterns and group similar patterns together. The load shape clustering results help illuminate potential areas for focused flexible load management and indicate what technology, incentives, or rates would best optimize the identified load patterns. For example, this allows programs to focus on customers with time-of-use patterns coincident with grid peaks.

Also, an AMI features-based approach to summarizing energy use data was assessed. Example features are baseload, heating and cooling degree day model slopes, and seasonal peak demand. Specific features can help identify use patterns which are well-suited for certain efficiency measures or programs. For example, baseload can capture equipment that is "always-on" such as refrigeration equipment that runs constantly. Lighting and HVAC, in contrast, tend to be utilized more heavily during occupied hours, and HVAC load tends to be driven by outdoor temperature. The features-based analysis allows the team to assess specific features of energy use depending on program requirements and potential customer benefits to identify customers with the most savings potential.

The research team used 331 previously completed commercial retrofit projects to test using AMI features for customer opportunity identification. Results show that using AMI features to focus program participation can increase overall program savings. Focused program participation driven by normalized energy features enable an increase in the depth of savings or percent savings relative to total usage. Evaluating customers based on normalized energy features provides visibility into ideal program customers beyond just the largest users. Results were also grouped by three primary measure categories: HVAC, lighting, and refrigeration. Not surprisingly, the relationship with energy usage and cooling degree days correlated much more strongly with increased HVAC savings than with lighting savings. Customer selection on the basis of summer peak demand also correlated with higher HVAC savings. Compared to other features, discretionary use and base use showed the greatest savings opportunity for lighting projects. Customer selection for refrigeration measures showed increased average daily savings with additional selections based on all other features except heating slope.

Opportunity exists to further scale the classification of load patterns which may improve the trends that can currently be identified in the data. There is also opportunity to conduct this analysis for specific regions on the grid, such as a substation that is particularly constrained during peaks. Building off these results, Efficiency Vermont will investigate how these methods can be applied to the residential sector, implemented equitably and effectively into its programs and processes, and employed to improve its customer opportunity identification.

Introduction

Efficiency Vermont's programs have consistently delivered the value of energy efficiency to customers for more than twenty years. Efficiency Vermont continues to evolve to maintain the high rates of return that efficiency is known for, and to help the State meet its decarbonization goals pursuant to the 2020 Global Warming Solutions Act. In addition to tried-and-true measures, such as lighting upgrades and weatherization, which reduce a building's energy load, Efficiency Vermont programs increasingly deliver benefits to customers through beneficial electrification—electrification of appliances, heating and hot water systems, and transportation. Switching from fossil fuels to efficient electric systems saves money for consumers and is essential for decarbonization. Yet, it also increases the demand for electricity, which can strain the grid if not properly managed.

As beneficial electrification grows and utilities integrate more intermittent renewable energy resources onto the grid, it is increasingly important to monitor and balance the timing of energy use. High-frequency reporting from advanced metering infrastructure (AMI) data supports data-driven energy management strategies. This study investigates and tests new ways to employ high-frequency AMI data to group customers with similar use patterns (load shape segmentation) and focus efficiency investments on those commercial customers with the highest potential savings—those who can achieve high magnitude savings, as well as those who can attain large savings relative to their size.

Efficiency Vermont can use effective customer opportunity identification to increase project savings potential and enable market transformation on the broadest possible scale by identifying new customers. Efficiency Vermont can also us this approach to inform program design and to focus customer outreach on customers that meet a certain criterion based on energy use characteristics that make them a better fit for the program.

In this study, Efficiency Vermont measured how a focused approach to opportunity identification affects average at-the-meter-savings. Using load clustering,[1] the research team identified typical load patterns and then grouped similar patterns together in a process known as load shape segmentation. With the information provided from the load shape segmentation analysis Efficiency Vermont seeks to strategically identify customers who can achieve high savings, support grid planning, and expand flexible load management (FLM) initiatives.[2]

Background

AMI data enables energy efficiency implementors to expand programs focused on meter-based savings, such as pay-for-performance, strategic energy management, and flexible load

[1] Load shape clustering identifies typical consumption patterns across a population.
[2] Since 2018, VEIC has partnered with Vermont distribution utilities to test a variety of approaches to FLM with commercial and industrial clients. Beginning in 2021, Efficiency Vermont is supporting up to 50 commercial and industrial customers enrolled in Phase II of Green Mountain Power's FLM 2.0 pilot program with advanced data analytics and modelling. [Read more about this work here]

management.[3] Energy consultants and account managers typically identify customers with significant savings potential based on their knowledge of specific facility operations and equipment. Program resources often flow to high-use commercial and industrial customers where the magnitude of electricity use creates significant savings opportunities. Yet, when trying to scale programs, it is not cost effective for programs to analyze each customer's energy use, individually. Efficiency Vermont can pinpoint untapped savings opportunities across portfolios of customers by using automated analytics and AMI data to identify customers with significant savings potential.

Across the country, program administrators now employ energy use analytics to test and implement enrollment and outreach strategies in both commercial and residential sectors. Pacific Gas and Electric (PG&E) conducted a retrospective case study which reported significantly improved potential energy savings in lighting, refrigeration, and HVAC programs when energy use guided customer outreach.[4] Based on these and similar results, PG&E identified data analytics and focused customer outreach as key strategies in their 2018-2025 Energy Efficiency Business Plan.[5] Power TakeOff, a leader in AMI data analytics, demonstrated a successful program design which proactively engaged small and medium business customers based on load shape characteristics.[6] TECH Clean California, a statewide initiative to advance electric space and water heating, uses electric and gas meter data to engage customers who will derive the most benefit from heat pump installations.[7]

Load shape segmentation has also shown significant potential for demand response. Lawrence Berkely National Laboratory developed a load shape clustering approach using AMI data as part of California's Demand Response Potential Study.[8]

Methods

The research team used the following research methods.

---

[3] Pay-for-performance programs incentivize customers based on savings performance at the meter, usually using whole facility savings calculations. Similarly, savings from strategic energy management programs are typically measured using whole facility meter calculations. Participating Vermont distribution utilities are using a pay-for-performance bill-credit to incentivize customers in the flexible load management pilots.

[4] Scheer, A., et al. (2018, August 10). *Energy Efficiency Program Targeting: Using AMI data analysis to improve at-the-meter savings for small and medium businesses*. Retrieved from CALMAC: https://www.calmac.org/publications/SMB_Targeting_Whitepaper_FINAL_8-10-18.pdf

[5] *Energy Efficiency Business Plan*. (2017, January 17). Retrieved from PG&E: https://www.pge.com/pge_global/common/pdfs/for-our-business-partners/energy-efficiency-solicitations/PGE-Energy-Efficiency-Business-Plan-2018-2025.pdf

[6] Widmer, P., & Tonielli, R. (n.d.). *Achieving Evaluated Behavioral, Retro-Commissioning, and Operational Savings with Small and Medium Businesses through AMI-Enabled Analytics and M&V 2.0.*

[7] Oppelstrup, M., et al. (2022). *Meter-Based Targeting for Beneficial Electrification at Scale*. Retrieved from ACEEE: https://aceee2022.conferencespot.org/event-data/pdf/catalyst_activity_32468/catalyst_activity_paper_20220810190535382_6bc589fa_e8b8_4cb3_9ce7_36b1c042fab3

[8] Murthy, S., et al. (2022). *A multi-level load shape clustering and disaggregation approach to characterize patterns of energy consumption behavior*. Retrieved from ACEEE: https://aceee2022.conferencespot.org/event-data/pdf/catalyst_activity_32662/catalyst_activity_paper_20220810191656362_a84e8e9f_f8ec_465c_bee3_c57ce688848a

Data

The research team examined data for commercial customers within Green Mountain Power (GMP) territory. GMP is the largest Vermont utility serving approximately 266,000 customers, with AMI data recorded every quarter-hour. The methods employed in this study can be applied to any AMI data recorded at least once per hour.

GMP's data set contained 32,246 active commercial utility meters with sufficient historical data to perform the analysis.[9] The research team conducted load shape segmentation for a random sample of 10,000 of those commercial meters.

Load Shape Clustering

Load shape clustering allows analysts to distill thousands of individual load shapes for each meter into representative load patterns across a population. Individual meters can then be grouped into their corresponding patterns. This helps program administrators understand what types of load patterns are most common in a particular sector or region and categorize usage types. These load patterns can help drive customer opportunity identification particularly for programs where time-of-use patterns are important.

Cluster analysis uses unsupervised learning algorithms to group together sets of data such that the data within a cluster are more similar to each other than they are to any other group. The research team used the K-Means algorithm to cluster a sample of utility AMI data.[10] The objective of K-Means is to group similar data points together given a pre-determined number of clusters (*k*). Each data point is allocated to the closest cluster based on the cluster *center*. The center is the mean of the data points within a cluster. Figure 1 shows an example of how the K-Means algorithm can take un-categorized data and group the points into similar clusters. The mean of each cluster, or the cluster center is shown in grey.

---

[9] Data was required from 2017 to 2022. Utility accounts with zero usage were removed from the analysis.
[10] Luo, X., Hong, T., Chen, Y., & Piette, M. A. (2017). *Electric load shape benchmarking for small- and medium-sized commercial*. Retrieved from https://www.sciencedirect.com/science/article/abs/pii/S0306261917309819

Figure 1. Example of using K-Means clusters to take un-categorized two-dimensional data and group that data into three clusters of similar data points.

The research team's analysis uses the approach described above to fit the K-Means algorithm and define the clusters but is based on the average seasonal 24-hour load shape, normalized by each meter's total daily use for each meter. Normalizing the data helped capture the representative load patterns and time-of-use characteristics by un-weighting magnitude. This study analyzed both summer and winter load shapes.[11]

With the training set of data defined, the team determined the optimal number of clusters (k) that most accurately represented the seasonal average daily load patterns in the data. The team used the within-cluster sum of squared error (SSE) to determine the optimal number of clusters for the analysis.[12] The team added additional clusters to the analysis until the added clusters no longer resulted in a significant reduction in the variation. At that point, the team determined that the optimal cluster amount had been reached.[13] Figure 2 shows an example of the SSE of summer seasonal load shapes for different numbers of clusters. The error begins to flatten out at six clusters, indicating that adding additional clusters does not explain that much more of the variation in the data.

[11] Summer load shapes used data during June – August and winter load shapes used data during December, January, and February. Shoulder season results were not displayed in the report as these seasons do not drive peak load as much as summer and winter.

[12] SSE is a statistical measure showing how much variation exists within a data set. The higher the SSE, the higher the variation in the data. In this context, the more clusters there are, the more accurately the cluster centers (average seasonal load pattern) will represent the data within that cluster (individual seasonal load pattern for each meter), reducing the SSE.

[13] If the number of clusters was too small, the load patterns would not be representative of the data within each cluster. On the other hand, if there were too many clusters, the analysis may result in many load shapes that look almost identical to each other for the purposes of demand management. Clusters that have very few members are also not meaningful for demand planning or customer opportunity identification purposes.

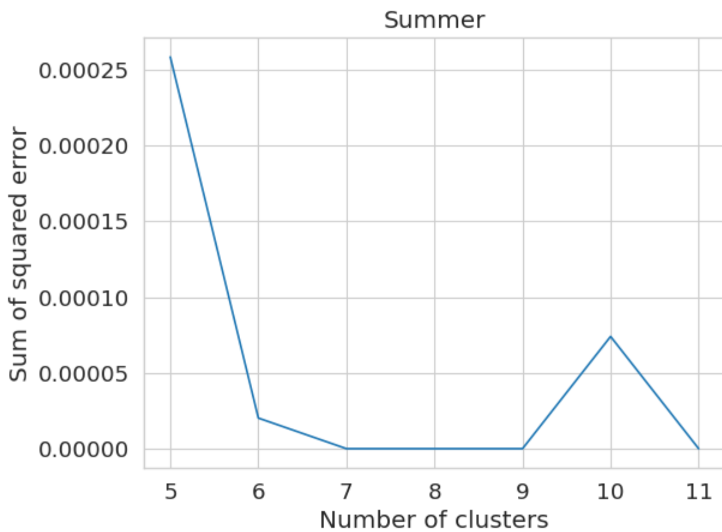

Figure 2 Within-cluster sum of squared error across a range of cluster amounts. As the number of clusters increases the variation in the data decreases.

After grouping the data into its relative cluster groups using the K-means algorithm, the research team labeled the cluster centers based on timing and width of cluster peaks. These labels help to identify features relevant to FLM and demand response applications.

Grouping the data into meaningful load patterns makes it possible to explore other interesting questions more deeply. For example:

- What kinds of buildings and building-uses comprise the identified load patterns?
- For a specific subset of the grid, for example a substation or town, what are the dominant load patterns?
- What proportion of meters fall into each load pattern?
- Which load patterns contribute most to grid constraints or greenhouse gas (GHG) emissions?
- What rates and technology would best incentivize non-optimal load patterns to become optimal load patterns for the grid?

Use Metrics

In addition to load shape clustering, this research team investigated a features-based approach to summarizing energy use data. Specific features of the load such as total use or winter peak demand can help identify use patterns which are well-suited for certain efficiency measures or programs. Table 1 summarizes the features that the team investigated. There is potential for more specific metrics based on program goals, but the table identifies those features most applicable to the projects analyzed in this study.

Table 1. Summary of AMI features and how they apply to different programs and technologies in the commercial sector.

Feature	Applicable programs	Applicable technologies	Definition	Units
Total use	EE	Lighting, refrigeration, HVAC	Total annual consumption	kWh
Summer on peak demand	EE/FLM	Lighting, refrigeration, HVAC	Average demand during summer peak periods	kW
Winter on peak demand	EE/FLM	Lighting, refrigeration, HVAC	Average demand during winter peak periods	kW
Discretionary energy use	EE	Lighting, HVAC	Non-baseload consumption (average daily consumption above the daily minimum)	kWh/day
Percent discretionary use	EE	Lighting, HVAC	Discretionary usage as percent of total consumption	%
Base use	EE	Lighting, refrigeration	Mean of daily minimum hourly demand *24	kWh/day
Percent base use	EE	Lighting, refrigeration	Base usage as a percent of total consumption	%
Upper demand range	EE/FLM	HVAC	97th percentile - 75th percentile identifies long demand tails	kW
Percent of energy on peak	FLM	HVAC	Ratio of average demand in hour 3 PM to 10 PM over average daily consumption	%
Ramp rate during peak hours	FLM	HVAC, lighting	Average ramp rate between 3 PM and 10 PM	kW
Cooling degree day (CDD) slope	EE / FLM	HVAC	The incremental change in energy use per day for every additional cooling degree day	kWh / day / CDD
Percent cooling degree day (CDD) slope	EE / FLM	HVAC	The incremental change in energy use per day for every additional cooling degree day as a percent of total consumption	%
Heating degree day (HDD) slope	EE / FLM	HVAC	The incremental change in energy use per day for every additional heating degree day	kWh / day / HDD
Percent heating degree day (HDD) slope	EE / FLM	HVAC	The incremental change in energy use per day for every additional heating degree day as a percent of total consumption	%

Efficiency Vermont can use these features to target a different component of a building's load, often at different times of the day or year. For example, baseload can capture equipment that is "always-on" such as refrigeration equipment that runs constantly. Lighting and HVAC, in contrast, tend to be utilized more heavily during occupied hours, and HVAC load tends to be driven by outdoor temperature.

The features-based analysis allows the team to assess specific features of energy use depending on program requirements and potential customer benefits. Outreach for a FLM program, for example, might focus on engaging customers based on their summer and winter peak demand since high peak loads may indicate larger grid impacts. A customer with a highly variable load might also be a good candidate for a FLM program with the implementation of utility controls and sequencing. Similarly, customers with long demand tails—a few hours with high demand—might have a few key hours per year when FLM could be highly effective in decreasing their demand charges.

The team can also use features-based analysis to identify generally high-energy customers who might provide more energy savings opportunity simply because they use more energy. But, instead of focusing only on the more obvious high-energy users, the analysis can also consider impact relative to total energy use and identify customers who could have large impacts and savings relative to their size.

Deploying the features-based approach could increase the reach of Efficiency Vermont's programs and engagement with customers, while continuing to deliver savings and grid benefits.

AMI Data for Customer Opportunity Identification

After running the analysis to summarize a customer's load patterns using AMI features, Efficiency Vermont can identify specific opportunities for intervention, based on feasibility, potential savings, and programmatic considerations.

To test the potential impacts of different features on at-the-meter savings after an efficiency upgrade, the research team conducted a retrospective case study using previously implemented custom commercial retrofit projects. The participant group consisted of commercial customers who had completed energy efficiency retrofit projects from 2018 to 2021. After filtering for utility data requirements and baseline uncertainty requirements, the analysis included 331 projects.[14]

[14] 12 months of pre- and post-period data are required following International Performance Measurement and Verification Protocol (IPMVP). ASHRAE-14 guidelines recommend fractional savings uncertainty must be within 50%. Industry experts, Lawrence Berkely National Laboratory recommend normalized mean bias error within 0.5%.

The team used AMI data to model pre-and post-savings for the retrofit projects,[15] and then used the AMI data (and the same pre-post assessment points) to compute savings for a matched comparison group which did not have intervention. The net savings of the interventions is estimated as the difference between participant group savings and comparison group savings.

The team used a nearest-neighbor matching algorithm to identify comparison group customers. The algorithm compared monthly use and average summer and winter on-peak demand and identified the group of customers whose use patterns most closely resembled those of the participant group. The comparison group customers had not received energy efficiency program incentives and were therefore assumed to have not completed any energy efficiency improvements. The team then used the matched comparison group to calculate net savings and adjusted the savings for biases in the baseline models and for confounding factors that may have influenced energy use outside of the intervention. For example, it was important to isolate the effects of COVID-19 on building energy use which were not captured in pre-post modeling alone.[16]

To estimate savings for the cohort of retrofit projects, the team computed average customer savings for the participant group and for the comparison group, and then subtracted the comparison group savings from the participant group savings. Table 2 shows these savings broken out by the primary measure type.[17]

Table 2. Pre-post savings by measure type for participant and comparison groups.

Projects with primary measure type	Number of projects	Participant average savings (kWh / day)	Comparison group average savings (kWh / day)	Net average savings (kWh / day)
Lighting	126	179	54	125
Refrigeration	31	260	122	138
HVAC	67	20	12	8
All projects	331	156	64	92

To verify that a features-based analysis can effectively identify customers with high potential savings, the team used a filtering mechanism. In a two-step process, the team first selected a certain percentage of the participant group that did not meet an established threshold for a

[15] Baselines for pre-post modeling use an hourly time-of-week and temperature model as implemented in the EEMeter open-source library http://docs.caltrack.org/en/latest/methods.html.

[16] Agnew, K., & Goldberg, M. (2013). *Chapter 8: Whole-Building Retrofit with Consumption Data Analysis Evaluation Protocol* . Retrieved from National Renewable Energy: https://www.nrel.gov/docs/fy17osti/68564.pdf

[17] Of the 331 projects analyzed, 42 had more than one measure type installed. To account for multiple measure types, the *primary* measure type is defined as the measure type that contributes to the most savings for the project. Only projects where the primary measure type contributed to more than 60% of the total project savings were broken out into groups for this analysis. Additional measure categories were present in the participant group but were not investigated explicitly. Positive savings within the comparison group could be contributed to modeling bias, random variation, or changes in energy usage due to COVID-19 shutdowns.

particular feature and then assessed savings for the remainder of the group. When a feature correlates with savings, removing a lower-performing segment from the total sample results in higher average savings for the remaining participant group. For example, if the team wanted to focus on the top 90 percent of the participant group, the threshold for outreach would have to exceed the tenth percentile of summer on-peak demand and customers who did not meet that 10 percent threshold, were removed from the sample. When the feature correlated to savings, the team observed that, by applying this method, higher average savings accrued for the remaining 90 percent of the participant group.

Figure 3 shows an example of how data can guide customer outreach. The graph shows heating degree day (HDD) and cooling degree day (CDD) slopes as a percentage of total daily energy use. These features could be used to identify customers for specialized outreach related to heating efficiency and weatherization.

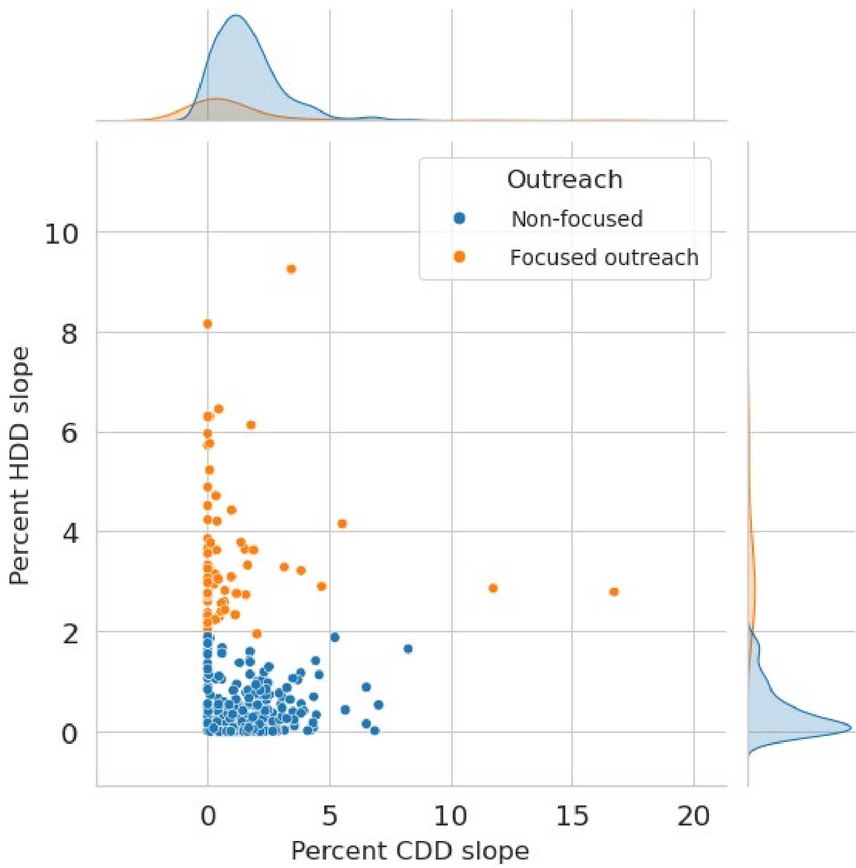

Figure 3. Customer cooling and heating degree day slopes as a percentage of total daily usage.

Orange dots in Figure 3 show customers with the highest 20 percent of heating degree day slopes. These customers may be candidates for weatherization or HVAC measures to increase heating efficiency. The y-axis shows the distribution of percent heating degree day slopes for customers most likely to benefit from outreach (orange) and those less likely to benefit (blue). The x-axis displays the distributions of percent cooling degree day slopes.

The team tested various features within each primary measure type—lighting, refrigeration, and HVAC—and investigated the effects of focused customer interventions on summer and winter peak demand savings.

Results

The featured-base analysis provided valuable results.

Load Shape Clustering

Figure 4 shows typical seasonal load patterns determined by the clustering algorithm using a random sample of 10,000 commercial GMP meters in Vermont. The load patterns represent the centers of each cluster. The vertical y-axis represents normalized demand over each hour. The normalized load is the ratio of average energy used for each hour over the average total daily energy use. This can be interpreted as the proportion of daily energy usage occurring at each hour. The analysis showed six load patterns typical of summer, and five load patterns typical of winter. Shown on the graphs below, these use patterns included increased use during *afternoon/evening,* increased use during *mid-day*, relatively *flat* use throughout the day, *morning and evening peak*, *night* loads, and increased use during *morning and afternoon*.

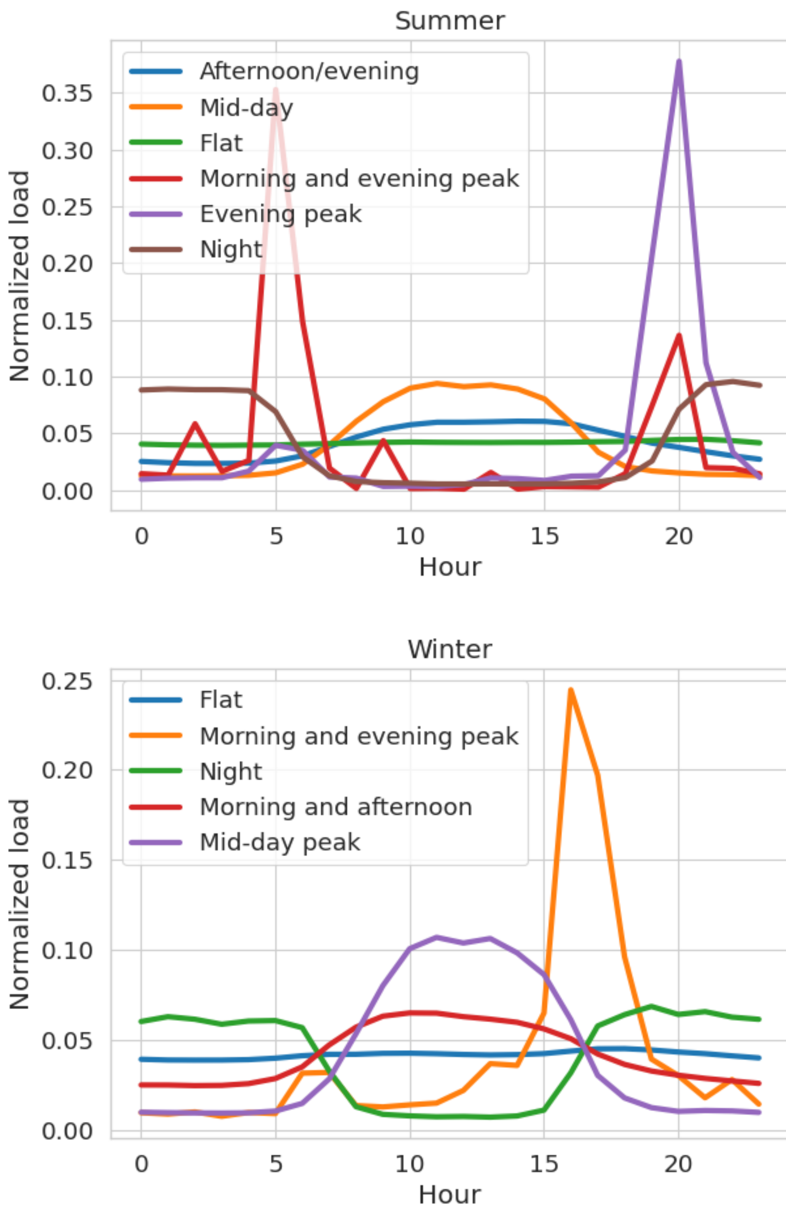

Figure 4. Summer (top) and winter (bottom) normalized load patterns.

Tables 3 and 4 provide additional information about each load pattern, showing the percentage of meters and the percent of total energy use within each pattern. This information can be used to identify customers with the most demand response potential. In the summer, customers with an *afternoon/evening* load pattern accounted for more than a third of the sample's total use. Because this load was coincident with evening summer peaks, these customers would be good candidates for demand response initiatives.

Table 3. Summer distribution of load patterns

Load pattern	% of sample within pattern	% of total usage within pattern
Flat	49.3	80.6
Afternoon/evening	30.3	16.2
Mid-day	12.9	3.0
Night	7.0	0.2
Evening peak	0.4	0.01
Morning and evening peak	0.1	0.00

Table 4. Winter distribution of load patterns

Load pattern	% of sample within pattern	% of total usage within pattern
Flat	61.1	86.3
Morning and afternoon	25.3	12.5
Night	8.0	0.3
Mid-day peak	5.0	0.8
Morning and evening peak	0.7	0.00

Flat load patterns comprised most commercial sector use in both summer and winter. There may be opportunities to shape flat load by incentivizing energy use at mid-day when solar is available and by reducing energy use during peak hours. Even though customers in the flat load pattern contributed less to steep grid peaks, they may be able to help reduce peak.

This type of analysis can be expanded to review flexible load potential across different regions and sectors of the grid.

The sample showed a very small percentage, less than 1 percent, of meters within the sample belong to the *morning and evening peak* and *evening peak* summer and winter load patterns. The summer load patterns are peaking from around 4 AM to 6 AM and 5 PM to 10 PM in summer. In winter, the *morning and evening peak* pattern is peaking from around 5 AM to 7 AM and 3 PM to 7 PM. Results like this, that is, load patterns with few associated meters, could indicate an outlier. The research team suspects that a larger sample size would show a clearer set of load patterns, but complications in scaling prevented further analysis during this iteration of the project. In the second iteration of this work, the team plans to incorporate a batched K-means algorithm to be able to analyze an increased sample of commercial metering data.[18] The team suspects that a larger sample size will create less noise in the load patterns which contain fewer meters.

[18] Mini Batch K-Means iterates over subsets of the data to optimize the K-Means algorithm, reducing computation resources required to identify clusters over large samples of data.

Identifying Customer Opportunities with AMI Data Analysis

Using load shape and AMI data to identify customer opportunities for efficiency measures has the potential to increase savings beyond what is currently achieved. The following sections demonstrate how opportunity identification techniques informed by AMI data analysis would affect the *average net savings per project* and the *average percent savings relative to load* under such a program model.

Figure 5 shows how using AMI data analysis to identify customers shifted the distribution of savings and increased the average net savings per project. In this example, the team categorized participants based on total annual use and focused on the 50 percent with the highest usage. Removing from the sample those participants with lower use nearly doubled the resulting average savings per day. For all participants, the average savings per project was 93 kWh/day. Focusing only on the top 50 percent of users resulted in an average savings per project of 180 kWh/day.

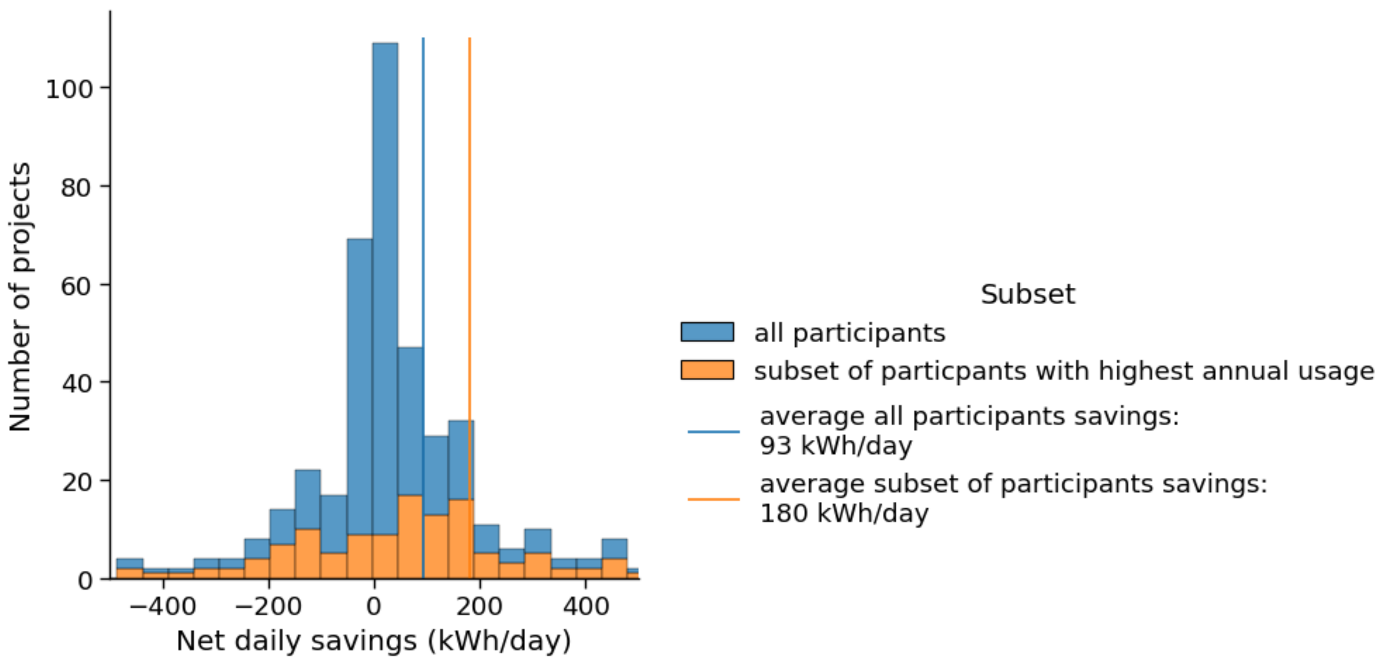

Figure 5. Average net daily savings (kWh/day) for all participants and for the 50 percent with the highest total annual use

Net savings per project

As shown in Figure 6, for nearly all order of magnitude features tested, the team's analysis showed that a customer opportunity identification approach based on AMI data analysis resulted in higher net daily savings per project.[19] In other words, focused program participation,

[19] Described in Table 1, magnitude features are the specific energy usage features—such as summer and winter on-peak demand, base demand, or cooling and heating degree day slopes—which can identify characteristics in the load that might make a particular project more suited to a certain efficiency measure.

resulted in higher average savings per project. The dotted grey line displays the average net savings results across all participants. Focusing on the top 40 percent of participants based on total usage, nearly doubled the average savings per project.

The team also tested for features which were normalized to reflect a percentage of total energy usage (bottom of Figure 6 and shown in Figure 7 and Figure 8) to identify customers with high relative use. Normalizing the features enables the research team to identify a smaller customer with load that is almost entirely cooling for specific cooling-related measures rather than selecting a larger customer with less cooling load relative to their total usage. These normalized feature-analyses showed mixed results. Selecting customers based on normalized *percent of base use* showed the highest positive correlation with average net daily savings. Other normalized features showed a negative correlation with this AMI-data informed customer opportunity identification. For example, the *percent of discretionary usage* correlated inversely with percent base use and showed a decrease in average net daily savings with stricter customer selection criteria. The results also showed a negative correlation between average net daily savings and *increased heating load*. The *Heating degree day (HDD) slope* correlated less strongly with increased average net daily savings for magnitude and normalized features. Fewer projects had significant increased load due to heating in part because heating often relies on other fuel sources. Given the small sample, the research team was unable to distinguish significant differences in the effectiveness of an AMI data informed approach for all other features except heating degree day slope.

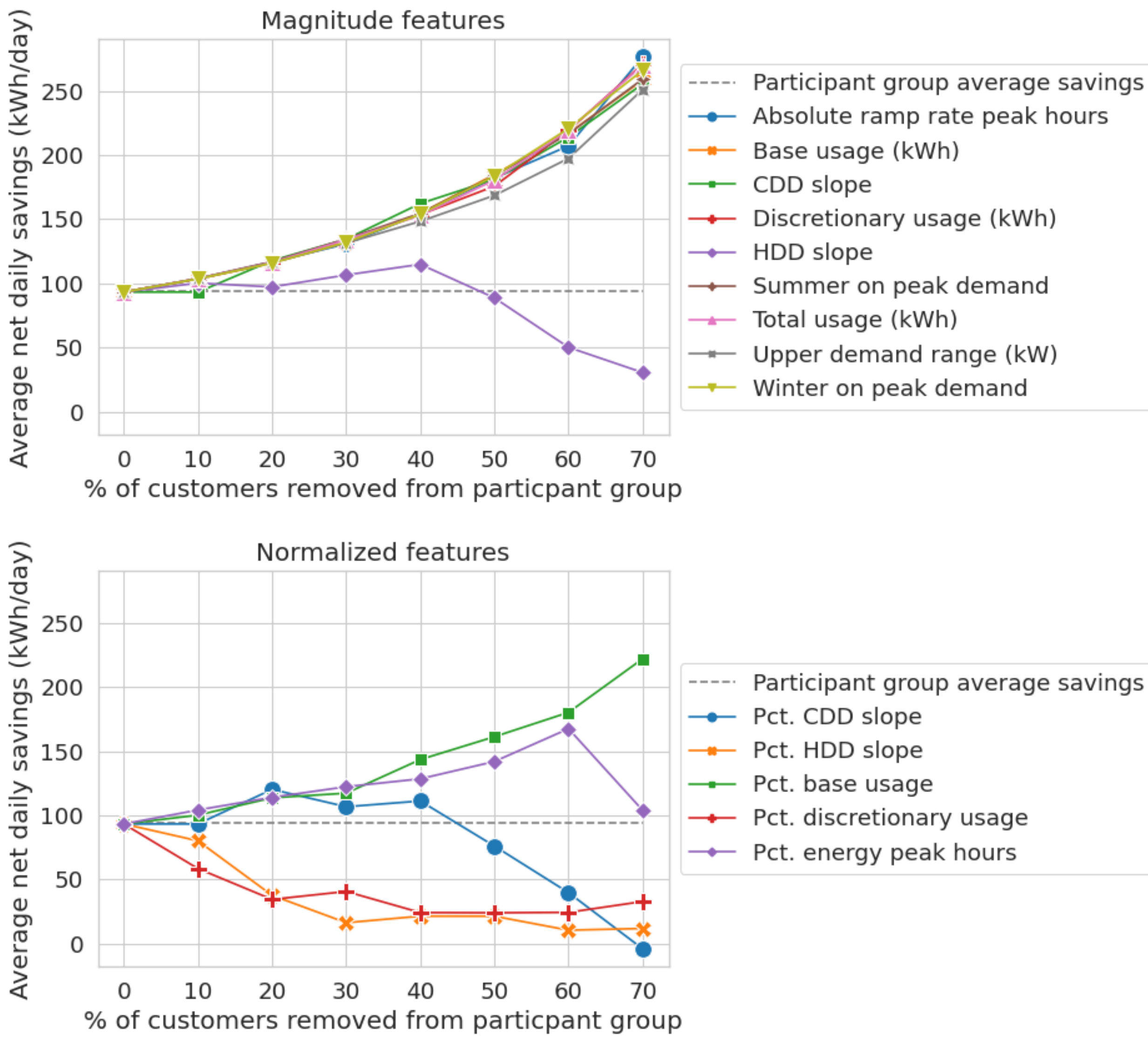

Figure 6. Average net daily savings (kWh/day) for a range of customer selection thresholds, and magnitude (top) and normalized (bottom) features

Figures 7 and 8 show results for winter and summer peak demand savings across various customer selection thresholds. Selecting the top 50 percent of projects based on total use resulted in a 1.5x increase in average summer peak demand reduction per project. Selecting the top 50 percent of projects based on total usage resulted in 2x the average winter peak demand savings.

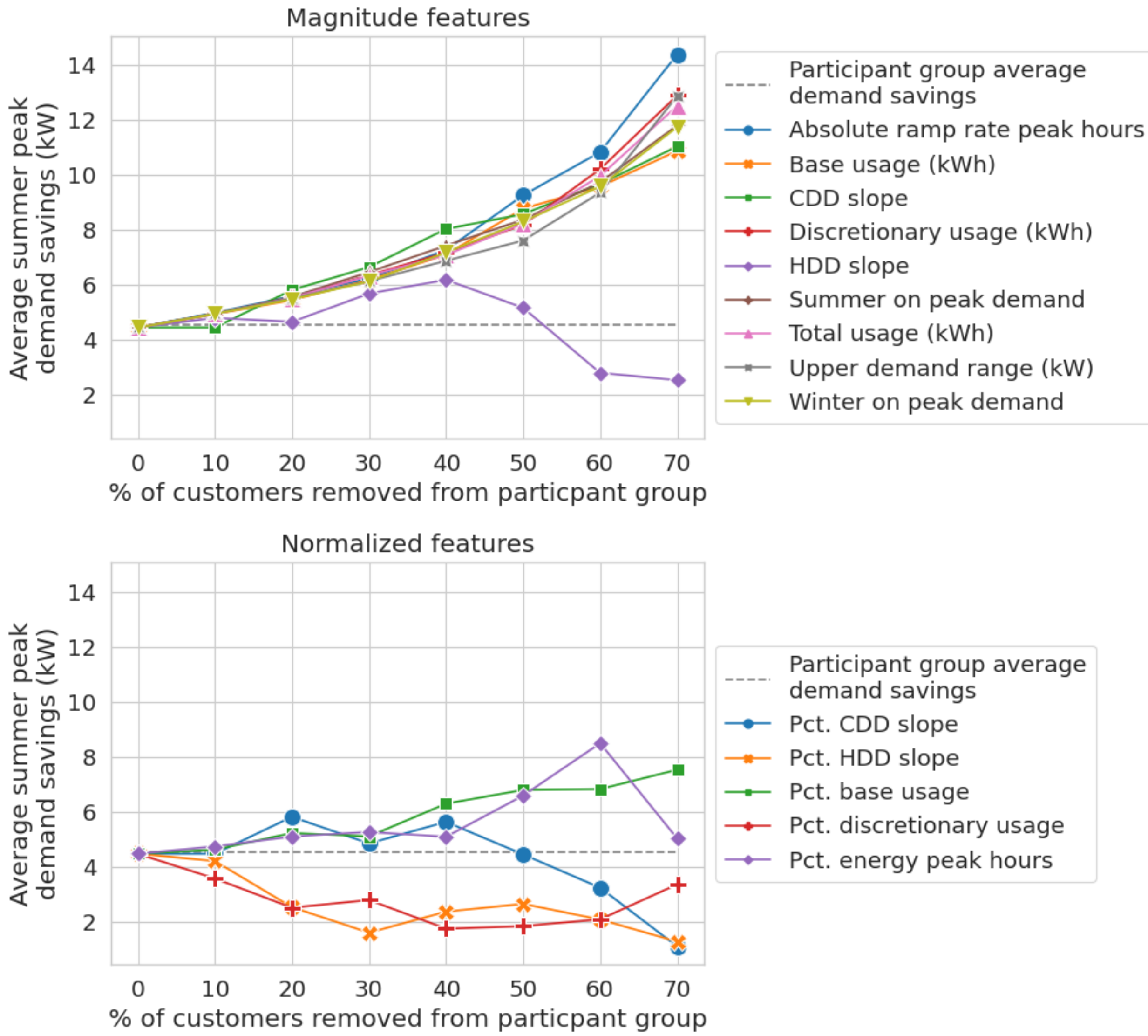

Figure 7. Average summer peak demand savings (kW) for a range of customer selection thresholds, and magnitude (top) and normalized (bottom) features

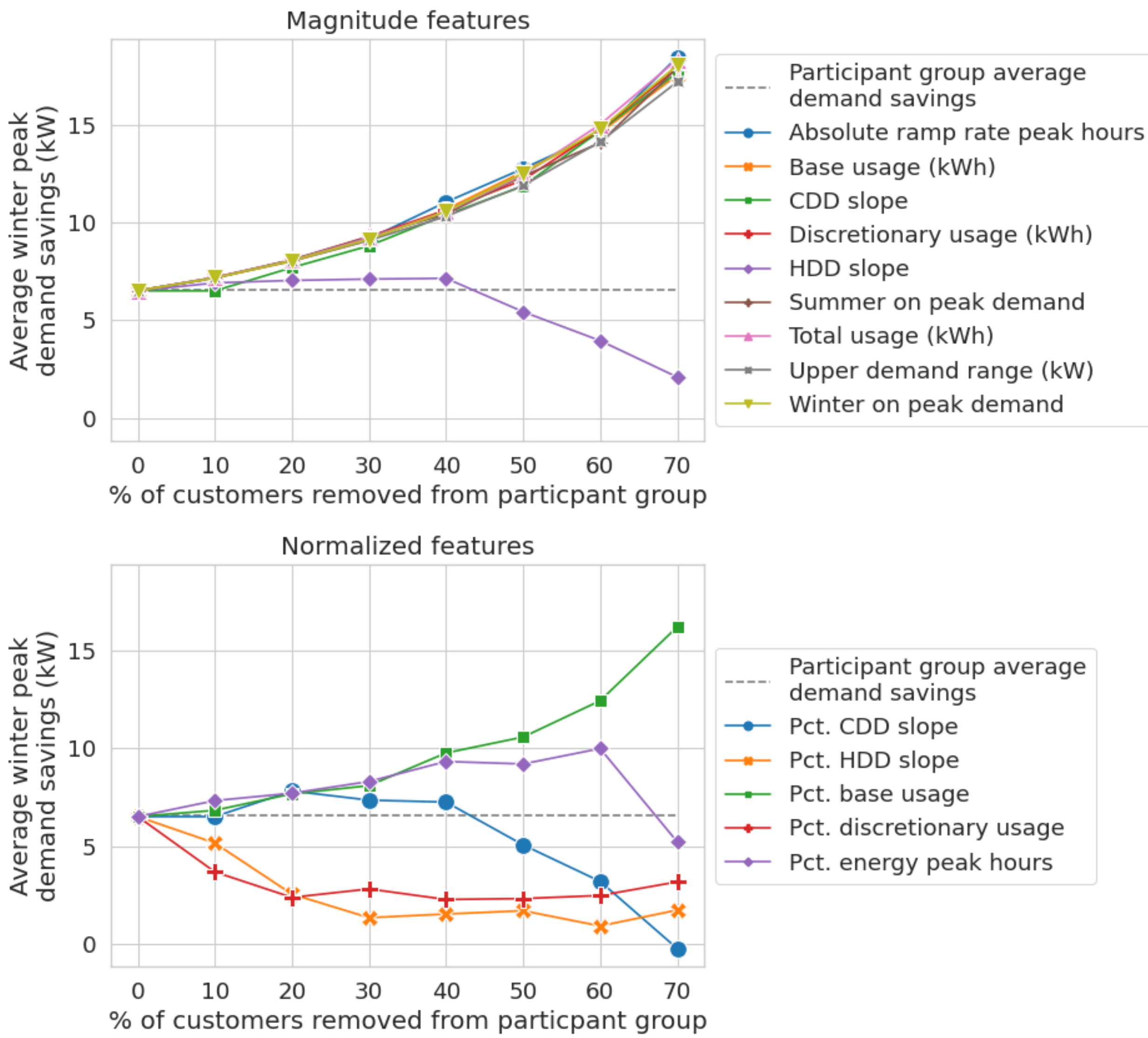

Figure 8. Average winter peak demand savings (kW) for a range of customer selection thresholds, and magnitude (top) and normalized (bottom) features

Percent savings per project

Figure 9 shows the average percent daily savings per project over a range of customer selection thresholds for each load feature. Most features showed a positive relationship with savings with stricter customer selection. Selecting the top 50 percent of projects based on total use increased average percent savings per project by a factor of 2. Even though the analysis selected customers based on total magnitude of energy use, the relative savings of the projects also increased, on average.

Normalized features showed less definitive results. With a customer selection threshold of 20 percent of total annual usage, *percent of energy during peak hours* and *cooling slope* showed increased average percent daily savings.

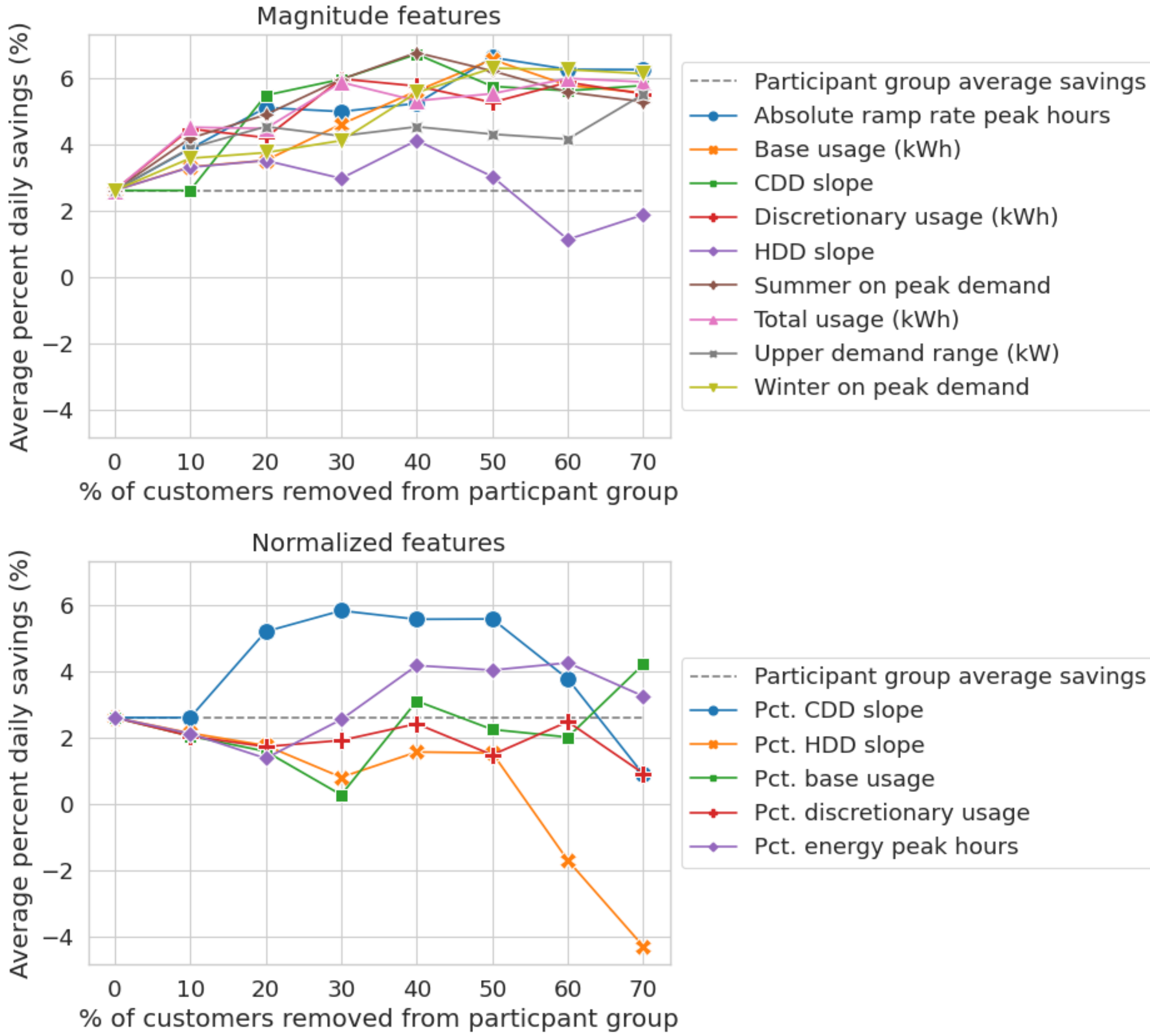

Figure 9. Average savings as a percent of baseline use over a range of customer selection thresholds, and magnitude (top) and normalized (bottom) features

Figures 10 and 11 show that customer selection affected average percent savings for *summer and winter peak demand*. Features associated with summer peak periods showed higher average percent savings with relatively low selection thresholds. *Summer on-peak demand* showed a 4x increase in average percent savings per project when selecting the top 80 percent of customers based on summer peak demand. Of the normalized features, *percent cooling degree day slope* showed the clearest positive trend with increased percent peak demand savings.

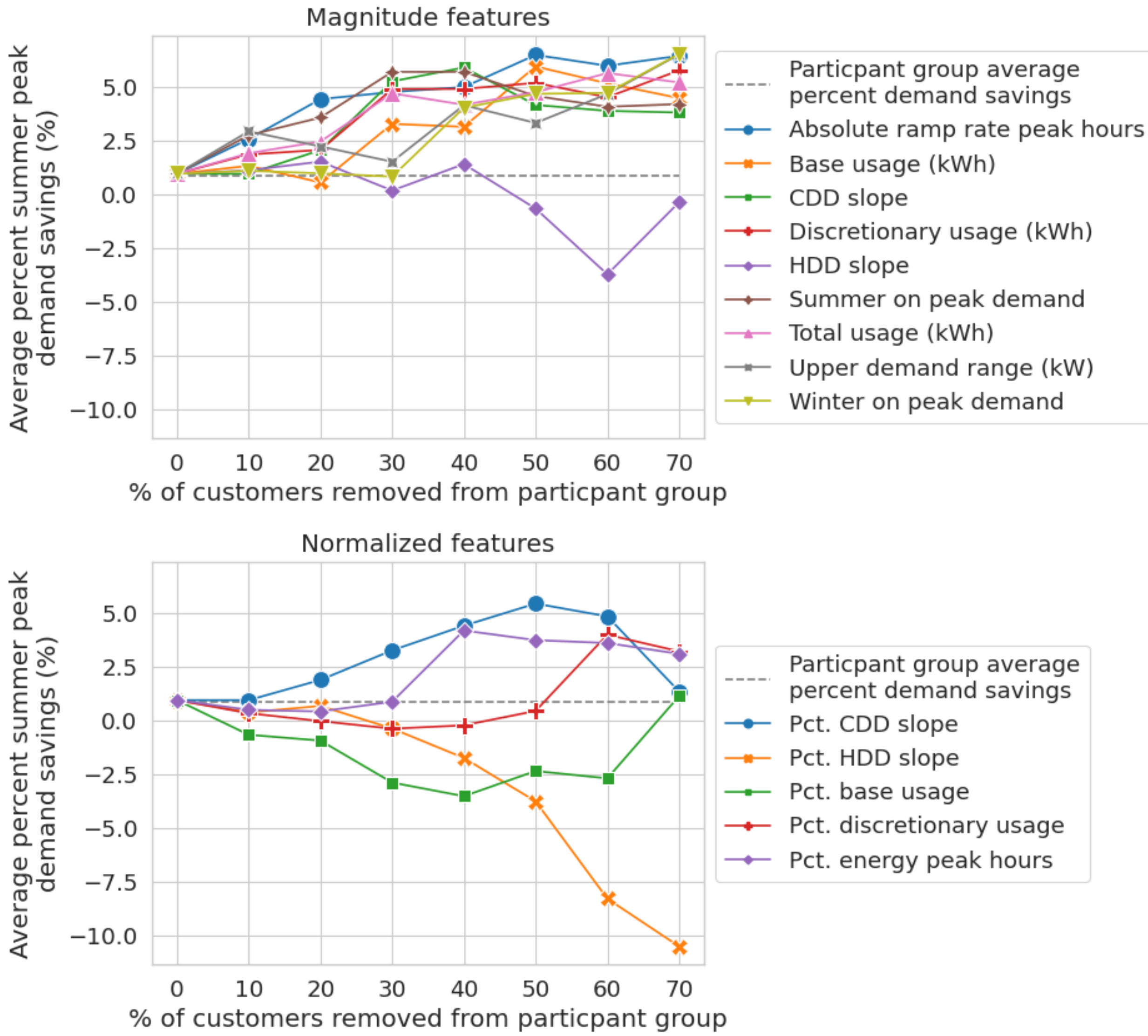

Figure 10. Average summer peak demand savings as a percent of baseline demand over a range of customer selection thresholds, and magnitude (top) and normalized (bottom) features

Selecting the top 50 percent of projects based on average winter peak demand showed around 2x the average percent of winter peak demand reduction. The normalized features did not show clear relationships with increased savings.

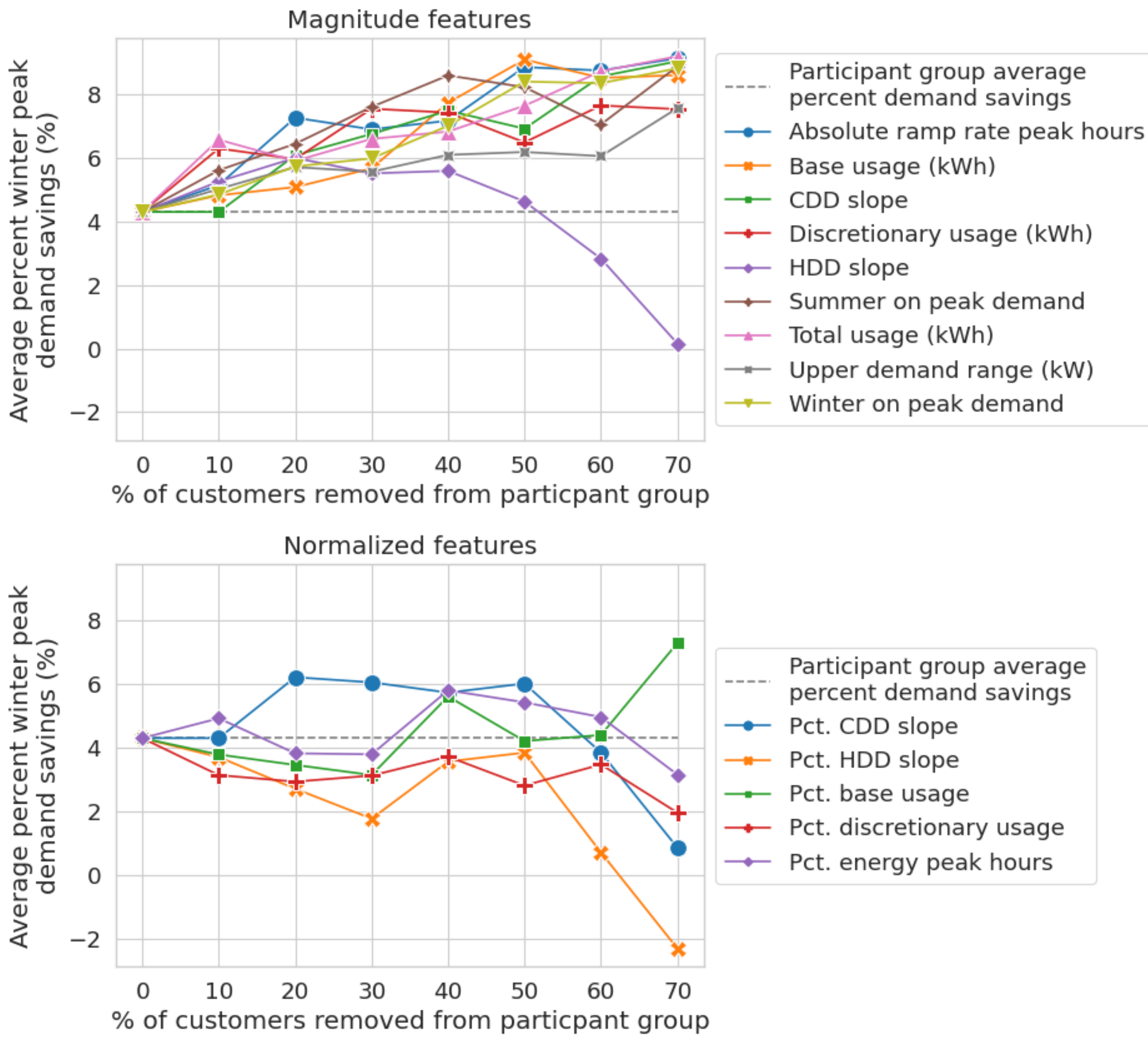

Figure 11. Average percent winter peak demand savings as a percent of baseline demand over a range of customer selection thresholds, and magnitude (top) and normalized (bottom) features

Measure Groups

In addition to analyses based on magnitude and normalized measures, the research team also grouped results on the basis of three primary categories: HVAC, lighting, and refrigeration. Measure descriptions that did not fit into these categories were not included in the following analysis. As Table 2 shows, sample sizes for each measure group were relatively small. This is an important factor to keep in mind since outliers can have disproportionate effects on results drawn from small data sets.

Even taking the small sample size into consideration, the results in Figure 12 showed that *cooling degree day slope* correlated much more strongly with increased HVAC savings than with lighting savings. Customer selection on the basis of *summer peak demand* also correlated with higher HVAC savings. These relationships are expected as electric HVAC is often used for cooling during summer.

Compared to other features, *discretionary use* and *base use* showed the greatest average daily savings as a result of lighting. Lighting can be occupancy driven and be highly correlated with discretionary load. Lighting can also be an always-on load and would be better identified with a base load feature. Customer selection for lighting efficiency projects might depend on the type of business. For example, office buildings might be more effectively selected in an analysis of discretionary load. If looking at warehouses or end uses where lights are always on, a feature such as percent of base load relative to total usage could be more effective.

Customer selection for *refrigeration measures* showed increased average daily savings which improved further with additional selections on the basis of all other features except heating slope.

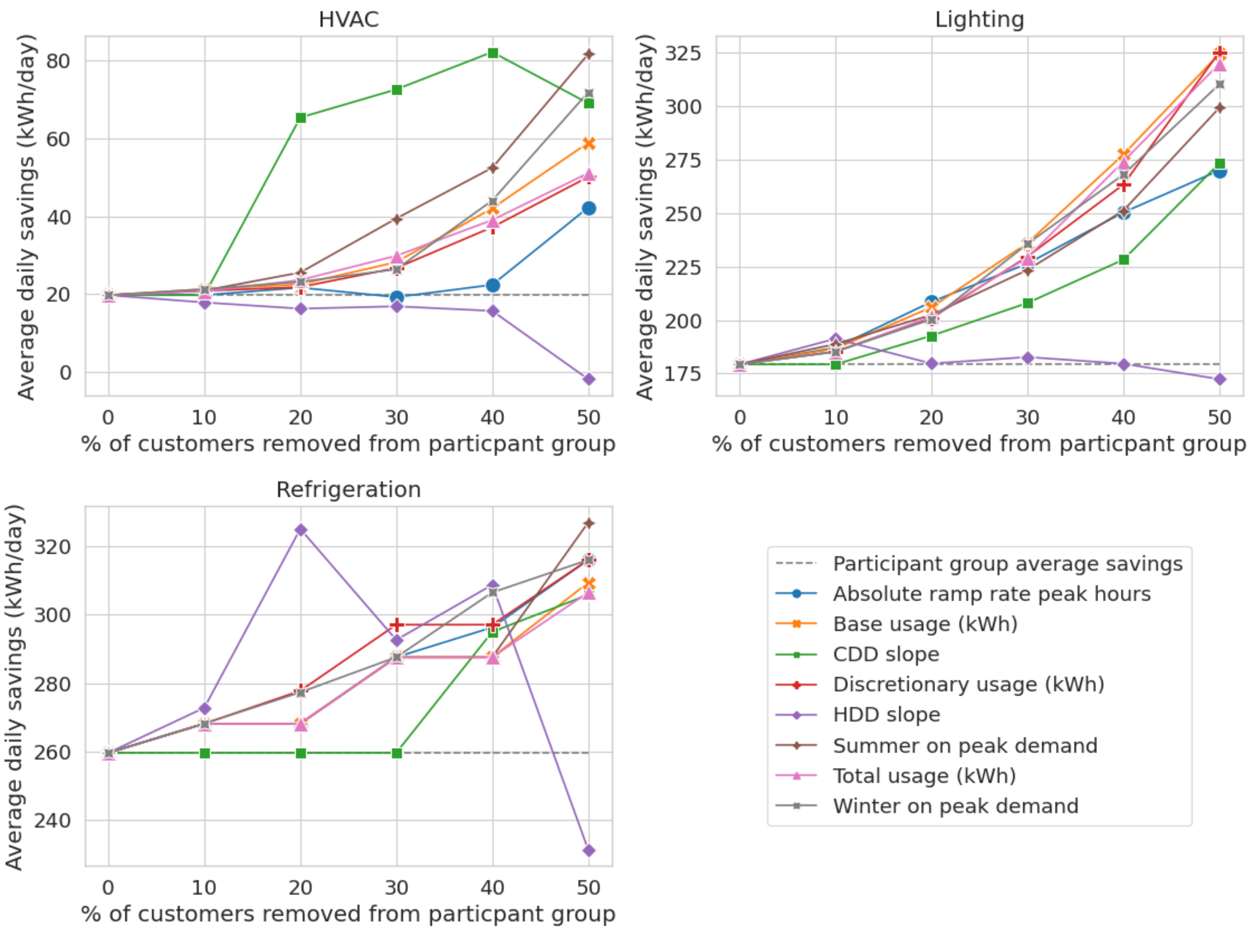

Figure 12. Average daily savings (kWh/day) over a range of customer selection thresholds and measure groups

Discussion

The research team's analysis shows that load shape segmentation presents an opportunity to gain a deeper understanding of the time-of-use patterns that exist in a sector or region. The results help illuminate potential areas for FLM and indicate what technology, incentives, or rates would best optimize these patterns. Opportunity exists to further scale the classification of load patterns which may improve the patterns that can currently be identified in the data. There is also opportunity to conduct this analysis for specific regions on the grid, such as a substation that is particularly constrained during peaks.

Testing approaches to customer opportunity identification with retrofit projects shows that load features can be used to identify projects with more potential for savings and grid impact. Given the small sample sizes in this analysis, the magnitude of savings increases could not be guaranteed, but these results deepen Efficiency Vermont's understanding of how particular features relate to different types of projects and customers. Even with smaller sample sizes, this analysis demonstrates how AMI data-informed customer selection could be implemented and lays the groundwork for expanding Efficiency Vermont's internal analytics infrastructure to begin using this approach in practice.

While this study tested how a new approach to customer selection could affect program energy savings, this approach could be expanded to support opportunity identification and inform outreach for FLM. Customers with high heating and cooling loads may have the ability to be flexible through thermal storage and controls. Customers with high coincidence during peak times are more likely to have increased demand impacts from load management. On the other hand, a steep negative ramp rate during peak hours could indicate that the customer may already be closing their facility for the day, and may not have additional load flexibility during these hours.

Using data analytics to inform program design and customer outreach, can enable Efficiency Vermont not only to reach customers with whom it may not already be connecting, but also to connect with those customers in a more personalized way based on the customers' energy needs. The opportunity also exists to overlay energy use data with other customer demographic and geographic data to ensure Efficiency Vermont is reaching customers who will benefit the most from its programs. For example, if Efficiency Vermont wanted to verify that Vermont schools follow a set-back schedule during non-occupied hours, it could use AMI features to identify schools that did not ramp down appropriately and help them implement controls to reduce energy use. The goals of the program and population of interest will largely drive what features are most useful.

Integration of this work into Efficiency Vermont's program process could look like building a self-service tool that allows program staff to filter by various features that relate to the goals of a program. A ranked list of identified customers could provide further focus and intention around which customers Efficiency Vermont reaches out to for enrollment in programs.

Conclusion

The increasingly urgent demands of climate change compel Efficiency Vermont to evolve its programs and work in new ways to help the State reach its carbon reduction goals. Employing data analytics to extend and focus its customer opportunity identification, could increase Efficiency Vermont's impact on greenhouse gas reductions. As Efficiency Vermont's portfolio continues to promote beneficial electrification, the timing of energy use will only become more critical. To better serve its customers, Efficiency Vermont can use data analytics to take advantage of the high-frequency meter data that is available for most of Vermont.

The results of this paper demonstrate that energy use patterns and features can identify commercial customers who would benefit from increased savings and deliver grid impacts if they employed certain measures. Building off these results, Efficiency Vermont will investigate how these methods can be applied to the residential sector, implemented equitably and effectively into its programs and processes, and employed to improve its customer opportunity identification.