Data-driven Characterization Of Cooling Needs In A Portfolio Of Co-located Commercial Buildings

Aqsa Naeem[1], Sally M. Benson[1], and Jacques A. de Chalendar[1]

Summary

Commercial buildings have growing cooling demands. Greater energy efficiency and flexibility are needed, especially in existing buildings that have a slow stock-turnover. We collect, analyze, and release a dataset on chilled water use from 119 co-located buildings in a warm-summer Mediterranean climate. Factoring out geography-driven differences, we observe a strong heterogeneity within and across different building types. The average estimated base cooling intensity at 18°C varies from 0.50 to 4.4 MJ/m^2/day across buildings, with the highest loads in healthcare and the lowest in residences. We find that simple, interpretable regressions can be used to model cooling load and provide data-driven benchmarks for comparing building performance. Over five years, these regressions explain over 70% of variance for buildings that collectively represent 85-94% of the portfolio's overall cooling load. Consumption increases by 7.6-9.8% for every 1°C increase in mean daily outside temperature and drops on weekends, by up to 27% in offices.

Keywords: Commercial building cooling, Building energy efficiency and flexibility, Data-driven urban energy models, Building energy end-use, Energy performance monitoring

Introduction

Deep energy efficiency and flexibility are needed in the commercial sector, both to ease the challenges from electrification and to adapt to a changing climate. Existing buildings are especially important because stock turnover is slow, but they remain poorly measured environments. Increasingly available data from distributed sensors can help unlock efficiency and flexibility opportunities through high quality data benchmarks and analytics.

Climate goals drive ambitious electrification targets and therefore require the conversion of much of our current energy infrastructure [1, 2]. These growing infrastructure needs, including for

[1]Stanford University, Department of Energy Science & Engineering

long-duration storage [3], are likely to be amplified by climate change. Requirements are typically based on peak rather than average loads. There is already strong evidence to confirm that along with mean temperatures, the upper tails of the temperature distribution are warming [4, 5]. Energy efficiency can help offset the growth in average demand, while flexibility in demand mitigates the peak-hour and peak-day challenges [6].

In the United States (US), commercial buildings represent 35% of 2021 electricity sales [7]. They include offices, retail sites, supermarkets, schools, laboratories, and data centers. Space cooling, refrigeration and ventilation represent half of the purchased electricity [8].

While significant effort has been invested into designing new buildings that will perform better, most existing buildings will continue to be used for a long time. The median year of construction for US commercial buildings is 1981 and only 29% of total floorspace was built after 2000 [9]. Better understanding and characterizing the cooling needs of existing buildings is critical to quickly decarbonize the current stock.

Physics-based simulation models are currently the tool of choice for assessing different flexibility and efficiency options at the building [10, 11] and urban [12, 13] scales. For example, simulations promise that upgraded controls could reduce energy requirements by 23-30% in half of the US commercial building stock [14]; that aggressive efficiency measures, electrification, and high renewable energy penetration can reduce 2050 carbon emissions from buildings by 72%-78% relative to 2005 levels [2]; that energy efficiency and flexibility measures can avoid 800 TWh of annual electricity use and 208 GW of daily peak load by 2050 [15], or that 10-20% energy can be saved from increasing temperature set points by 2.2°C in a range of climate zones [16].

But comparisons of simulations to real world measurements show they better capture collective behavior than the often idiosyncratic behavior of commercial buildings [17], including during controlled experiments [18, 19]. Simulation models are expensive to set up. The state of the art in model calibration relies on design parameters and whole building electricity consumption [20, 21]. Models are largely used during building design and construction, seldom for operations [22]. Building behavior also changes in time. The impacts of retrofits or demand management measures will likely vary from building to building. The behavior of real-world commercial buildings is complex and determined by the interaction of the weather, occupancy, building physical characteristics, control systems and set points. For example, recent in-the-wild flexibility experiments found strong

heterogeneity in the response of daily chilled water use to a 1.1°C cooling setpoint adjustment [23].

Increasingly available on-the-ground measurements [24, 25] bring tremendous added value to simulation approaches for analyzing energy consumption and efficiency in buildings. They could greatly enhance the current capabilities of modeling tools [26, 27].

Results

We collect, analyze and release a large new data set on chilled water use from 119 co-located buildings in a warm-summer Mediterranean climate. Five years of real-world measurements of chilled water use are analyzed to explore how the summertime (May - September) cooling consumption varies across this portfolio of commercial buildings, located on a university campus. We leverage data-driven regression models to characterize cooling requirements both at the individual building and portfolio scales. These models are made to adapt across different buildings by identifying the features that influence the cooling consumption through a three-step feature curation process described in the Supplemental Information (SI). We repeat this process every year to track the cooling consumption behavior of buildings over time. Modeling and analysis focuses on three main questions.

1. *How well can commercial cooling consumption be modeled by weather and occupancy-related predictors in a simple but adaptive regression framework?* Overall, cooling loads are well explained by simple, interpretable regression models. These models explain over 70% of the variance in chilled water use for buildings that, across time, collectively represent 85 to 94% of the total cooling load of the building portfolio. The models disaggregate cooling load into two components, a fixed base load that reflects the system design and a variable component that is operations and weather dependent.

2. *What are the drivers of cooling consumption in commercial buildings?* Over five years, the portfolio's base load intensity, which we define as the cooling use intensity at 18°C (64.4°F) is 2.6-3.0 MJ/m^2/day. We identify Outside Air Temperature (OAT) as the most important driver of variance in cooling water consumption. An increase of 1°C (1.8°F) in the mean daily OAT increases the cooling consumption intensity by 7.6-9.8%. A weekend indicator variable is the next most significant feature, especially in small-sized buildings. Overall

consumption reduces by 7-8% on the weekends. Solar radiation, wind, and relative humidity (RH) were minor drivers for cooling load. In the warm-summer Mediterranean climate we consider, RH has the least influence on the cooling load except for one year, where it helped in characterizing an additional 3% of variance in cooling load.

3. *How heterogeneous are the cooling requirements for different commercial buildings located in the same geographical region, and what are the major drivers of these differences?* There is strong heterogeneity across individual buildings, both in base load intensity and in sensitivity to the drivers that were identified. Over the years, the average estimated base load intensities across buildings range from 0.50 MJ/m^2/day (standard deviation: $\pm$0.05 MJ/m^2/day) to 4.4 MJ/m^2/day ($\pm$0.75 MJ/m^2/day). The healthcare buildings consume the most, and the multi-unit residential buildings consume the least. The average temperature sensitivity ranges from +6.8% ($\pm$1.3%) to +18% ($\pm$5.5%) per degree Celsius. Office buildings reduce their cooling consumption by an average of 26.7% ($\pm$1.9%) on the weekends, while healthcare loads are simultaneously reduced by only 2% ($\pm$0.24%).

A large new dataset on chilled water use from 119 co-located commercial buildings

We collect and release data from a varied sample of 119 co-located commercial buildings. Our dataset includes floorspace, building category, and daily cooling consumption between May and September (Figure 1). A little less than half of the buildings are smaller-sized Offices and Classrooms buildings, 37 buildings of varying sizes house laboratories (Labs), 12 buildings offer healthcare facilities, while the remaining 17 buildings are small-sized buildings that are used for various purposes.

Figure 1C shows the variability in cooling consumption of buildings in the three categories that have the highest consumption. In all three plots, the curves with a higher color saturation indicate the overall category-level profiles, which we obtain by aggregating the consumption and normalizing it by total area of buildings within each category. Comparing the categorical profiles, the Offices and Classrooms buildings have the lowest consumption, while laboratories and healthcare buildings have a higher cooling load intensity due to more stringent HVAC design regulations and ventilation requirements for public health and safety. Additionally, these buildings have process equipment (e.g., computed tomography, X-ray, sterilizers etc.), which increases their chilled water

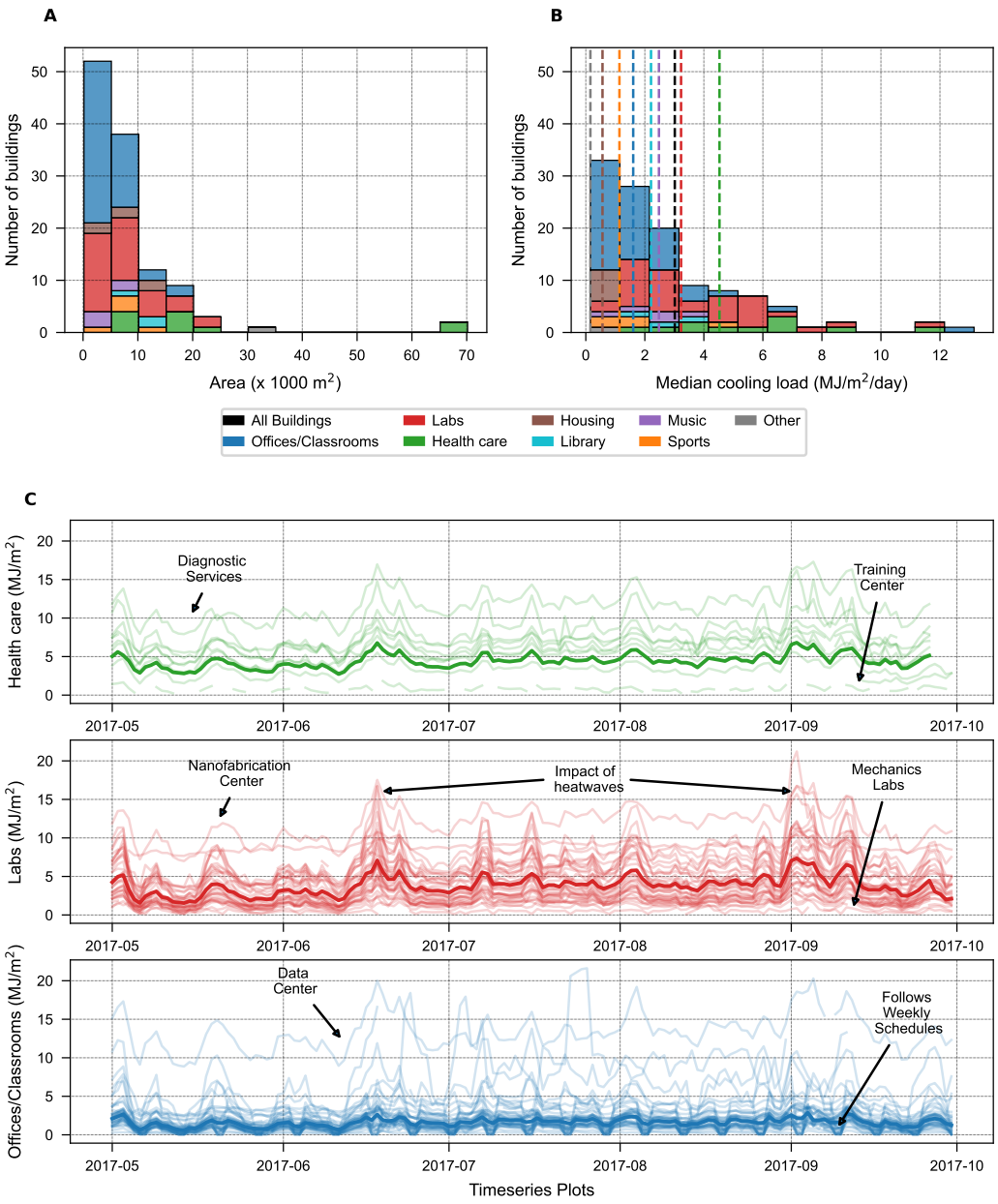

Figure 1: Characteristics of 119 co-located heterogenous commercial buildings. (A) Histogram for floorspace by building category. (B) Histogram for the median daily cooling consumption (in MJ/m^2/day) by buildings in each category during the summer season (May - September). (C) Time series plot of cooling consumption intensity of the top three categories in the portfolio that have the highest consumption.

consumption. Further, the cooling loads in laboratories and healthcare buildings are more impacted by heat waves that occurred in June and September 2017 than the Offices and Classrooms

buildings due to more frequent air changes.

Even within a building category, there is a wide range of cooling intensity (MJ/m^2/day) due to differences in the equipment these buildings house and the type of operations carried out in them. For instance, hospitals have a large proportion of floor space reserved for operating rooms, diagnostic facilities, isolation spaces and 24-hour medical care that may have more critical requirements (e.g., minimum ventilation, maximum indoor temperature, supply and return air flows and locations etc.) than the healthcare buildings that have a greater number of outpatient clinics and offices of medical practitioners. In Figure 1C, the building at the top has a diagnostic facility, while the one with the lowest consumption is an educational training building with some space reserved for a psychiatry clinic. Similarly, the Offices and Classrooms building that has the highest consumption houses data center and information technology offices, which is why its cooling consumption is higher than other buildings in this category. In the Laboratories category, the average daily cooling consumption varies from 0.96 MJ/m^2/day (mechanics laboratories) to approximately 12 MJ/m^2/day (nanofabrication facilities).

A simple regression framework explains a large fraction of the variance in commercial building cooling loads

We develop adaptive, data-driven explanatory models using the measured building data. We model cooling consumption as a function of OAT and a set $\mathscr{F}$ of other weather and occupancy-related features using the regression equation

$$\log y_{i,t} = \alpha_i + \beta_i X_{OAT,t} + \sum_{k \in \mathscr{F}} \gamma_{i,k} X_{k,t} + \epsilon_{i,t}, \quad \forall i = 1 \cdots N, \tag{1}$$

where $y_{i,t}$ is the cooling demand intensity of building i in MJ/m^2/day on day t, $X_{OAT,t}$ is OAT in degrees Celsius, and $X_{k,t}$ is a feature in the set $\mathscr{F}$, which contains solar radiation (denoted by, sun_rad), wind speed, RH, Tdew, and a weekend indicator variable. The weather data we consider are presented in Figure S3. The unknown model parameters to be estimated include α_i, β_i, and γ_i, while $\epsilon_{i,t}$ represents an error term for building i. The temperature sensitivity, weekend sensitivity, and base load intensity are calculated from β_i, γ_i, and α_i as given in equations (2), (3), and (4)

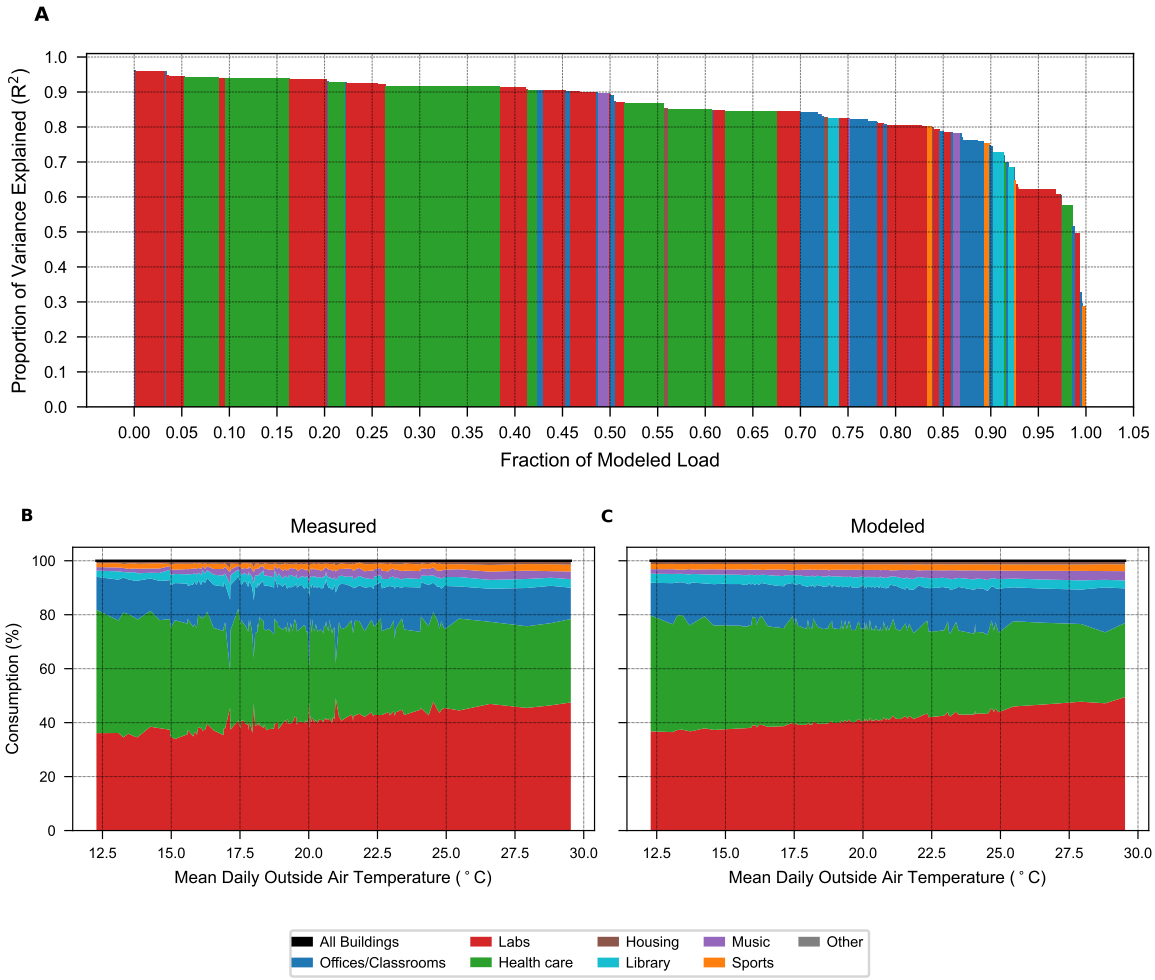

Figure 2: (A) Proportion of variance explained as described by the coefficient of determination (R^2), plotted against the fraction of modeled load for year 2017. The vertical bars indicate different buildings and their colors represent the type of buildings. The width of each bar indicates the fraction of total cooling load that the building consumes. Stack plot showing the percentage change in cooling consumption of buildings in the portfolio with mean daily outside air temperature for the (B) Measured cooling load and (C) Modeled cooling load. The fluctuations observed in the modeled cooling load are attributed to the influence of different features in the feature set, particularly weekends. See Figure S8 for performance in other years and Figure S9 for comparison with linear models.

respectively,

$$\Delta y_i^{\beta} = (e^{\beta_i} - 1)100\%, \quad \forall i = 1 \cdots N, \tag{2}$$

$$\Delta y_i^{\gamma} = (e^{\gamma_i} - 1)100\%, \quad \forall i = 1 \cdots N, \tag{3}$$

$$y_i = e^{(\alpha_i + 18\beta_i + \sum_{k \in \mathscr{F}} \gamma_{i,k} \bar{X}_k)}, \quad \forall i = 1 \cdots N. \tag{4}$$

Δy_i^{β} and Δy_i^{γ} represent the percentage change in cooling demand intensity of building i due to a unit change in OAT and weekend indicator, respectively. Further, y_i denotes the base load intensity of building i and $\bar{X}_k$ represents the average daily values of the k^{th} variable. Models are estimated using the standard Ordinary Least Squares method, implemented in the Python package `statsmodels` [28]. A model is selected and trained for each building in each year (Experimental Procedures).

To evaluate the overall explanatory power of the models that were selected, we rank the buildings in order of descending R^2 and determine the total observed cooling load for each building, $Y_i = \sum_t y_{i,t}$. Figure 2A shows the R^2 achieved for each building i as a function of the fraction of cumulative load up to that building, $\sum_{j \leq i} Y_j / \sum_k Y_k$. The bars correspond to different buildings on which the model was trained in 2017, while the width of each bar corresponds to the fraction of total load that the building represents. The bar colors indicate the building category. The trained models can explain at least 70% of the variance in cooling load for 92% of the total cooling load in 2017 (Figure 2A). The models provide higher quality predictions for larger buildings, e.g. the healthcare buildings and several laboratories, than for smaller buildings such as Offices and Classrooms. Similar performance data are shown for the other years in our dataset (Figure S8). However, if data from all five years are used to train the models, only 69% of the modeled cooling load has an R^2 of greater than 0.7. Shifts in consumption patterns are found in some of the buildings, which is reflected in year-to-year changes to the model parameters and poor matches to the five-year data set.

We compute aggregate predictions for the entire portfolio over the range of outside temperatures (Figures 2B and 2C). The healthcare buildings and laboratories make up about 80% of the total cooling load. Due to a higher temperature sensitivity, the laboratories consume a greater percentage of cooling load at higher temperatures. As shown, the modeled cooling load closely matches the variation observed in the measured load at the portfolio scale. However, at the level of individual buildings, the model's performance may show greater variability.

The models for the portfolio (All Buildings) explain 85-94% of the variance in cooling load over the five years we consider, while the proportion of variance explained by buildings in different categories is tabulated in Table S2. The average training and testing performance scores that are achieved during the model selection process are discussed in the Experimental Procedures. It

is also important to highlight that many of these models show a superior performance at lower temperatures because that is where most of the data points are concentrated. This issue could be resolved by using weighted linear regression, but was not done here.

Identifying the drivers of cooling load

For the following analysis, we focus on buildings where the R^2 is over 70%, the threshold over which we consider a building's cooling loads to be well-explained. Mean daily OAT is the main driver for cooling consumption for the building portfolio (Figure 3A-B). Alone, it explains over 70% of the variance in cooling load for buildings that represent 69-84% of portfolio-level cooling loads. However, while temperature explains most of the cooling load variance at the portfolio scale, only 40 to 65% of the buildings are well explained by temperature alone. This suggests that buildings with smaller cooling loads are more influenced by additional factors. For these buildings, adding more features improves the model performance, bringing the "well-matched" fraction of the portfolio to 85-95% of the cooling load and 71-83% of the buildings.

The weekend indicator is the second most important feature, followed by relative humidity (RH, Figure 3C). Including the remaining features such as the dew point temperature (Tdew), wind, and solar radiation (sun_rad) does not improve performance. The weekend indicator remained an important feature across the years (Figure S12). The RH was not a significant variable across all years. Instead, in some years, Tdew had more influence on the model performance, while in other years, solar radiation and wind play a more prominent role in characterizing the cooling loads.

Statistical decomposition of commercial building cooling loads

The regression models provide a statistical decomposition of cooling loads that allows characterization of the heterogeneity across buildings (Figure 4). For most buildings, the three main components that we obtain from each model are the base load intensity in equation (4), temperature sensitivity in equation (2), and the weekend sensitivity in equation (3). We show this decomposition for the year 2017 in Figure 4.

We define base load intensity as the expected value of cooling demand at 18°C (64.4°F). The average portfolio's base load intensity in 2017 is 2.63 MJ/m^2/day; across years it is 2.6-3.0 MJ/m^2/day. The average base load intensity varies 8-fold across different building types, indicating a high degree of heterogeneity between buildings (Figure 4A). Healthcare buildings are

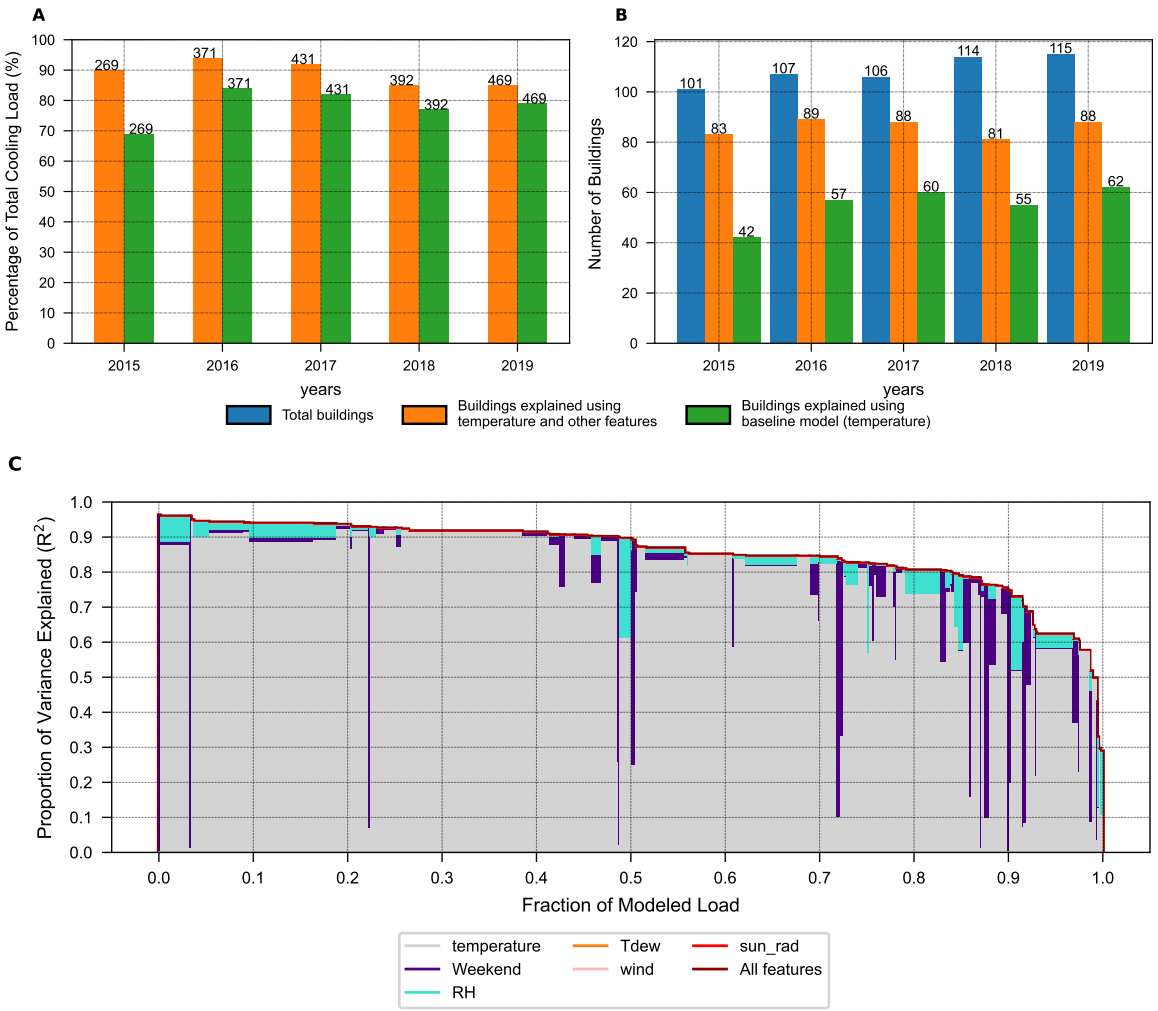

Figure 3: (A) Quantifying the percentage of modeled load that can be explained with an R^2 of more than 70% using the baseline model and the model that uses multiple features from the feature set. The values on bar tops indicate the total cooling load (in TJ) of buildings. (B) The bars indicate the number of buildings that correspond to the percentage of modeled load in (A). Note that different number and type of buildings contribute to the modeled load. The bigger difference in the green and orange bars suggests that the improvement in model performance occurs mainly due to small-sized (having smaller load) buildings which are better explained with more features. (B) Contribution of different features to overall explanatory power in 2017. The baseline model, which only uses outside air temperature as a predictor, helps explain about 82% of the modeled load in 88 buildings with an R^2 of over 70%. However, without the weekend indicator and RH, the model in some buildings performs badly (which is why we see the blue and turquoise dips). Figure S12 shows the contribution of different features in different years.

the most cooling-intensive (3.93 MJ/m^2/day), while residential buildings consume the least (0.48 MJ/m^2/day).

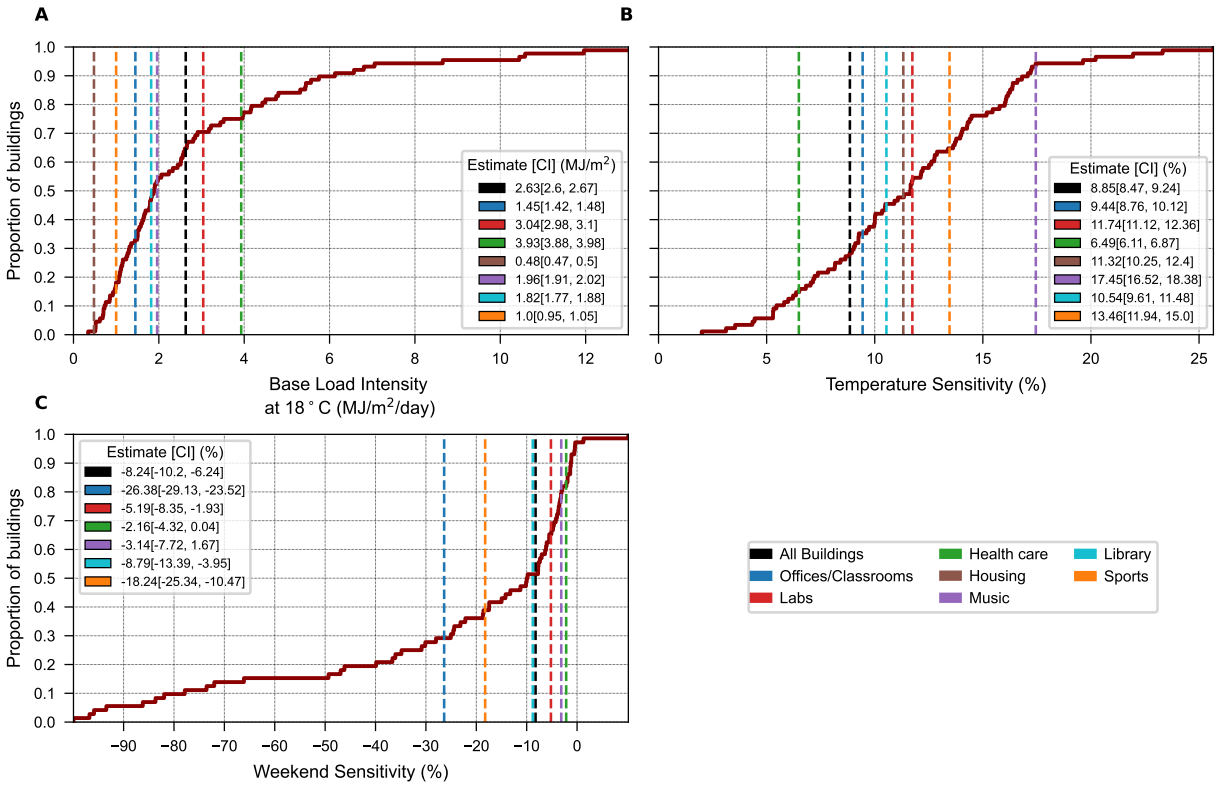

Figure 4: Statistical decomposition of cooling load using empirical cumulative distribution functions in 2017. The vertical dashed lines capture the responses at the portfolio-level and at the aggregate level of different building types. The estimates and their 95% Confidence Intervals (denoted by CI) for different building types are also tabulated on each subplot. (A) Base load intensity at 18°C (1.8°F): About 65% of buildings, which mostly comprises Offices and Classrooms buildings, are less cooling-intensive than the portfolio-level cooling loads. (B) Temperature sensitivity: About 28% of buildings were less temperature sensitive than the portfolio-level cooling loads. (C) Weekend sensitivity: The buildings containing offices and classrooms exhibit the largest reduction in their cooling demand on the weekends.

We express the temperature sensitivity of a building as the percentage increase in cooling consumption intensity (MJ/m^2/day) due to a 1°C (1.8°F) increase in mean daily OAT (Figure 4B). In 2017, the portfolio-level temperature sensitivity was 8.9%; across years it was 7.6-9.8%. A wide range is also measured for temperature sensitivity, with nearly 2-fold variation across building types. Heath care facilities have the lowest temperature sensitivity (6.5%), while special-purpose facilities for sports (13.5%) and music had the largest (17.5%).

The sensitivity to weekends, which we are using as a proxy variable for occupancy, is also very heterogeneous across buildings (Figure 4C). The overall portfolio's cooling consumption decreases

by about 8% on the weekends in 2017; and by 7-8% across years. The office and classroom buildings are the most affected (26%), and the healthcare buildings are the least sensitive (2%).

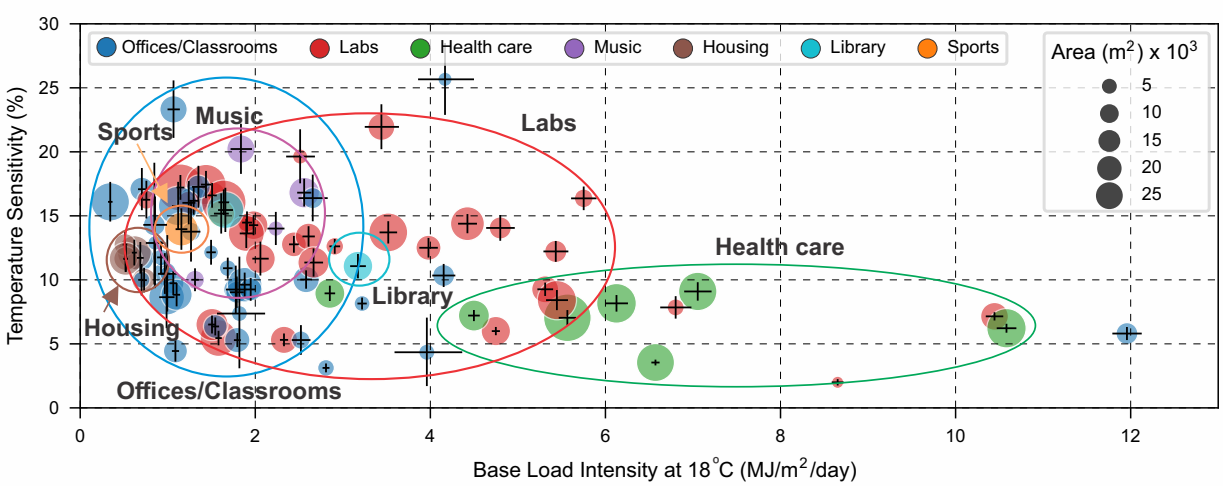

Figure 5: Associating temperature sensitivity and base load intensity of buildings to the building categories. Different colors of circles indicate different building types, while the size of the circles represent the area of building. The vertical and horizontal lines on each circle represent 95% confidence intervals. See also Figure S13.

Relationship between base load intensity, temperature sensitivity, building category and floorspace

Our results indicate no correlation between the building floorspace and estimated base load intensities and temperature sensitivities (Figure S13). However, since each building offers different types of services, we evaluate the relationship between the estimated model parameters and different building categories in Figure 5. The plot shows the temperature sensitivity and base load requirements for all of the buildings with an R^2 of greater than 0.7 in 2017. Additionally, the size of the circles represents building floorspace in meters squared.

The model parameters shown in Figure 5 demonstrate that there are common features within specific types of buildings (e.g. high base load intensity for the healthcare buildings) but also, a wide range of parameters for buildings within a particular category. For example, the laboratory buildings are mostly characterized by above-average temperature sensitivity (10-15%), but the base load intensity varies over a factor of nearly 10. Buildings with a high temperature sensitivity and low base load intensity likely contain laboratories that conduct frequent air changes, while those with a higher base load intensity house laboratories that use process cooling. The healthcare buildings uniformly demonstrate low temperature sensitivity and high base load intensity due to

the necessity of maintaining a consistent indoor environment, ensuring patient well-being, and facilitating a continuous and reliable operation of medical equipment. Residential buildings have the most uniform parameters, with very low base load intensity and 11 to 13% temperature sensitivity. Office buildings are the most heterogeneous category, with large variations in base load intensity and temperature sensitivity, reflecting the variation in occupancy and building use. For instance, the building with the highest base load intensity houses a data center and information technology offices.

Discussion

Uncertainty and trends in energy consumption drivers

Buildings with a smaller base load intensity tend to have a higher relative uncertainty in the estimates of temperature sensitivity and base load intensity, as shown by the 95% confidence intervals included in Figure 5. In contrast, the estimates are more accurate for buildings with lower temperature sensitivity. Note that there is a similarity in uncertainty patterns of both temperature sensitivity and base load intensity. This similarity could be attributed to our method for computing base load intensity, which may be introducing interactions between the variables, thus influencing the consistency in uncertainty patterns.

Year to year variability in the baseline intensity, temperature sensitivity, and weekend sensitivity for different categories of building vary from 8-17%, 6-19%, and 4-12%, respectively (Figure S14). In general, base load intensity varies less year to year than the temperature sensitivity. This suggests the baseline operations of the buildings is relatively stable, and a more variable response to fluctuations in the weather. Further, buildings show the greatest dispersion in their weekend sensitivity, probably reflecting local decision making on whether to cool the buildings on the weekends or not.

With the exception of the period from 2015 to 2016, there are no statistically significant temporal trends over the 5 year period in base load intensities and temperature sensitivities for buildings within the three major categories: Offices and Classrooms, Laboratories, and Healthcare (Figure S15, Table S3). For this analysis we included only buildings that have data from all 5 years (e.g. the new hospital is not included). Between 2015 and 2016, base load intensities went down, and temperature sensitivities increased among all three categories of buildings. Interestingly, these

changes coincide with the transition of all analyzed buildings to an electrified district energy system in 2015, suggesting that they could be due to an initial period of adaptation and fine-tuning of operations. Additionally, the health care buildings exhibit a distinct pattern as their base load intensities in 2017 are statistically different from 2015. This discrepancy might be attributed to the introduction of a new hospital building, which could have influenced the consumption patterns and operations in the existing health care buildings.

Implications for improving chilled water demand management

The models presented in this paper can be used to establish building-specific data-driven baselines that are a necessary first step for evaluating options for efficiency improvements and demand flexibility. The existing fleet of commercial buildings will remain for decades. Better understanding the drivers of energy consumption in the commercial building stock is needed to drive decarbonization efforts: energy efficiency benchmarking [29], retrofits [30, 31], demand response programs [32], resilience planning [13], and general urban sustainability policy making [33]. Our findings have important implications for electrical utility planners, district energy system operators, and for managing the energy transition.

Commercial buildings are typically larger and much more complex than residential buildings. But our results indicate their overall energy consumption may be easier to predict, which is likely explainable by aggregation effects. One of our main findings is that a simple regression framework with a small number of predictors can explain a large fraction of variance in cooling loads. This is in contrast with the much more complex models that are used in the literature to explain the energy consumption of residential buildings (SI, Review of related literature). Operations and usage appear to drive heterogeneity between commercial buildings. Their influence will not always be easy to identify a-priori in mixed-use commercial buildings, further motivating the need for measurement-based, data-driven baseline energy consumption models. Our results also indicate strong value in collecting and analyzing thermal utility as a complement to the electric load analyses that are increasingly widespread in the commercial sector.

These data-driven benchmarks serve as efficient tools for promptly forecasting the future cooling demand in response to warming temperatures. For instance, these models project that increasing temperatures would impact Labs the most, causing an 11.9% growth in their average daily cooling consumption intensity if mean daily temperatures increase by 1°C. Similarly, Offices and

Classrooms would also experience temperature-related increases, with a 1°C increase in mean daily OAT leading to a 9.3% growth, while Health care buildings, due to their lowest temperature sensitivity among the three categories, would encounter a 6.3% increase in their average daily cooling consumption intensity. As temperatures continue to rise, these insights would be invaluable for planning and optimizing space conditioning methods in different types of buildings.

Our characterization of chilled water use in commercial buildings helps identify candidates for building energy retrofits. For example, the buildings with higher temperature sensitivity in Figure 5 could be retrofitted with window upgrades, insulated walls and roofing, and installation of awnings and other shading devices that would limit the exposure to direct sunlight and reduce the thermal sensitivity of the building. Alternatively, retrofitting buildings with low thermal sensitivity and high baseline intensity requires a different approach to reduce cooling loads, for example on decreasing the cooling required for process loads or ventilation. Options include using independent cooling systems for zones with the largest requirements, upgrading air handlers to enable humidity-following temperature set points and to cool only the outside air and not the recycled air, adding occupancy sensors to reduce ventilation requirements for facilities that are not be used, and adding heat-recovery chillers to reduce reheat requirements [34].

These models are also a useful foundation for fault diagnosis tools. Since co-located buildings, even within the same category, exhibit widely varying temperature sensitivities and base load intensities, the information in Figure 5 can help flag unusual behavior in buildings and provide a preliminary diagnosis for atypical energy consumption. This information can be used to assess whether a mechanical retrofit and/or changes to the operating regime would be beneficial. Further, changes in the temperature sensitivity and baseline intensity can be monitored to assess the impact of energy-efficiency measures and energy retrofits that have been implemented.

Finally, the data from these models will aid building managers in devising appropriate Demand Side Management (DSM) strategies for portfolios of different types of buildings. For example, in buildings with high temperature sensitivity, reduction in cooling is likely to make the buildings uncomfortable for occupants during heat waves but will also have a high impact. Whether they are suitable candidates for DSM depends on whether demand response (DR) events can be synchronized with periods of reduced or low occupancy. In the case of this study, to achieve significant portfolio-scale DR, participation of laboratory and healthcare buildings would be needed since they represent

about 80% of the total cooling load (Figure 2B). However, only their non-critical loads could be shed during the peak cooling consumption hours. Careful assessment of the impact of DR measures would be required in advance of implementing any DR strategy.

Experimental Procedures

Resource Availability

Lead Contact. Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Aqsa Naeem (anaeem@stanford.edu).

Materials Availability. No materials were used in this study.

Data and Code Availability. The dataset used for this study is available at `https://data.mendeley.com/preview/pxv3724ry3?a=b4be5688-a5da-4415-b33d-47c30b8e2762`.

Data acquisition and pre-processing

This study uses time series data measured in 119 commercial buildings to evaluate and demonstrate a set of building characterization tools at an urban scale. The buildings are located in a warm-summer Mediterranean climate where energy is served by an electrified district energy network. To meet the thermal demands of buildings, hot water and chilled water are produced on site at the Central Energy Facility (CEF) and distributed to the buildings through a network of underground pipes. The hot and chilled water are used by HVAC systems to condition air that is then used to maintain indoor zone temperatures within specified bounds.

Five years of hourly building meter data are obtained from the university's data historian. Weather data are available from an on-site weather station and include OAT, solar radiation, wind speed, Tdew, and RH. Our building-level dataset includes building-level cooling load, floorspace area, and information on the type of buildings. Cooling load is measured from the flow rate of chilled water going through the building and the difference in temperature of water in the supply and return pipes. This includes loads from the building HVAC systems, process chilled water loads, and losses. Portfolio-level load, which we denote by 'All Buildings' in Figure 4 and Figure S14, is similarly measured using the flow rate and temperature difference between the supply and return water at the CEF.

As outlined in Table 1, the building meter data undergo a series of filters to ensure that they are smooth and have fewer outliers and instances of missing data points. The first step involves passing the data through fixed upper and lower bounds that are unique to each meter to eliminate unwarranted spikes and incorrect negative measurements. Then a moving window of size 10 (240 hourly values) is used to compute rolling mean and standard deviation. All the points that lay beyond the rolling bounds of mean $\pm$ 4 x standard deviations are removed. The third step involves filling in the missing values using linear interpolation. If more than 3 consecutive observations are missing, then that day is excluded from the data.

The dataset is then transformed in the fourth step to match meters to buildings because there is a small percentage of buildings that either share a single meter or have multiple meters each. In these cases, data reconciliation between the building's actual use and meters is made possible by consolidating the meter-splitting rules that are used for billing purposes and ensuring that the consumption across all buildings sum to the total portfolio-level load.

Finally, the dataset is further cleaned by removing outliers that are identified by fitting a linear regression model on the cooling consumption of individual buildings and inspecting the residuals. These models use outside air temperature along with a combination of other weather- and occupancy-related variables to model the cooling load. For each building, outliers are detected using an absolute studentized residual threshold of 3 and then eliminated from the data.

Table 1: Data pre-processing: Building raw meter data are passed through a series of cascaded filters to eliminate outliers. Note IDs denote Identities.

Label	Transformation	Method
Step 1	Filtering	Fixed upper and lower bounds (variable across buildings).
Step 2	Filtering	Variable: Mean $\pm$ 4$\times$ standard deviation.
Step 3	Linear interpolation	To fill in 3 consecutive missing values.
Step 4	Reconciliation	Mapping meters to unique building IDs.
Step 5	Outlier removal	Regression-based detection and removal of outliers.

Building cooling load model selection process

We now describe model selection and evaluation (Figure 6). First, we use a three-step feature curation process to obtain feature sets in each year that could explain the buildings' cooling load (Figure S5). For p features, the feature curation block outputs 2^p curated sets, denoted by $\mathscr{F}_i$ where $i \in \{0, \ldots, 2^p - 1\}$. These sets are then made input to k-fold cross validation block, which generates 2^p models corresponding to each $\mathscr{F}_i$ for the n^{th} building in a total of N buildings. The third step involves the selection of an optimal model, M^n_{opt}. For the n^{th} building, we define the optimal model as the most parsimonious model that has the lowest cross-validated error, which is selected using the one standard-error rule i.e., by calculating the average mean squared error (MSE) and standard errors of MSE of the folds [35]. We also evaluate the impact of using a different model selection criterion e.g., using adjusted coefficient of determination (R^2_{adj}) and it was found that the selected model through a different criterion does not result in the most parsimonious model. In the last step, the selected models are evaluated using R-squared error. These steps are repeated for each year to account for the differences in the cooling load of buildings. The details for each of these steps are provided in the SI.

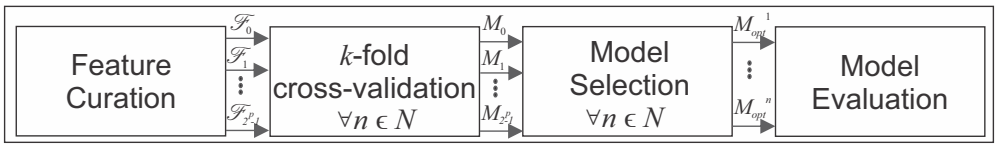

Figure 6: Schematic of model selection and evaluation process. Figure S5 illustrates the feature curation process.

Model formulation

We opted for using log-linear models in our main specification. These models offer a comparable goodness-of-fit and predictive performance when contrasted with linear models, as illustrated in Figure S9. In addition, the log-linear models facilitate a more intuitive understanding of the effect of temperature increase by allowing us to express the temperature sensitivities in terms of percentage changes in cooling load. However, it's important to note that if they are used to extrapolate to extremely high temperatures not observed in our dataset, these models could lead to implausible results since cooling consumption cannot increase indefinitely and may surpass the limits of the cooling systems.

Supplemental Information

The supplemental PDF contains additional information on our Results and on our Experimental Procedures, Figures S1-S18, and Tables S1-S3.

Acknowledgments

This research was supported by Stanford Land, Buildings, and Real Estate, and TotalEnergies SE.

Declaration of interest

The authors declare no competing interests.

References

[1] M. Huismans, "Electrification." `https://www.iea.org/reports/electrification`, 2022. Online; accessed 2022-11-11.

[2] J. Langevin, C. B. Harris, and J. L. Reyna, "Assessing the potential to reduce us building co2 emissions 80% by 2050," *Joule*, vol. 3, no. 10, pp. 2403–2424, 2019.

[3] P. Albertus, J. S. Manser, and S. Litzelman, "Long-Duration Electricity Storage Applications, Economics, and Technologies," *Joule*, vol. 4, no. 1, pp. 21–32, 2020.

[4] N. S. Diffenbaugh and M. Burke, "Global warming has increased global economic inequality," *Proceedings of the National Academy of Sciences*, vol. 116, no. 20, pp. 9808–9813, 2019.

[5] M. Subbarao, "Shifting distribution of land temperature anomalies, 1951-2020." `https://svs.gsfc.nasa.gov/4891`, 2021. Online; accessed 2022-11-11.

[6] R. C. Triolo, R. Rajagopal, F. A. Wolak, and J. A. de Chalendar, "Estimating cooling demand flexibility in a district energy system using temperature set point changes from selected buildings," *Applied Energy*, vol. 336, p. 120816, 2023.

[7] EIA, "Electric power annual 2021," tech. rep., US Energy Information Administration, 2022.

[8] EIA, "Annual energy outlook 2022 with projections to 2050.." `https://www.eia.gov/outlooks/aeo/pdf/AEO2022_Narrative.pdf`, 2022. Online; accessed 2023-04-26.

[9] EIA, "2018 commercial buildings energy consumption survey (cbecs)," tech. rep., Energy Information Administration, 2021.

[10] U.S. Department of Energy, "Commercial reference buildings." `https://www.energy.gov/eere/buildings/commercial-reference-buildings`.

[11] J. New, M. Adams, B. Bass, N. Clinton, A. Berres, E. Garrison, T. Guo, and M. Allen-Dumas, "Model america: A crude energy model and data for nearly every us building," *Available at SSRN 4220628*, 2022.

[12] C. F. Reinhart and C. C. Davila, "Urban building energy modeling–a review of a nascent field," *Building and Environment*, vol. 97, pp. 196–202, 2016.

[13] T. Hong, Y. Chen, X. Luo, N. Luo, and S. H. Lee, "Ten questions on urban building energy modeling," *Building and Environment*, vol. 168, p. 106508, 2020.

[14] N. E. Fernandez, S. Katipamula, W. Wang, Y. Xie, M. Zhao, and C. D. Corbin, "Impacts of commercial building controls on energy savings and peak load reduction," tech. rep., Pacific Northwest National Lab.(PNNL), Richland, WA (United States), 2017.

[15] J. Langevin, C. B. Harris, A. Satre-Meloy, H. Chandra-Putra, A. Speake, E. Present, R. Adhikari, E. J. Wilson, and A. J. Satchwell, "US building energy efficiency and flexibility as an electric grid resource," *Joule*, vol. 5, pp. 2102–2128, Aug. 2021.

[16] T. Hoyt, E. Arens, and H. Zhang, "Extending air temperature setpoints: Simulated energy savings and design considerations for new and retrofit buildings," *Building and Environment*, vol. 88, pp. 89–96, 2015.

[17] B. Bass, J. New, N. Clinton, M. Adams, B. Copeland, and C. Amoo, "How close are urban scale building simulations to measured data? examining bias derived from building metadata in urban building energy modeling," *Applied Energy*, vol. 327, p. 120049, 2022.

[18] R. Yin, S. Kiliccote, and M. A. Piette, "Linking measurements and models in commercial buildings: A case study for model calibration and demand response strategy evaluation," *Energy and Buildings*, vol. 124, pp. 222–235, 2016.

[19] J. S. MacDonald, E. Vrettos, and D. S. Callaway, "A critical exploration of the efficiency impacts of demand response from hvac in commercial buildings," *Proceedings of the IEEE*, vol. 108, no. 9, pp. 1623–1639, 2020.

[20] D. Coakley, P. Raftery, and M. Keane, "A review of methods to match building energy simulation models to measured data," *Renewable and sustainable energy reviews*, vol. 37, pp. 123–141, 2014.

[21] A. Chong, Y. Gu, and H. Jia, "Calibrating building energy simulation models: A review of the basics to guide future work," *Energy and Buildings*, vol. 253, p. 111533, 2021.

[22] N. Luo, M. Pritoni, and T. Hong, "An overview of data tools for representing and managing building information and performance data," *Renewable and Sustainable Energy Reviews*, vol. 147, p. 111224, 2021.

[23] J. A. de Chalendar, C. McMahon, L. F. Valenzuela, P. W. Glynn, and S. M. Benson, "Unlocking demand response in commercial buildings: Empirical response of commercial buildings to daily cooling set point adjustments," *Energy and Buildings*, p. 112599, 2022.

[24] C. Miller, A. Kathirgamanathan, B. Picchetti, P. Arjunan, J. Y. Park, Z. Nagy, P. Raftery, B. W. Hobson, Z. Shi, and F. Meggers, "The building data genome project 2, energy meter data from the ashrae great energy predictor iii competition," *Scientific data*, vol. 7, no. 1, pp. 1–13, 2020.

[25] N. Luo, Z. Wang, D. Blum, C. Weyandt, N. Bourassa, M. A. Piette, and T. Hong, "A three-year dataset supporting research on building energy management and occupancy analytics," *Scientific Data*, vol. 9, no. 1, pp. 1–15, 2022.

[26] A. Nutkiewicz, Z. Yang, and R. K. Jain, "Data-driven urban energy simulation (due-s): A framework for integrating engineering simulation and machine learning methods in a multi-scale urban energy modeling workflow," *Applied energy*, vol. 225, pp. 1176–1189, 2018.

[27] D. L. Villa, "Institutional heat wave analysis by building energy modeling fleet and meter data," *Energy and Buildings*, vol. 237, p. 110774, 2021.

[28] S. Seabold and J. Perktold, "statsmodels: Econometric and statistical modeling with python," in *9th Python in Science Conference*, 2010.

[29] H. Li and X. Li, "Benchmarking energy performance for cooling in large commercial buildings," *Energy and Buildings*, vol. 176, pp. 179–193, 2018.

[30] S. Nagpal, J. Hanson, and C. Reinhart, "A framework for using calibrated campus-wide building energy models for continuous planning and greenhouse gas emissions reduction tracking," *Applied Energy*, vol. 241, pp. 82–97, 2019.

[31] X. Li and R. Yao, "Modelling heating and cooling energy demand for building stock using a hybrid approach," *Energy and Buildings*, vol. 235, p. 110740, 2021.

[32] J. L. Mathieu, P. N. Price, S. Kiliccote, and M. A. Piette, "Quantifying changes in building electricity use, with application to demand response," *IEEE Transactions on Smart Grid*, vol. 2, no. 3, pp. 507–518, 2011.

[33] U. Ali, M. H. Shamsi, C. Hoare, E. Mangina, and J. O'Donnell, "A data-driven approach for multi-scale building archetypes development," *Energy and buildings*, vol. 202, p. 109364, 2019.

[34] E. Bonnema, M. Leach, and S. Pless, "Technical support document: Development of the advanced energy design guide for large hospitals – 50savings." `https://www.nrel.gov/docs/fy13osti/52588.pdf`, 2013. Online; accessed 2023-08-14.

[35] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning: With Applications in R*. Springer Publishing Company, Incorporated, 2014.