Title

Comparing simulated demand flexibility against actual performance in commercial office buildings

Authors

Yin, Rongxin
Liu, Jingjing
Piette, Mary Ann
et al.

Publication Date

2023-09-01

DOI

10.1016/j.buildenv.2023.110663

Peer reviewed

Comparing Simulated Demand Flexibility against Actual Performance in Commercial Office Buildings

Rongxin Yin[a,*], Jingjing Liu[a], Mary Ann Piette[a], Jiarong Xie[b], Marco Pritoni[a], Armando Casillas[a], Lili Yu[a], Peter Schwartz[a]

[a]*Lawrence Berkeley National Laboratory, Berkeley, California, USA*
[b]*School of Architecture, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA*

Abstract

Commercial building energy benchmarking has been used as a mechanism to evaluate energy use of a single building over time, relative to other similar buildings, or to simulations of a reference building conforming to various energy standards. Lack of empirical demand flexibility data and consistent flexibility metrics has limited the ability to compare demand flexibility performance with estimated demand flexibility in buildings. In this study, we collected demand response performance data for a total of 831 demand response events from 192 sites as a first step to build such a building demand flexibility dataset, and propose a standard core data schema to consolidate field data from different sources. We also performed parametric simulations of a control strategy called "global temperature adjustment" using commercial office prototype building models. We then compared the simulated demand flexibility performance against the actual data for offices with global temperature adjustment strategy implemented. During demand response events with an average outside air temperature of 34°C (range 23°C-42°C), the measured demand decrease intensity of the demand flexibility metrics were 6.1 watts per square meter (W/m^2), 10.0 W/m^2, 11.1 W/m^2, 7.1 W/m^2, and 4.7 W/m^2 for small, small-medium, medium, medium-large, and large office buildings, respectively. Compared to the measured data in medium- and large-size buildings, the simulated demand decrease intensity was 0.7 W/m^2 (17%) lower on average. The discrepancy between simulated and measured peak demand intensities fell within one standard deviation of the mean mea-

*Corresponding author: Rongxin Yin
Email address: `ryin@lbl.gov` (Rongxin Yin)

sured data. The comparison results validate the credibility of simulations in capturing real building data for assessing the technical potential of building demand flexibility.

Keywords: demand flexibility, commercial office building, cross validation, control strategy, global temperature adjustment, field-testing, prototype building model

---

Nomenclature

AC Air conditioner

ASHRAE American Society of Heating, Refrigerating and Air-Conditioning Engineers

BAS Building automation systems

CAV Constant air volume

CDD Cooling degree days

COP Coefficient of performance

CZ Climate zone

DDI Demand decrease intensity

DDP Demand decrease percentage

DER Distributed energy resources

DF Demand flexibility

DII Demand increase intensity

DOE Department of Energy

DR Demand response

EE Energy efficiency

FERC Federal Energy Regulatory Commission

GEB Grid-interactive efficient buildings

GTA Global temperature adjustment

HVAC Heating, ventilation and air-conditioning

IMF Input marco file

MELs Miscellaneous electric loads

OAT Outside air temperature

RTU Rooftop unit

TCLs Thermostatically controlled loads

TES Thermal energy storage

VAV Variable air volume

VFD Variable frequency drive

1. Introduction

Demand flexibility is a relatively new term used to categorize different ways of managing demand-side loads to provide demand-side flexibility and grid interactivity. The recent national roadmap for Grid-Interactive Efficient Buildings (GEB) [1] provides a definition: “demand flexibility, also sometimes referred to as load flexibility, is the capability provided by on-site distributed energy resources (DERs) to reduce, shed, shift, modulate, or generate electricity.” DERs include energy efficiency, energy storage, demand response, electric vehicles, grid-interactive efficient buildings, combined heat and power, and renewable energy such as solar photovoltaics. In the past few decades, the electricity market has begun to consider demand-side resources as valuable assets for meeting capacity needs, improving reliability, reducing wholesale and retail costs, and supporting grids with higher levels of renewable energy distributed generation. More recently, increased levels of renewable energy have begun to create instances of oversupply, more often during low-load, shoulder seasons — the months between the winter heating and summer cooling seasons. These instances of oversupply can cause unstable grid conditions and negative energy prices on the wholesale market, and are usually met with energy curtailment. As a result, manipulating demand (shape, shift, shed, and shimmy) to participate in various programs in the electricity market has seen increased interest from grid operators and regulators [2]. A notable example of demand flexibility includes demand response (DR), which has traditionally been defined as load shedding or shifting by consumers in response to higher electricity prices or grid supply shortages, usually during extreme hot or cold weather events. The U.S. Federal Energy Regulatory Commission (FERC) has defined DR as “Changes in electric usage by demand-side resources from their normal consumption patterns in response to changes in the price of electricity over time, or to incentive payments designed to induce lower electricity use at times of high wholesale market prices or when system reliability is jeopardized” [3]. Since 2008, federal policy has progressively introduced demand response as a dispatchable resource that, in contrast to other resources, is able to participate in organized energy markets [3]. For DR participation in the U.S. wholesale markets, the potential peak demand in 2011 peaked at 32,488 megawatts (MW) and reached 30,788 MW in 2019, accounting for approximately 5.3%-7.0% of the peak demand [4]. From 2012 to 2020, potential peak demand savings from retail demand response programs in the United States increased by approximately 2,517 MW, or 8.8%, from 28,503 MW to approximately 31,020 MW. In 2019, utilities reported over 15,000 MW of potential peak demand savings from the residential and commercial customer class, roughly 51% of the reported total retail potential peak demand savings [4].

As one form of demand flexibility in buildings, DR has been playing a significant role in reducing peak demand for residential and commercial customers. In particular, thermostatically controlled loads (TCLs) have been the primary flexible demand resource in buildings. Thermostat setpoint adjustment is a common DR control strategy that utilizes the building thermal mass to reduce the building cooling load in summer. Since the 1990s, simulations and laboratory tests have demonstrated the use of building thermal mass to reduce peak demand for cooling loads (10% to 40%) [5, 6, 7, 8, 9]. Furthermore, optimal zonal temperature strategy (such as a linear, step, or exponential reset of thermostat temperature setpoint) can reduce the peak demand about 25% to 45% and still deliver acceptable occupant thermal comfort [10, 11]. Similar simulation studies have shown that a simple zonal temperature adjustment strategy can reduce chiller power use 80%-100% (10-23 W/m^2) during peak hours [12]. Considering a linear relationship between zone temperatures and cooling loads, the near-optimal setpoint trajectory from the simplified inverse building model reduced peak cooling power by an average of 31.6% over the four test days [13]. The same approach was deployed in three representative small, medium, and large commercial buildings, reducing peak cooling loads by 33%, 42%, and 51%, respectively. Based on the simplified building model, the estimated peak cooling load reduction ranges from 22 W/m^2 to 32 W/m^2 [14]. In the same medium-size office building, researchers conducted a follow-up simulation using EnergyPlus to evaluate the effect of nighttime and morning pre-cooling on the following day's peak demand shed [15]. Simulation results show that increasing the zone temperature setpoint by 2.2°C (4°F) can reduce chiller electricity consumption by about 33%. A recent study [16] developed a novel split air-conditioner (AC) load model for both constant-speed and variable-speed compressors. The 2°C (3.6°F) thermostat reset control strategy achieves an average AC load shed of 0.2 kW per household, which is about 25% of the air conditioner's rated power demand. A large-scale parametric simulation was used in a global sensitivity study of demand response in medium-sized office buildings to identify critical building design factors impacting the demand flexibility performance of new constructions[17]. Results indicated that design internal loads, internal thermal mass, the chiller plant sizing factor and coefficient of performance are ranked as critical parameters for achieving significant potential in demand flexibility. It is worth noting that most of the cited studies conducted field tests or simulations on hot days, and the results may not reflect the effects of humidity on the DR performance.

The aforementioned simulated and experimental results show that building thermal loads can be shifted or reduced to provide demand flexibility with appropriate control strategies, which have inspired many follow-up simulations [18, 19, 20, 21], laboratory [22, 23] and field tests [16, 24, 25, 26, 27], and pilot studies [28, 29, 30, 31] on the potential of building demand flexibility (DF). Each method (simulation, laboratory/field testing) deployed in the above studies has unique advantages from different aspects of interest of DF influential factors, magnitudes, and variations. In general, simulation can evaluate the impact of different influential factors on the DF performance while maintaining the consistency of the remaining model inputs. On the other hand, field-testing can be used to quantify the actual DF performance while considering the characteristics of the building itself. By contrast, laboratory testing can explore DF performance by using different aspects in the testbed setup. A recent study [32] conducted a comprehensive review of methods for quantifying energy flexibility in residential buildings. It concluded that 85% of the reviewed studies were simulations, only about 10% of the reviewed studies involved actual measurements in the quantification, and the rest deployed both simulations and measurements. Another review paper presented several open-source building energy management datasets for the use of reinforcement learning and data-driven modeling, while there was very limited information about demand flexibility performance [33]. In the context of parametric analysis of DF performance, simulations can be used to estimate DF characteristics for different building types, vintages, and climate conditions through parametric/sensitivity analysis [18, 19]. Existing large datasets are available to help estimate flexibility across the national building stock [20, 34, 35]. Several previous simulation studies have used EnergyPlus to perform large-scale parametric simulations of demand flexibility strategies such as global temperature adjustment (GTA) in commercial buildings [21, 11], while also analyzing its relation to time of day and outdoor weather conditions such as outside air temperature (OAT). A national-level assessment study analyzed the demand response potential for seven types of small commercial buildings in different vintages and climates, including various control strategies such as pre-cooling, shading, and lighting power reduction [34]. A similar simulation study utilized both residential and commercial prototype building models to project the baseline annual electricity consumption and net peak demand for U.S. buildings by 2030 [35]. By incorporating the best possible energy efficiency and flexibility measures, it was estimated that up to 800 TWh of annual electricity usage and 208 GW of daily net summer peak demand could potentially be avoided. It is important to note that the accuracy of the bottom-level modeling of demand flexibility is critical to deploying this building-grid resource in the future.A promising approach was presented to deploy automatic calibration of urban building energy models by learning building characteristics and energy performance from a building energy database [36]. In contrast to the use of physical energy models in this area, a recent study delpoyed a data-driven simulation apporach to evaluate the performance of demand-driven control strategies at the district scale [37]. It was found that when building occupancy decreases by 25% to 75%, there's a 5% to 15% reduction in space cooling demand at the campus scale. A similar study employed a data-driven model for model predictive control of large-scale HVAC systems in providing demand flexibility [38]. It can be expected that large demand flexibility datasets will be needed for use in data-driven modeling approaches. Despite the broad scope of these analyses, the models on which they are based have not been validated against sufficient field data. To some extent, their applicability to actual building stock is limited.

From the perspective of DF measurement and verification, there are some similarities in the DR performance quantification metrics framework, such as demand reduction units (W/m^2) or the percentage demand reduction (%) for the entire building or cooling load. A recent study [39] conducted an in-depth review of data-driven key performance indicators (KPIs) for building energy flexibility. These KPIs, which include metrics for load shedding and shifting at the building level, are commonly used in energy flexibility quantification using a baseline model. As highlighted in a recent review study [40], most building codes and standards primarily focus on energy performance, with limited emphasis on energy flexibility metrics. To fill this gap, in our recent companion paper [41], we defined a set of metrics for both "load shed" (Shed) and "load shift" (Shift), providing single-event metrics with an associated set of metric attributes. While some of the evaluation metrics used in these studies align, overall there is a need for a standard data schema to document, organize, and standardize DR performance-related data. Such a standard data schema can be used to integrate these datasets and compare DR performance by building type, vintage, and climate location. In addition, the literature lacks standard field datasets to compare model simulations against measurements or benchmarking flexibility performance against similar buildings. Local utility companies and aggregators manage sizable portfolios of buildings that participate in different grid services, but the data they process are not public.

The above summary identified several research gaps within the field of building demand flexibility, including: (1) lack of standard data schemes for compiling simulation and field test datasets, (2) lack of sufficient empirical demand flexibility data, and (3) limited in-depth analysis of the difference between simulated and measured DF performance. In this study, we attempted to leverage past research and field/pilot studies from nearly 200 commercial buildings as a basis for addressing these gaps. The main contributions of the paper are the following:

- We propose a standard data schema for integrating the field demand response performance data to build a commercial building demand flexibility database.

- We compare the field DF dataset of hundreds of buildings by building type, vintage, and climate to help identify performance characteristics.

- We compare simulated DF results with field data to justify the use of prototype models in the assessment of DF potential for commercial buildings, and provide a set of recommendations for DF modeling in small, medium, and large offices.

The rest of this paper is organized as follows: Section 2 describes the methods deployed in this study, including field data schema standardization and prototype model simulations. Including an exploratory data screening and analysis in section 3, we present a data schema for defining common building characteristics and demand flexibility performance metrics in commercial buildings. In section 4, we discuss the use of this demand flexibility dataset and present a use case by comparing the prototype building model simulations against similar field-testing buildings. Last, we provide an in-depth discussion about the discrepancy between the simulated and field DF performance dataset.

2. Methodology

As the first attempt to build such a building demand flexibility dataset, it was essential to develop a standard data schema consistent with DF metrics framework [41] by leveraging existing field studies. The goal was to facilitate more field studies into this dataset with the standard data schema, and to improve the availability and consistency of these data across buildings and markets and among stakeholders. This data schema contains information about the physical characteristics, location, and DF performance of the building and serves as the core data schema of a commercial building demand flexibility database.Similar to the modeling methods used in previous studies [11, 20, 21], we propose to use the commercial prototype building models [42] for assessing the DF characteristics by different building types, vintages, and climates.

2.1. Streamlining field data collection with standard data schemas

Comprehensive building field-testing datasets are valuable; however, collecting, cleaning, and incorporating them into standard datasets is very challenging. It is necessary to design a standard data schema to describe basic information about building and DF performance characteristics, especially for compiling data from various data sources. Figure 1 illustrates the proposed entity–relationship data schema for compiling all collected data into a database. The data were categorized into four primary groups: building information, DF measures, program events, and DF performance metrics. Note that the focus of this data schema is to catalog information for demand flexibility in commercial buildings. Details about the components defined in the data schema are discussed in section 3.

2.2. Prototype model simulations

In this study, we conducted parametric simulations of the common strategy "global temperature adjustment" for small, medium, and large prototype office buildings in two climates, as shown in Table 1. As defined in [43], GTA is a strategy that allows building operators to adjust the space temperature setpoints for an entire facility. The reasons for our focus on office sites that implement only the GTA control strategy include: (1) GTA is the most commonly used DF strategy in office buildings, and (2) the different combinations of DF control strategies for each building present a significant challenge of

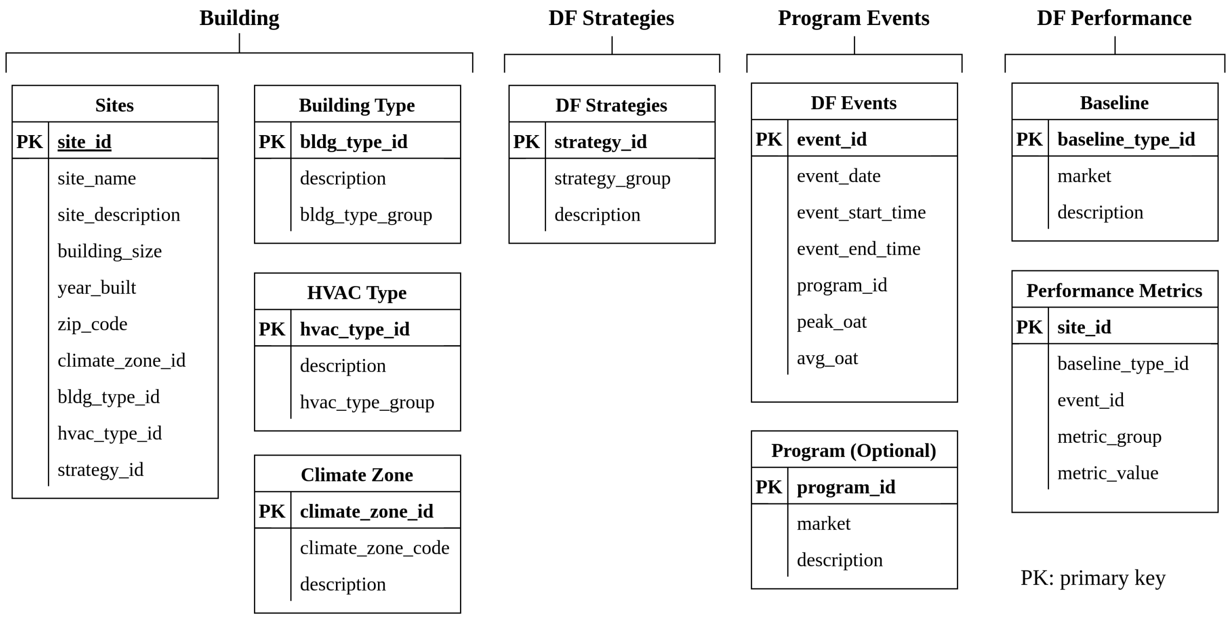


Figure 1: Proposed common data schema for building demand flexibility

modeling capability when using prototype building models. Three DF metrics were calculated for each DF simulation run: (1) DDI (demand decrease intensity: W/m^2, which is calculated as the amount of demand shed per building floor area), (2) DII (demand increase intensity [optional for "shift" service]: W/m^2, which is calculated as the amount of demand increase per building floor area), and (3) DDP (demand decrease percentage: %, which is calculated as the percent of load shed over the whole building power baseline). Those metrics are the most frequently used energy and demand related metrics in field studies [41]. With respect to the location, two climate zones are selected from the 19 climate zones for the U.S. as defined in the ASHRAE Standard 169-2013. The selection of these building types, vintages, and climates was based on the available field sites' characteristics.

By comparing prototype model simulation results with the collected field data, we expected to gain insight into the validity and limitation of the prototype simulation approach for the estimation of DF potential, such as large scale simulations [20, 21, 11]. Figure 2 depicts the simulation data flow of DF modeling using two kinds of parametric methods for a large scale of prototype model simulations and a cross-comparison framework between simulation, laboratory testing, and field data. We used EP-Macro to param-

Table 1: Summary of DF simulations for comparison with field datasets

Building Types	**Building Vintages**	**Climates**	**DF Control Strategies**	**Simulation Period**	**DF Metrics**
Small Office (511 m^2), Medium Office (4,980 m^2), and Large Office (46,338 m^2)	Pre-1980, 1980-2004, 90.1-2004	3B (warm-dry), 3C (warm-marine)	GTA	Full year (one DR event per weekday)	DDI, DII, DDP

eterize the IMF (input macro file) files for the pre-1980 prototype building models and used the OpenStudio [44] platform to generate each simulation model with built-in measures of demand flexibility for the post-1980 prototype building models. Then we adopted the proposed DF metrics framework [41] to calculate the DF performance metrics and attributes for each simulation case by following a DF metrics calculation procedure. Note that we compared performance in terms of magnitude (e.g., demand decrease intensity W/m^2) and variation (e.g., building type, vintage, and climate) rather than detailed model calibration for each building.

On the other hand, the DF metrics used in each field-testing study vary by their aspects of interest in understanding the DF performance. In the recent work, we defined a set of DF performance metrics for various grid service products (e.g., shed and shift) [45, 41]. However, it is quite a challenge to standardize the field performance data into the same set of metrics with limited access to the raw field data. Therefore, we propose the use of DDI (Demand Decrease Intensity, W/m^2) as the DF metric in the comparison of prototype simulation results against the field data.

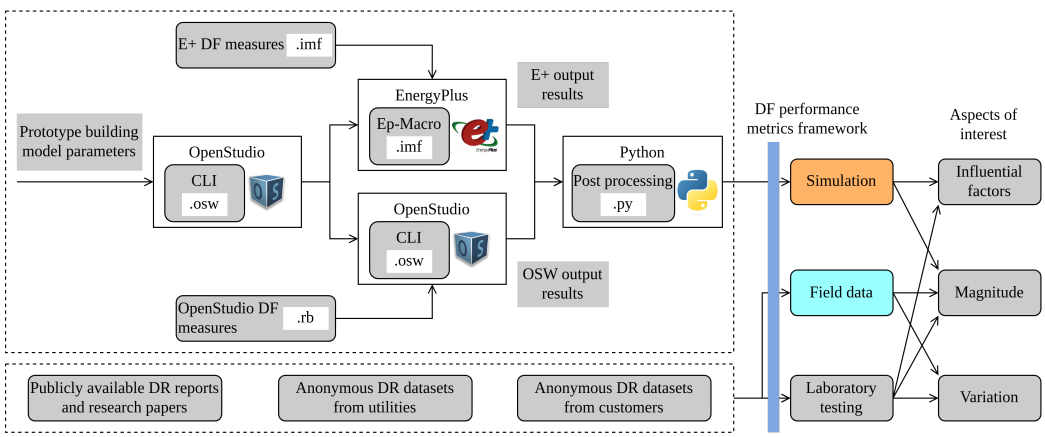


Figure 2: Cross-comparison framework between simulation, laboratory test, and field data

3. Field-testing Data Summary

For the field data presented in this study, most of the effort was spent on data cleaning and mapping to a common data schema because of the inconsistent DF performance metrics reported in previous studies. Through an extensive literature review and engagement with one utility in California and individual customers, we collected DR performance data for a total of 831 DR event days from 192 sites. Following the standard core data schema, minimum data requirements for each site include: (1) building characteristics, such as building type, vintage, floor area, location (either zip code or climate zone), and heating, ventilation and air-conditioning (HVAC) system type; (2) DF strategies, such as end-use and control sequence; (3) demand-related approaches, such as peak demand kW, intensity W/m^2, or peak demand reduction kW; and (4) event-related approaches such as event start and end date/time, as well as weather data during the event hours.

Table 2: Collected DF performance datasets as of August 2022

Data Source	**Number of Sites**	**Number of DR events**	**Reference**
Lab studies (2003-2015)	101	447	[28, 29, 30, 46, 47, 48, 19, 49, 50, 51, 52, 53, 54, 55]
Utility company	68	257	Anonymized data
Other published reports/papers	19	109	[5, 7, 10, 56, 57, 58, 26, 59, 60, 61, 22, 62]
Customers	4	18	Anonymized data
All	192	831	See above

3.1. Site Description

With respect to field building characteristics, Figure 3 presents a summary of field sites by building type, vintage, and ASHRAE climate zone. Over 90% of office sites are located in warm climate zone (CZ) 3B (warm-dry) and 3C (warm-marine). Except for one site in CZ 5A (cold-humid), the rest of the sites are in mixed climates 4A (mixed humid) and 4C (mixed marine). In this study, our primary focus is on the 97 offices, of which approximately 20% were built before 1980, 73% were constructed between 1980 and 2004, and the remaining 7% were built after 2004.

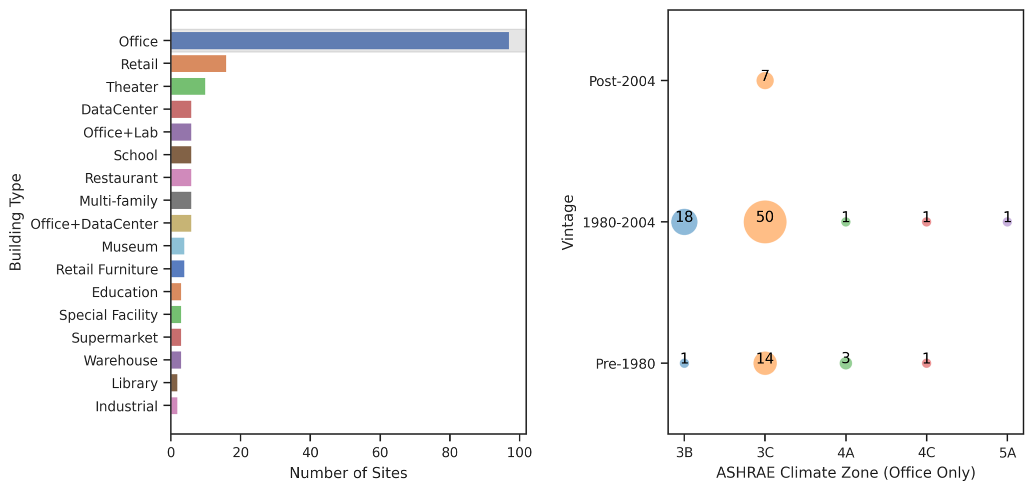


Figure 3: Summary of field demonstration sites by vintage and climate

As shown in Figure 4, about 95% of field sites have a variable air volume (VAV) system as their air distribution system. Large office buildings with a floor area of about 10,000 m^2 and above choose air-cooled or water-cooled chillers as the cold source equipment. For small-medium and medium-size office buildings with a floor area between approximately 1,000 m^2 and 10,000 m^2, packaged rooftop units or air-cooled chillers are their favored cooling equipment. Single- or multiple-zone rooftop units are an ideal and cost-effective HVAC option for small buildings. The reason for HVAC system characterization here is that DF performance can vary widely between two different HVAC system types, such as packaged rooftop units with constant air volume (CAV) or VAV systems. In the case of a VAV system, thermostat setpoint adjustment can reduce the VAV terminal's airflow rate to the minimum airflow setpoint. This is the most effective because it reduces the load of all associated air handling and cooling equipment, while keeping the zone temperature within the thermostat's control.

Figure 5 shows the ranking of field site peak demand intensity (left) and a box plot by vintages and climates (right). Here's the breakdown of office buildings by vintage and climate: 18 sites in "1980-2004, 3B", 50 sites in "1980-2004, 3C", 14 sites in "Pre-1980, 3C", 7 sites in "Post-2004, 3C", and

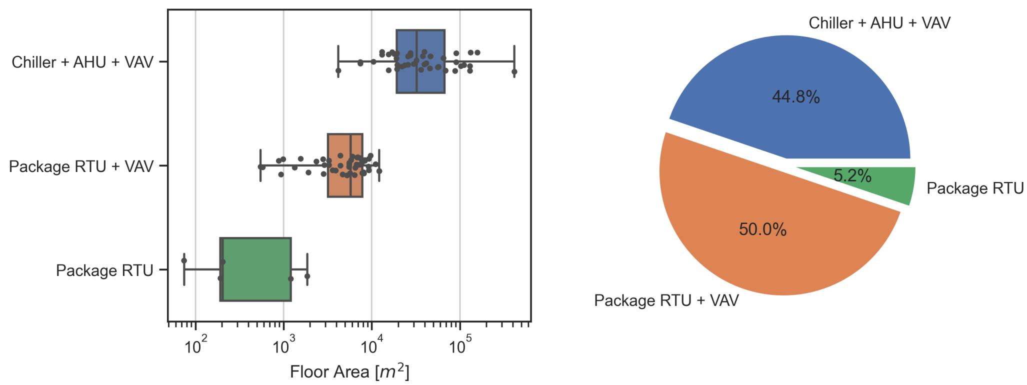


Figure 4: Summary of HVAC type by site floor area (left) and percentage of sites per HVAC type (right)

3 sites in "Pre-1980, 4A". Among offices located in the same climate zone 3C, newer buildings (represented by 50 sites from the '1980-2004' period and only 7 sites from the 'post-2004' period) exhibit a higher average peak demand intensity than those constructed before 1980. This observation was a bit unexpected, as newer buildings are more compliant with building energy efficiency codes. Possible reasons could be that: (1) newer buildings have installed more plug loads, such as computers and servers; (2) older buildings may have had efficiency retrofits installed in recent years; and (3) in addition to the huge amount of glass used in the new office building (steel-framed), most older buildings built before the 1980s have a mass/concrete envelope with a significant amount of building thermal mass to reduce high temperature fluctuations during heat waves. This observation is consistent with the commercial buildings energy consumption survey data [63]. For buildings constructed in the similar period, warmer climate (climate zone 3B and 3C) results in a relatively high peak demand intensity of building HVAC load in comparison with mixed humid climate zone 4A. When compared to the office building peak demand intensity in climate zone 3B and 3C, they are relatively lower by 24-30%. A possible reason for this difference is that climate zones 3B ($2500 < \text{CDD10°C} < 3500$) and 3C ($\text{CDD10°C} - 2500$) have higher cooling degree days (CDD) than the mixed-humid climate zone 4A ($\text{CDD10°C} < 2500$). However, as we only have data from three sites located in Climate Zone 4A, a fair comparison can't be achieved. Consequently, sites located in Zone 4A are excluded from the following comparative analysis of DF performance.

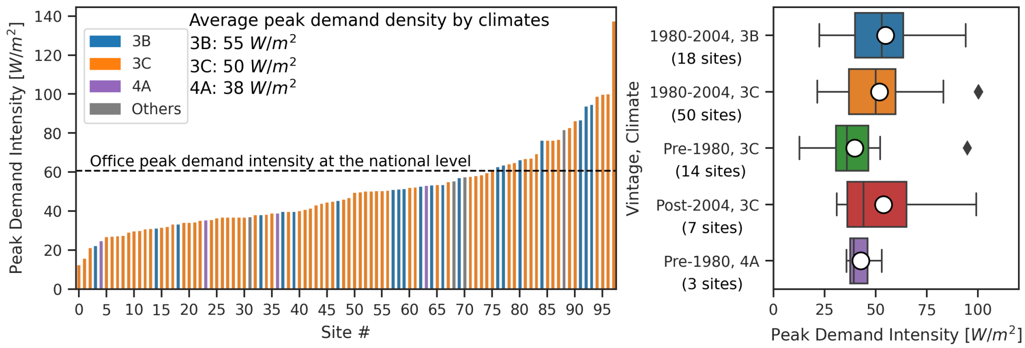


Figure 5: Summary of field demonstration office sites by peak demand intensity

3.2. DR strategies, events and performance

As described in the collected field data, DR related information includes: (1) DR control strategies, (2) DR events (event start date time and end date time, or duration), and (3) DR performance (kW shed) per event for each site. For DR control strategies implemented in commercial office buildings, there are several major groups of control strategies by end-use sectors, including building envelope, HVAC system and plant, lighting, water heater, and plug loads. In this study, we mainly focused on the performance of control strategies such as HVAC, and lighting in the commercial office sector. Table 3 presents a summary of DF control strategies for each specific strategy deployed in field-tested commercial office buildings. The top three DF control strategies are HVAC-A1 (global temperature adjustment [cooling - raise zone temperature by 2°F-6°F]), HVAC-P3 (cycle on/off compressors by 30%, 50%, and 100%), and LTG-A1 (Dimming control [continuous and step 20%-60%]). Furthermore, approximately 12% of field sites have implemented pre-cooling using building thermal mass along with HVAC-A1 control sequence. Reducing static pressure is another common control sequence in HVAC systems with constant airflow distribution, or combined with the HVAC-A1 control sequence for additional fan power savings. A detailed summary of DF control strategies are reported in Appendix A, Table A.1. We created a unique identifier for each control strategy in the data schema, which can also be extended with new control strategies. More details on commercial building control strategies and recommended control sequences for demand response can be found in a previous study [43].

Figure 6 shows the DR performance of office sites sorted by floor area in

Table 3: Summary of DF control strategies in commercial office buildings (note: some buildings implement more than one DF control strategy)

Strategy Group	Strategy Subgroup	Strategy Id	Count of Sites
Building Envelope	Thermal Mass	BE-A1	16
HVAC	Air Distribution	HVAC-A1	**62**
		HVAC-A3	**22**
		HVAC-A6	15
		HVAC-A4	5
	Plant	HVAC-P3	**19**
		HVAC-P2	5
		HVAC-P1	3
Lighting	Interior Lighting	LTG-A1	**19**
MELs (miscellaneous electric loads)	MELs	MEL-A1	2

participating DR events, with DDI ranging from 0 to 32 W/m^2 (0-3 W/ft^2). On average, the DDI of DF metrics are 6.1 W/m^2, 10.0 W/m^2, 11.1 W/m^2, 7.1 W/m^2, and 4.7 W/m^2 for small ($\leq$465 m^2), small-medium (>465 m^2 and $\leq$2,323 m^2), medium (>2,323 m^2 and $\leq$4,645 m^2), medium-large (>4,645 m^2 and $\leq$9,290 m^2), and large office buildings (>9,290 m^2), respectively. It is clear that the DR performance of small and medium office buildings exhibits greater variability compared to that of large office buildings. Large office buildings have the lowest DDI magnitude among small to medium, medium to large, and large office buildings. Given the different DR performance characteristics between small and large office buildings, it can guide us to classify these buildings by different event durations to take full advantage of their maximum DR potential.

For the data schema of DF performance metrics, the other important field is the baseline model option for quantifying the load change during the event. Common baseline model options include an average baseline with or without the morning adjustment (3/10, 5/10, and 10/10), a weather-matching baseline, a weather regression baseline, and more [64]. The average baseline is calculated from either 3, 5, or 10 days with the highest average load during the event period. These days are selected from the previous 10 days of good data (excluding weekends, holidays, a DR event day, and any operation off day). Additionally, The morning adjustment is a ratio of (a) the average load of the first three of four hours before the event to (b) the average load of the same hours from the selected baseline days. The

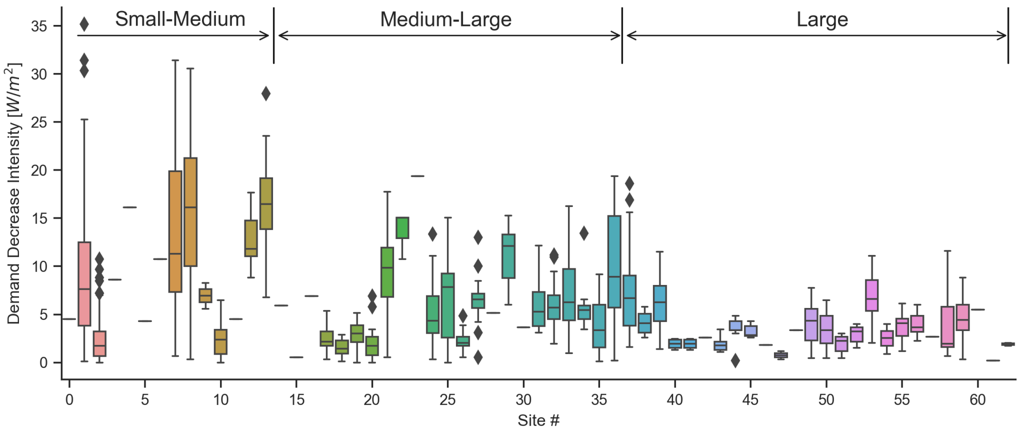


Figure 6: Summary of DR performance by sites with GTA strategy implemented

adjustment factor is limited to ±20% of the customer baseline.

4. Results

4.1. Comparing the simulation results against the field data

Figure 7 shows the comparison of office building peak demand intensity by size and climate. In the plot, the field data has error bars with 95% confidence intervals. It can be seen that small offices have a wider range of peak demand intensities than medium and large offices. For offices in climate 3B, differences between simulated and measured peak demand intensity range from -41% to 31%. In contrast, the difference is relatively small for offices in climate 3C, ranging from 4% to 22%. On the other hand, for all office buildings in climates 3B and 3C, the difference between the simulated and measured peak demand intensities is less than one standard deviation.

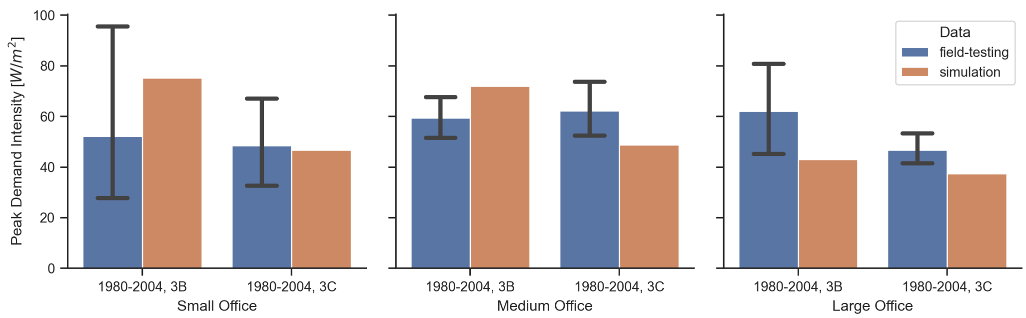


Figure 7: Comparison of office building peak demand intensity by size, vintage, and climate

4.1.1. Comparison by the vintage and climate

As described in section 3.2, global temperature adjustment (GTA) as a standalone strategy or combined with pre-cooling (pre-cooling+GTA) is the primary focus of this comparison study, as this control strategy can serve as load shed and shift service, respectively. Therefore, office sites with GTA only were selected for a fair comparative analysis with prototype building simulation results in this study. On the other hand, the DF metrics used in each field-testing study vary by their aspects of interest in understanding the DF performance. In the recent work [45], we defined a set of DF performance metrics for various grid service products (e.g., shed and shift). However, it is quite a challenge to standardize the field performance data into the same set of metrics with limited access to the field raw data. Therefore, we propose the use of DDI (Demand Decrease Intensity, W/m^2) as the DF metric in the benchmarking of prototype simulation results against the field data.

Figure 8 shows a comparison of DR performance between field tests and simulations for small offices (HVAC type: Packaged Rooftop Unit [RTU]) in ASHRAE climate zone 3C. It is worth noting that we only compared the DR performance of sites located in climate zone 3B, as small office sites in other climates are very limited (only one small office in climate zone 3B implemented the GTA). On the other hand, we selected simulated DR performance on days with a maximum daily OAT above 29.4°C [85°F] to align with field-test DR events. On average per event, the DDI ranged from 4.3 W/m^2 to 16.1 W/m^2. Three sites participated in more than five events over the summer, while the participation of the remaining four sites was unclear (only the average DDI was provided). In comparison with the average DDI across all sites, the simulated average DDI was 1.6 W/m^2 higher than the measured value, which is approximately 120% of the average DDI of the seven available sites.

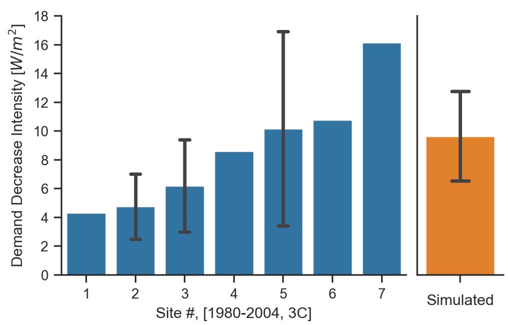
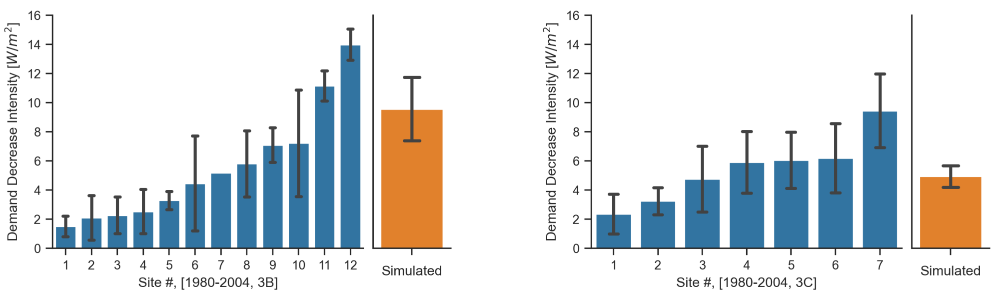


Figure 8: Comparison of DF performance in small offices in ASHRAE climate zone 3C (7 sites, 21 events)

Figure 9 shows a comparison of DR performance between field tests and simulations for medium to large offices (HVAC type: Packaged RTU + VAV) in ASHRAE climate zones 3B (left) and 3C (right). We used the post-1980 prototype medium-size office building model for the comparison and simulated the GTA control strategy for the entire year. This plot compares simulated DDI values with field data under the same range of weather conditions (daily peak OAT). The 12 medium-large office buildings located in climate zone 3B exhibited a wide range of DDI performance, with an average DDI of 4.5 W/m^2 over 80 events. Only two field sites outperformed the prototype simulation model (Avg. DDI 9.6 W/m^2) by approximately 3.1 W/m^2 on average. For sites 1-4, the estimated DDI by the validated simulation model was still 1-4.8 W/m^2 higher than the measured data on average [19]. From the 15-min whole building meter data, the load profile did not show significant load shedding upon the activation of DR control sequences. For medium and medium-large offices in climate zone 3C, both the measured and simulated DDI metrics showed relative smaller variations by events. The simulated average DDI is 1.2 W/m^2 lower than the measured value, which is approximately 80% of the average DDI of the seven available sites. In summary, it is a big challenge to identify abnormal DR performance without additional data such as building automation system (BAS) trend logs and sub-metering data. However, it drives us to collect more field datasets to create an outlier to identify and remove abnormal DR performance data points.

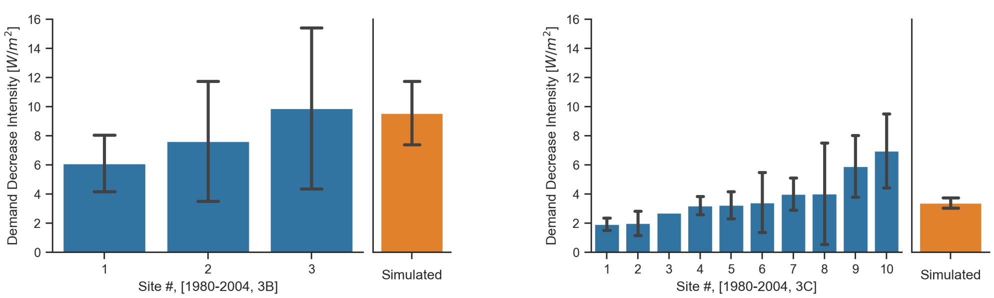


(a) (b)

Figure 9: Comparison of DF performance in medium and medium-large offices in ASHRAE climate zone 3B (left: 12 sites, 86 events) and 3C (right: 7 sites, 63 events

)

Figure 10 shows a comparison of DR performance between field tests and simulations for large offices in ASHRAE climate zones 3B (left) and 3C (right). Although the only three sites located in the climate zone 3B are marked as large offices, the HVAC type is the air-cooled chiller with VAV system. Instead of using large office simulation results, we compared the results with field data in climate zone 3B using the same medium office building model described above. All the sites located in climate zone 3C have water-cooled chillers as their cooling plant. This is the same as in the prototype large office building model. It should be noted that the coefficient of performance (COP) of water-cooled chillers (5.2-6.3) is much higher than that of typical packaged rooftop direct expansion (DX) units found in medium offices (COP 2.8-3.4). As a result, each watt of cooling load reduction in large offices translates into a smaller kilowatt demand shed, despite the additional demand shed from accessory equipment like pumps and cooling towers. The average DDI per field site was about 2-7 W/m^2. In contrast, the simulated DDI was 17% (0.7 W/m^2) lower than the measured value, on average. For large offices located in climate 3B, the simulated DDI was 16% (1.3 W/m^2) higher than the measured value. The results of this comparison provide evidence for the use of prototype building models in DR performance evaluation, especially in large-scale deployments.



(a) (b)

Figure 10: Comparison of DF performance in large offices in ASHRAE climate zones 3B (left: 3 sites, 25 events) and 3C (right: 10 sites, 111 events)

4.2. In-depth analysis of simulation results

Figure 11 shows the results across three representative vintages (1980-2004 (existing buildings constructed in or after 1980), 90.1-2004 [65], and 90.1-2013 [66]) for small, medium, and large office buildings located in the climate zone 3B. The medium office has the greatest DDI (6-10 W/m^2), followed by the small office (3-8 W/m^2), and then the large office (about 3 W/m^2), in that order. The magnitude of DDI decreases with advanced EE measures in newer vintage models, especially for thermally driven loads such as HVAC. A previous study [21] reported similar comparison results of load flexibility between different vintage models. The main reason is the reduced HVAC cooling load from energy efficient upgrades in building envelope, HVAC system, and plant.

To answer differences of DR performance shown in Figure 11 across building types, we conducted an in-depth analysis of GTA impact on the building cooling load.

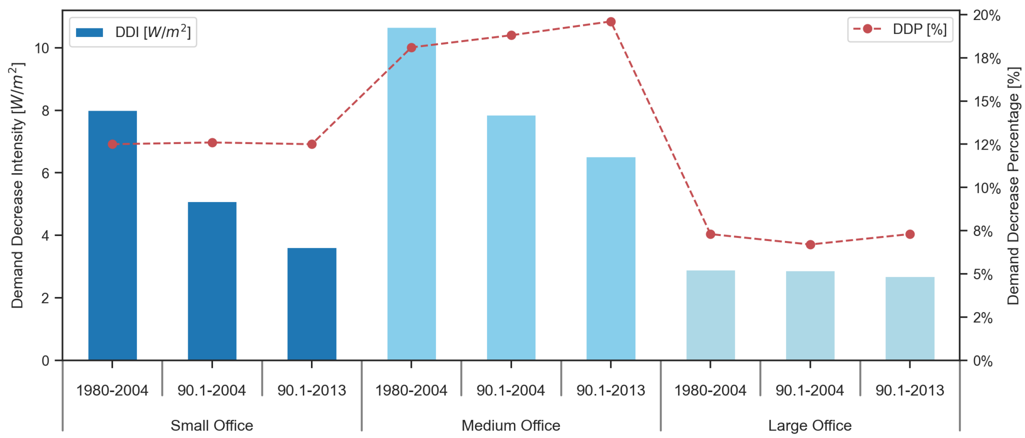


Figure 11: DDI from 2.2°C (4°F) GTA in office buildings (3B climate zone) by building size and vintage

4.2.1. Comparison between medium- and large-size offices

Using our expertise gained from past simulation and field-testing experience, as well as literature reviews, we concluded that there are two main factors that contribute to the lower per-floor-area shed in large office buildings compared to medium office buildings. First, the cooling system efficiency is significantly higher in large offices with water-cooled chillers (coefficient of performance [COP] 5.2-6.3) compared with the packaged rooftop direct expansion (DX) units in medium offices (COP 2.8-3.4). Therefore, each watt of cooling load reduction translates to a smaller kilowatt demand shed in large offices. The second reason is that core zones take a larger proportion of the total floor area in large offices (40%) compared to medium offices (29%), and cooling load reduction from GTA is significantly smaller in core zones compared to perimeter zones in general. One additional nuance here is that the core zone load shed in medium offices (3 floors) is larger than that in large offices (12 floors) due to the fewer number of floors in comparison, because core zones on the top and bottom floors are not truly isolated from thermal conduction with the outdoor environment.

Figure 12 shows that for both medium and large offices, the VAV dampers in the core zones stayed at their minimum positions during the entire GTA period, which means raising the GTA ceiling would not increase load shed; to the contrary, VAV dampers in perimeter zones gradually open above the minimum positions after about an hour into the GTA period, making additional load shed possible with deeper GTA. The blue and orange lines in Figure 12 represent the average VAV damper positions for the VAV in the perimeter and core zones during the 12 hottest weekdays of the year. The solid and dashed lines represent the VAV damper positions for the baseline and "shift" event modes (e.g., "10 am-2 pm precooling"+"2 pm-6 pm shed"). The profile of the core zone VAV damper keeps consistent with internal heat gains from occupants, lighting, and plug loads. In contrast, the zone VAV dampers in the perimeter areas receive additional effects from the outside weather, especially higher solar heat gain and OAT during peak hours. As a result, the large perimeter area offers a greater potential of cooling load reduction from the GTA control strategy. Compared to less core zone area in medium offices, the overall per-floor-area cooling load reduction is significantly smaller in large offices. Therefore, a relatively accurate representation of the perimeter/core zoning is recommended for DF modeling of buildings with VAV systems.

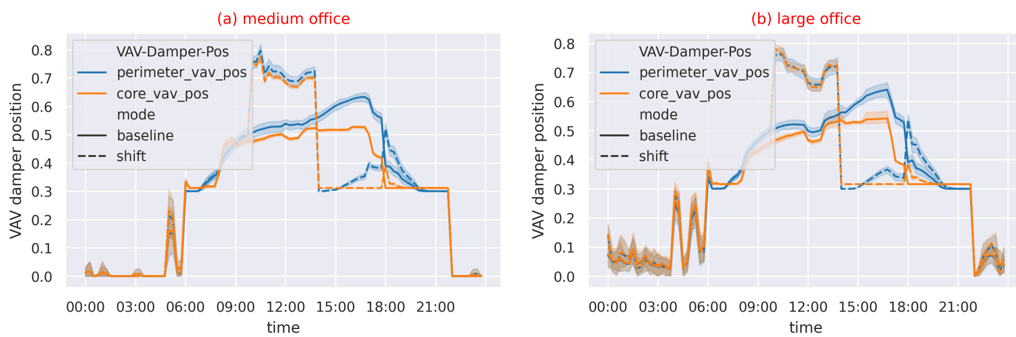


Figure 12: VAV Damper positions in perimeter and core zones under GTA and precooling+GTA in 90.1-2004 medium and large offices in the 3B climate zone (average during the 12 hottest weekdays)

4.2.2. Comparison between small- and medium-size offices

As mentioned in Figure 11, the DDI from 2.2°C (4°F) GTA is lower in small offices compared to medium offices on a per floor area basis. There are many factors that come into play; however, the primary reason for this is that the supply fans in small offices are running at constant speed, whereas medium offices use variable speed fans driven by variable frequency drives (VFDs). As shown in Figure 13, the blue and orange lines represent the cooling and fan power densities on hot days that may dispatch DF events, respectively; the solid and dashed lines represent the power curves for the baseline and “shift” event modes (e.g., “10 am-2 pm precooling”+”2 pm-6 pm shed”). Figure 13 (b) shows that in medium offices, the demand shed from supply fans during the 2-6 pm GTA was about 30% of the total demand shed. In small offices, there was no demand shed from the supply fans, as shown in Figure 13 (a).

Other significant factors that drive the differences between small and medium offices include: (1) the smaller window-to-wall ratios in small offices and (2) the minimum cooling effects with the VAV systems in medium offices. These two factors drive results in opposite directions: factor 1 decreases shed ability, while factor 2 increases shed ability in small offices relative to medium office. However, the constant speed fan is the most important determinant factor that results in DDI difference between small- and medium-size offices, as discussed above.

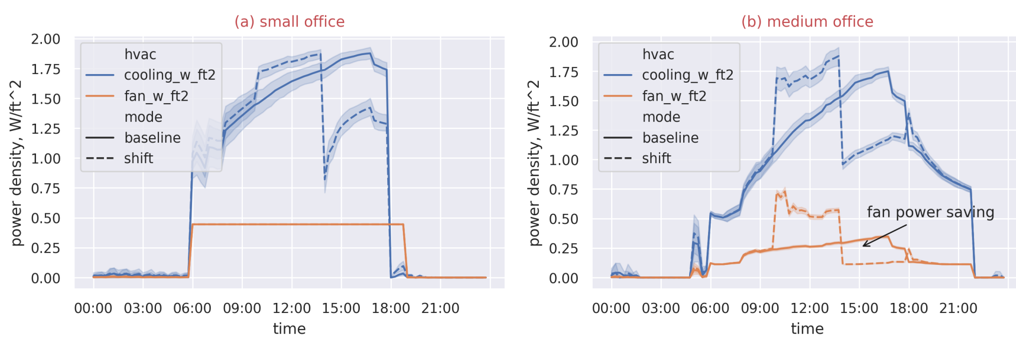


Figure 13: Comparing HVAC demand changes from pre-cooling 1.1°C (2°F) + 2.2°C (4°F) GTA in 90.1-2004 small and medium offices in the 3B climate zone (average during the 12 hottest weekdays)

5. Discussions

5.1. Core data model/schema for compiling DF datasets from various sources

During the field data collection/mapping/cleaning process, there were far more challenges than initially anticipated. First, the publicly available datasets about building demand flexibility are very limited. Second, each study proposes DF performance metrics according to their interests and research goals, which results in inconsistent DF performance metrics reported across all field studies. Third, only very limited site-specific information, such as vintage, is available in published reports or papers. Fourth, not all field sites participate in actual DR program events. Last, events with different start times and durations posed challenges for the consolidation of those datasets for comparative analysis.

A common data schema is essential for compiling data from different data sources. Besides the proposed data schema depicted in Figure 1, it is worth noting that it is necessary to have controlled floor area as a sub-attribute of building floor area, as the DF control strategy may only be implemented for a portion of the site. Among other categories, DF related information such as control strategies and performance metrics are data fields that are unique from existing databases of building energy use [67]. A set of DF performance metrics with a specified baseline model option is acceptable as a simplified core data model defined here. To develop baseline models for quantifying the DF performance, it is suggested that users collect the interval meter data for cooling or heating seasons (based on seasonal events or at least 30-45 weekdays prior to the first event day [64]; even a shorter period of two weeks prior to the event day may be enough [68]). In collaboration with colleagues [21], we summarized a list of common DF control strategies in commercial buildings (see Appendix A). This table can be expanded easily with a new strategy ID for a specific DF control sequence based on the standard data format.

5.2. DF performance influence factors

As we know from our literature review, there are multiple factors that affect DF performance in commercial buildings, including building type and size, building geometry and envelope, HVAC system/plant type and capacity, internal heat gains, building operation schedule, etc. In practice, field measurements show greater variability in similar buildings under the same climatic conditions. However, buildings of the same type and size show similar DF performance by orders of magnitude, regardless of other factors. The field data shows that small & medium and large office buildings have the highest and lowest DDI magnitudes. The simulated results show the same pattern of DDI for small, medium, and large offices in the same climate. In theory, the potential of DF from HVAC end use is determined by the amount of HVAC heating/cooling loads, which indicates that the DF potential of the same building in different climate zones is very similar under the same weather conditions. As described in subsection 4.2, several recommendations are made for the use of prototype building models in demand flexibility analysis, including: (1) a relatively accurate representation of the perimeter/core zoning in medium and large offices, and (2) a precise HVAC controller for the GTA control strategy and cycling RTUs on and off in small and medium offices.

5.3. Comparison between the measured and simulated DF performance datasets

The comparison results suggest that the proposed DF simulation framework can give a reasonable estimate of the magnitude of the mean demand decrease intensity, although the field measured data show a larger variability across different event days and sites. We have very limited site-specific information to understand what led to these differences. However, several factors could have led to greater variability in the field data calculated metrics. These include but are not limited to:

1. Selection of the baseline method and its accuracy. In simulation, baseline is not an issue because consistent building operation can be easily achieved by running the model with the DF strategies disabled. However, with actual buildings, baselines are far more complex because the building's operation without shedding load on the exact same event day cannot be recreated in reality but only emulated with modeling methods. For example, the common "average" baseline method may not represent a good reference when the building's load variability and weather sensitivity are high and the previous week's weather has been greatly different.
2. Significant differences among the sites. In the DF simulation framework, the DOE reference prototype models require various inputs related to building vintage, climate location, and DF strategy details. There are variations in an actual building's geometry, construction,

window-to-wall ratio, thermal mass, internal load, occupancy, HVAC system configuration, efficiency and control settings, and many other building characteristics. These discrepancies between the prototype simulation model and a specific actual building can also lead to significant differences in the results. Another reason is that the associated typical meteorological year data (TMY) weather files with prototype building models may not fully capture the microclimate conditions experienced at the field sites.

3. Uncertainties in the DF control sequences. It was observed that the same DF strategy performed well for the same site on some event days, but worse on other days. It is difficult to diagnose the unpredictable operational behavior in the DF performance due to lack of building operation trend logs. For example, might changes in schedules, occupancy, or setpoints explain the variability? Or did the building operator change the actual setpoint from the plan due to occupant complaints? By contrast, in a lab environment, the uncertainties mentioned above can be minimized or removed, and the impacts from uncertainties can be evaluated and quantified by conducting sensitivity analysis in both simulation and lab testing.
4. Another observation of DF performance in the small office building is that the simulated cooling power profile is relatively smooth over the hours of operation. By contrast, there are significant power fluctuations due to the compressor cycling of packaged constant volume RTUs with either single or two stage compressors, especially for small offices with only one or two packaged RTUs. A field study [47] presents an example of power fluctuations for a small office with constant air volume packaged RTU system. Another contributing factor is the relatively light thermal mass inherent in small offices. It may cause a high degree of fluctuation in DF performance, especially on event days with varying weather conditions. In addition to insufficient thermal mass in small offices, the performance fluctuations may also be caused by leaking ducts or undersized HVAC equipment. To achieve a relatively consistent load shed performance, one alternative method involves coordinating the on/off cycles of each RTU in a periodic order. However, this control strategy requires multiple RTUs on a single site and may not maintain the indoor temperature at the desired level. In terms of load participation in the electricity market, larger office buildings are well-suited for longer grid event participation, providing consistent DR

performance due to the substantial amount of thermal mass from the building envelope and interior furnishings. In contrast, smaller office buildings can contribute their maximum DR capacity by participating in short-term grid events. On the other hand, the aggregation of small office buildings can offer a relatively cost-effective and consistent DR performance, despite unsynchronized variability in each individual building.

6. Conclusions and Future Work

In this study, we collected DR performance data for a total of 831 DR event days from 192 sites as a first step to build a building demand flexibility dataset. We proposed a standard core data model/schema to merge the field DF data from different sources. For the comparison study, our primary focus is on the 97 offices, of which approximately 20% were built before 1980, 73% were constructed between 1980 and 2004, and the remaining 7% were built after 2004. In the second part, we compared the actual performance of 62 office sites (which only implemented GTA) in terms of one DF metric (demand decrease intensity, W/m^2), so we were able to draw several conclusions from the data available to date:

- Demand flexibility by building size. The DDI of DF metrics were 6.1 W/m^2, 10.0 W/m^2, 11.1 W/m^2, 7.1 W/m^2, and 4.7 W/m^2 for small, small-medium, medium, medium-large, and large office buildings, respectively. For medium- and large-size buildings, the simulated DDI was 17% (0.7 W/m^2) lower than the measured value, on average. For large offices located in the 3B climate zone, the simulated DDI was 16% (1.3 W/m^2) higher than the measured value. In general, medium-sized offices can provide the highest DDI with a consistent performance over extended event hours. Large office buildings can independently provide high load shedding kW in the building-to-grid service participation. Small-size offices can shed their loads for a short-term event period. Aggregation in small offices can achieve DF performance similar to medium-size offices in terms of kW capacity and consistency.

- Demand flexibility by building vintage and climate. Specifically, the simulated average DDI was 1.6 W/m^2 higher than the measured value in small offices (1980-2004, climate zone 3C), approximating 120% of the average DDI. In climate zone 3C, medium to medium-large offices exhibited smaller variations in both measured and simulated DDI metrics. The simulated average DDI was 1.2 W/m^2 less than the measured value, representing about 80% of the average DDI across available sites. Medium-large office buildings (1980-2004, climate zone 3B) exhibited diverse DDI performance, averaging 4.5 W/m^2 across 80 events. Only two sites outperformed the prototype simulation model's average DDI of 9.6 W/m^2, by about 3.1 W/m^2. For large office buildings (1980-2004, climate zone 3C), the average DDI per field site was about 2-7 W/m^2.

The simulated DDI was, on average, 17% (0.7 W/m^2) lower than the measured value. In contrast, for large offices located in climate zone 3B, the simulated DDI was 16% (1.3W/m^2) higher than the measured value.

- Climate impacts on peak demand and demand flexibility. Medium-large office buildings located in climate zone 3B exhibited a wider range of DDI performance in comparison with the DF performance in climate zone 3C. In comparing simulated and measured peak demand intensity for offices in a warm and dry climate, discrepancies ranged from -41% to 31%. In contrast, the difference was relatively small for offices in climate zone 3C (warm and marine), ranging from 4% to 22%. On the other hand, for all office buildings in climate zones 3B and 3C, the difference between the simulated and measured peak demand intensities was less than one standard deviation. Another observation is that small offices exhibited a similarly high DDI in the relatively mild climates of 3C, while all field-test event days had a maximum daily OAT above 29.4°C [85°F]. This suggests that the DF performance on hot days in cold climates can serve as a reference for similar buildings in the same weather conditions in hot climates.

- Differences of DF performance across building types. Simulation results indicate that the magnitude of DDI decreases with advanced EE measures in newer vintage models, especially for thermally driven loads such as HVAC. Additionally, we drew several specific conclusions: (a) High-efficiency HVAC systems and plants lead to a smaller kilowatt demand shed for the same amount of cooling load reduction; (b) an accurate representation of the perimeter/core zoning is recommended for DF modeling of buildings with VAV systems; (c) The air-side HVAC system type (e.g., CAV vs. VAV) is the most crucial determinant of the DDI difference between small- and medium-size offices.

- Regarding the utilization of flexible loads as a grid resource, large office buildings are suitable for longer grid event participation while providing consistent DR performance, due to the large amount of thermal mass from the building envelope and interior furnishings. In contrast, small office buildings can participate in short-term grid events while contributing their largest DR capacity. On the other hand, the aggregation of small office buildings can provide relatively cost-effective, consistent, and high DR performance, with unsynchronized variability in each individual buildings.

Data collection efforts in the building demand flexibility area still have a long way to go. Despite the current data limitations in this study, it is the first attempt to compile actual DF data from field-testing sites. This dataset can be used to compare DF performance across different building types, vintages, and climates for identifying DF performance influence factors, or to groups of similar buildings for benchmarking. Additionally, a comparison between the measured and simulated DF performance data can help increase the credibility of DF potential estimates, particularly with respect to the large-scale analysis. With respect to the simulation of building demand flexibility, a few factors have been overlooked in previous simulation studies, such as internal thermal mass level, perimeter/core zoning, HVAC system/plant sizing, and COP.

There are several gaps in the publicly available datasets that can limit the scope of analysis and innovation in this area. Notable gaps include but not limit to: (1) incomplete data (e.g., customer demographics, control strategies, and programs), (2) inconsistent data formats and metrics, (3) insufficient temporal resolution of DF related dataset. Actually, there are available field DF performance datasets under various DR programs managed by each utility and system operators across the national electricity market. It is expected that this data schema can be used by stakeholders such as utilities, aggregators, facility managers, DR program managers, and policymakers to compare a building's DF performance against similar buildings, identify DF opportunities, and estimate DF potential. Furthermore, we expect to identify the drivers of the discrepancy between the field dataset and the simulated DF results. The next step of this work is to promote the need for a national building demand flexibility performance database between stakeholders, with the goal of building electrification/decarbonization.

Acknowledgements

This work was supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, Building Technologies Office, of the U.S. Department of Energy under Contract No. DE-AC02–05CH11231. The authors want to thank Ms. Monica Neukomm and Dr. Amir Roth for funding this research work. We also wish to acknowledge the Pacific Gas and Electric company included in this study for providing valuable DR performance data and sharing their experiences and insights to support our DF benchmarking research.

Appendix A. DF control strategies

Table A.1 summarizes the recommended DF control sequences in commercial buildings by each group of DF control strategies.

Table A.1: Summary of DF control strategies in commercial buildings

End-use Category	**Subcategory**	**DF Control Strategy**	Level of Response		
			Low (1)	Medium (2)	High (3)
Building envelope	Exterior/interior wall and internal thermal mass	Passive thermal mass storage (start HVAC system earlier than normal operation for pre-cooling)	2 hours	4 hours	6 hours
		Active thermal mass storage (pre-cooling phase change materials by 2-6ºF)	2ºF	4ºF	6ºF
		Night flushing/economizer (pre-cooling)	N/A	N/A	N/A
	Smart windows	Blind control to reduce the cooling load while maintaining the daylight level	N/A	N/A	N/A
		Electrochromic - control thermal performance to reduce the cooling load while maintaining the daylight level	N/A	N/A	N/A
HVAC Systems	VAV	Global temperature adjustment (cooling - raise zone temperature by 2ºF-6ºF)	2ºF	4ºF	6ºF
		Global temperature adjustment (heating - Decrease thermostat heating setpoint by 2ºF-4ºF)	2ºF	3ºF	4ºF
		Duct static pressure decrease from 1.5" to 1.0"	1.4"	1.2"	1.0"
		Supply air temperature increased from 55ºF to 65ºF	2ºF	6ºF	10ºF
		Limit AHU cooling valve position to 70%	90%	80%	70%
		Limit VFD fans and pumps speed to 70%	90%	80%	70%
	CAV	Supply air temperature increased from 55ºF to 65ºF	N/A	N/A	N/A
		Lock cooling valve position at the AHU	90%	80%	70%
HVAC Plant	Water/air-cooled chiller	Chilled water temperature reset (increase 5ºF)	2ºF	4ºF	6ºF
		Chiller demand limit to 50%-90%	90%	70%	50%
	Packaged RTU	Cycle on/off compressors by 30%, 50% and 100%	30%	50%	100%
	Partial TES system	Shut off 1/3 or 1/2 of multiple chillers	N/A	1/3	1/2
Lighting	Interior lighting	Dimming control (continuous and step 20%-60%)	20%	40%	60%
		Switch on/off	N/A	N/A	100%
	Task lighting	Switch on/off	N/A	N/A	100%
Water Heater	Electric	Setpoint adjustment (Decrease water temperature setpoint by 5-15ºF from 120ºF)	5ºF	10ºF	15ºF
		Setpoint adjustment (Decrease water temperature setpoint by 5ºF-15ºF from 120ºF) with preheat (increase water temperature by 10ºF)	5ºF	10ºF	15ºF
		Switch on/off	N/A	N/A	100%
	Heat pump	Setpoint adjustment (Decrease water temperature setpoint by 5ºF-15ºF from 120ºF)	5ºF	10ºF	15ºF
		Reduce deadband for heat pump to 1ºF	1ºF	1ºF	1ºF
		Limit heat pump duty cycling (0-100%)	30%	20%	10%
MELs (miscellaneous electric loads)	Non-critical process load	Stand-by equipment load reduction	N/A	N/A	100%

References

[1] DOE, A National Roadmap for Grid-Interactive Efficient Buildings, Tech. rep., U.S. Department of Energy (DOE) Building Technologies Office (2021).

[2] P. Alstone, J. Potter, M. A. Piette, P. Schwartz, M. A. Berger, L. N. Dunn, S. J. Smith, M. D. Sohn, A. Aghajanzadeh, S. Stensson, J. Szinai, T. Walter, L. Mckenzie, L. Lavin, B. Schneiderman, A. Mileva, E. Cutter, A. Olson, J. Bode, A. Ciccone, A. Jain, Final Report on Phase 2 Results - 2025 California Demand Response Potential Study, Tech. rep., Lawrence Berkeley National Laboratory (2017).

[3] FERC, Order No. 745: Demand Response Compensation in Organized Wholesale Energy Markets, Tech. rep., Federal Energy Regulatory Commission (2011).

[4] FERC, 2021 Assessment of Demand Response and Advanced Metering Staff Report, Tech. rep., Federal Energy Regulatory Commission (2021).

[5] J. E. Braun, Reducing Energy Costs and Peak Electrical Demand through Optimal Control of Building Thermal Storage, ASHRAE Transactions 96 (2) (1990).
URL http://eta-publications.lbl.gov/sites/default/files/dbc-paper90.pdf

[6] A. Rabl, L. K. Norford, Peak load reduction by preconditioning buildings at night, International Journal of Energy Research 15 (9) (1991) 781–798. doi:10.1002/ER.4440150909.

[7] F. B. Morris, J. E. Braun, S. J. Treado, Experimental and simulated performance of optimal control of building thermal storage, ASHRAE Transactions 100 (1) (1994) 402–414. doi:10.2/JQUERY.MIN.JS.

[8] K. Keeney, J. Braun, A Simplified Method for Determining Optimal Cooling Control Strategies for Thermal Storage in Building Mass, HVACR Research 2 (1) (1996) 59–78. doi:10.1080/10789669.1996.10391333.
URL http://www.tandfonline.com/doi/abs/10.1080/10789669.1996.10391333

[9] J. E. Braun, K. W. Montgomery Nitin Chaturvedi, Evaluating the Performance of Building Thermal Mass Control Strategies, HVACR

Research 7 (4) (2001) 403.
URL http://eta-publications.lbl.gov/sites/default/files/409braun-october.pdf

[10] K.-H. Lee, J. E. Braun, R. W. Herrick, Development and Application of an Inverse Building Model for Demand Response in Small Commercial Buildings, in: SimBuild 2004 IBPSA-USA National Conference, Boulder, CO, 2004.
URL http://eta-publications.lbl.gov/sites/default/files/ibpsa-usa-simbuild.pdf

[11] R. Yin, E. C. Kara, Y. Li, N. DeForest, K. Wang, T. Yong, M. Stadler, Quantifying flexibility of commercial and residential loads for demand response using setpoint changes, Applied Energy 177 (2016) 149–164. doi:10.1016/j.apenergy.2016.05.090.
URL http://linkinghub.elsevier.com/retrieve/pii/S0306261916306870

[12] P. Xu, P. Haves, M. A. Piette, Peak Demand Reduction from Pre-Cooling with Zone Temperature Reset in an Office Building, in: 2004 ACEEE Summer Study on Energy Efficiency in Buildings, Pacific Grove, CA, 2004.

[13] K.-H. Lee, J. E. Braun, A Data-Driven Method for Determining Zone Temperature Trajectories that Minimize Peak Electrical Demand, ASHRAE Transactions 114 (2) (2008).
URL https://gig.lbl.gov/sites/all/files/lee-braun-data-driven.pdf

[14] K.-H. Lee, J. E. Braun, Evaluation of methods for determining demand-limiting setpoint trajectories in buildings using short-term measurements, Building and Environment 43 (10) (2008) 1769–1783. doi:10.1016/J.BUILDENV.2007.11.003.
URL https://www.sciencedirect.com/science/article/pii/S0360132307002144

[15] P. Xu, P. Haves, M. Piette, L. Zagreus, Demand Shifting With Thermal Mass in Large Commercial Buildings: Field Tests, Simulations and Audits, Tech. rep. (2005).

[16] Z. Jiang, J. Peng, R. Yin, M. Hu, J. Cao, B. Zou, Stochastic modelling of flexible load characteristics of split-type air conditioners using grey-box modelling and random forest method, Energy and Buildings 273 (2022) 112370. doi:10.1016/J.ENBUILD.2022.112370.

[17] J. Song, J. Peng, J. Cao, R. Yin, Y. He, B. Zou, W. Zhao, Global sensitivity analysis of fan coil air conditioning demand response—a case study of medium-sized office buildings, Applied Thermal Engineering 230 (2023) 120721. doi:10.1016/J.APPLTHERMALENG.2023.120721.

[18] G. P. Henze, D. E. Kalz, S. Liu, C. Felsmann, Experimental analysis of model-based predictive optimal control for active and passive building thermal storage inventory, HVAC and R Research 11 (2) (2005) 189–213. doi:10.1080/10789669.2005.10391134.

[19] R. Yin, S. Kiliccote, M. A. Piette, K. Parrish, Scenario Analysis of Peak Demand Savings for Commercial Buildings with Thermal Mass in California Description of Prototypical Models, in: ACEEE 2010 Summer Study on Energy Efficiency in Buildings, no. May, Pacific Grove, California, USA, 2010, pp. 374–387.

[20] S. Huang, Y. Ye, D. Wu, W. Zuo, An assessment of power flexibility from commercial building cooling systems in the United States, Energy 221 (2021) 119571. doi:10.1016/J.ENERGY.2020.119571.

[21] N. Luo, J. Langevin, H. Chandra-Putra, S. H. Lee, Quantifying the effect of multiple load flexibility strategies on commercial building electricity demand and services via surrogate modeling, Applied Energy 309 (mar 2022). doi:10.1016/j.apenergy.2021.118372.

[22] J. Cai, J. E. Braun, Laboratory-based assessment of HVAC equipment for power grid frequency regulation: Methods, regulation performance, economics, indoor comfort and energy efficiency, Energy and Buildings 185 (2019) 148–161. doi:10.1016/j.enbuild.2018.12.022.

[23] A. L. Hjortland, J. E. Braun, Load-based testing methodology for fixed-speed and variable-speed unitary air conditioning equipment, Science and Technology for the Built Environment 25 (2) (2019) 233–244. doi:10.1080/23744731.2018.1520564.
URL `https://www.tandfonline.com/doi/full/10.1080/23744731.2018.1520564`

[24] P. Xu, L. Zagreus, Demand Shifting With Thermal Mass in Light and Heavy Mass Commercial Buildings, Tech. rep., LBNL-61172, Lawrence Berkeley National Laboratory (2007).
URL `http://eta-publications.lbl.gov/sites/default/files/61172.pdf`

[25] R. Yin, P. Xu, S. Kiliccote, S. California Edison, Auto-DR and Pre-cooling of Buildings at Tri-City Corporate Center, Tech. rep. (2008).

[26] K. R. Keeney, J. E. Braun, Application of Building Precooling to Reduce Peak Cooling Requirements, ASHRAE Transactions 114 (2) (2008) 75–84.
URL `http://eta-publications.lbl.gov/sites/default/files/ameritech-96.pdf`

[27] J. Jihyun Kim, R. Yin, S. Kiliccote, Automated Price and Demand Response Demonstration for Large Customers in New York City using OpenADR, in: International Conference for Enhanced Building Operations, LBNL-6472E, Montreal, Quebec, 2013.
URL `http://eta-publications.lbl.gov/sites/default/files/lbnl-6472e.pdf`

[28] M. A. Piette, R. Kukulka, R. L. Therkelsen, Development and Evaluation of Fully Automated Demand Response in Large Facilities, Tech. rep., LBNL-55085, Lawrence Berkeley National Laboratory (2005).
URL `http://www.energy.ca.gov/research/index.html`

[29] M. A. Piette, D. S. Watson, N. Motegi, S. Kiliccote, E. Linkugel, Automated Demand Response Strategies and Commissioning Commercial Building Controls, in: 14th National Conference on Building Commissioning, LBNL-61013, San Francisco, CA, 2006.
URL `http://eta-publications.lbl.gov/sites/default/files/61013 scan.pdf`

[30] M. A. Piette, D. Watson, N. Motegi, S. Kiliccote, M. A. Piette, D. Watson, N. Motegi, S. Kiliccote, Automated Critical Peak Pricing Field Tests: 2006 Pilot Program Description and Results, Tech. rep., LBNL-62218. Lawrence Berkeley National Laboratory (2007).
URL `http://eta-publications.lbl.gov/sites/default/files/62218.pdf`

[31] M. A. Piette, G. Ghatikar, S. Kiliccote, D. Watson, E. Koch, D. H. Akuacom, Design and Operation of an Open, Interoperable Automated Demand Response Infrastructure for Commercial Buildings, Journal of Computing Science and Information Engineering 9 (2) (2009).
URL `http://eta-publications.lbl.gov/sites/default/files/lbnl-2340e.pdf`

[32] H. Li, Z. Wang, T. Hong, M. A. Piette, Energy flexibility of residential buildings: A systematic review of character-

ization and quantification methods and applications (aug 2021). doi:10.1016/j.adapen.2021.100054.

[33] Z. Nagy, G. Henze, S. Dey, J. Arroyo, L. Helsen, X. Zhang, B. Chen, K. Amasyali, K. Kurte, A. Zamzam, H. Zandi, J. Drgoňa, M. Quintana, S. McCullogh, J. Y. Park, H. Li, T. Hong, S. Brandi, G. Pinto, A. Capozzoli, D. Vrabie, M. Bergés, K. Nweye, T. Marzullo, A. Bernstein, Ten questions concerning reinforcement learning for building energy management, Building and Environment 241 (2023) 110435. doi:10.1016/J.BUILDENV.2023.110435.

[34] J. Cai, J. E. Braun, Assessments of demand response potential in small commercial buildings across the United States Assessments of demand response potential in small commercial buildings across the United States, Science and Technology for the Built Environment 0 (0) (2019) 1–18. doi:10.1080/23744731.2019.1629245.
URL https://doi.org/10.1080/23744731.2019.1629245

[35] J. Langevin, C. B. Harris, A. Satre-Meloy, H. Chandra-Putra, A. Speake, E. Present, R. Adhikari, E. J. H. Wilson, A. J. Satchwell, Us building energy efficiency and flexibility as an electric grid resource (2021). doi:10.1016/j.joule.2021.06.002.
URL https://doi.org/10.1016/j.joule.2021.06.002

[36] Y. Chen, Z. Deng, T. Hong, Automatic and rapid calibration of urban building energy models by learning from energy performance database, Applied Energy 277 (2020) 115584. doi:10.1016/J.APENERGY.2020.115584.

[37] M. Mosteiro-Romero, C. Miller, A. Chong, R. Stouffs, Elastic buildings: Calibrated district-scale simulation of occupant-flexible campus operation for hybrid work optimization, Building and Environment 237 (2023) 110318. doi:10.1016/J.BUILDENV.2023.110318.
URL http://arxiv.org/abs/2210.06124

[38] Z. Li, J. Zhang, Data-oriented distributed overall optimization for large-scale hvac systems with dynamic supply capability and distributed demand response, Building and Environment 221 (2022) 109322. doi:10.1016/J.BUILDENV.2022.109322.

[39] H. Li, H. Johra, F. de Andrade Pereira, T. Hong, J. Le Dréau, A. Maturo, M. Wei, Y. Liu, A. Saberi-Derakhtenjani, Z. Nagy, A. Marszal-Pomianowska, D. Finn, S. Miyata, K. Kaspar, K. Nweye, Z. O'Neill, F. Pallonetto, B. Dong, Data-driven key performance indicators and datasets for building energy flexibility: A review and perspectives, Applied Energy 343 (2023) 121217. doi:https://doi.org/10.1016/j.apenergy.2023.121217.
URL https://www.sciencedirect.com/science/article/pii/S0306261923005810

[40] R. Li, A. J. Satchwell, D. Finn, T. H. Christensen, M. Kummert, J. L. Dréau, R. A. Lopes, H. Madsen, J. Salom, G. Henze, K. Wittchen, Ten questions concerning energy flexibility in buildings, Building and Environment 223 (2022) 109461. doi:10.1016/J.BUILDENV.2022.109461.

[41] J. Liu, R. Yin, L. Yu, M. A. Piette, M. Pritoni, A. Casillas, J. Xie, T. Hong, M. Neukomm, P. Schwartz, Defining and applying an electricity demand flexibility benchmarking metrics framework for grid-interactive efficient commercial buildings, Advances in Applied Energy 8 (2022) 100107. doi:10.1016/J.ADAPEN.2022.100107.

[42] M. Deru, K. Field, D. Studer, K. Benne, B. Griffith, P. Torcellini, B. Liu, M. Halverson, D. Winiarski, M. Rosenberg, M. Yazdanian, J. Huang, D. Crawley, U.s. department of energy commercial reference building models of the national building stock (2011).

[43] N. Motegi, M. Piette, D. Watson, S. Kiliccote, P. Xu, Introduction to Commercial Building Control Strategies and Techniques for Demand Response, Tech. rep., LBNL-59975, Lawrence Berkeley National Laboratory (2007).
URL http://eta-publications.lbl.gov/sites/default/files/59975.pdf

[44] R. Guglielmetti, D. Macumber, N. Long, Openstudio: An open source integrated analysis platform, 2011.
URL http://www.osti.gov/bridge

[45] J. Liu, R. Yin, M. Pritoni, M. A. Piette, M. Neukomm, Developing and Evaluating Metrics for Demand Flexibility in Buildings: Comparing Simulations and Field Data Energy Technologies Area, in: ACEEE 2020 Summer Study on Energy Efficiency in Buildings VIRTUAL, 2020. doi:10.20357/B7WW34.

[46] A. Chiu, G. Wikler, I. Bran, J. Prijyanonda, S. Yoshida, K. Smith, M. A. Piette, S. Kiliccote, G. Ghatikar, C. Thomas, 2007 Auto-DR Program, Task 13 Deliverable, Auto-DR Assessment Study, Tech. rep., PGE (2007).
URL www.auto-dr.com

[47] J. Page, S. Kiliccote, H. Dudley, M. A. Piette, A. K. Chiu, B. Kellow, E. Koch, P. Lipkin, Automated Demand Response Technology Demonstration for Small and Medium Commercial Buildings, Tech. rep., LBNL-4982E. Lawrence Berkeley National Laboratory (2011).
URL http://eta-publications.lbl.gov/sites/default/files/LBNL-4982E.pdf

[48] G. Ghatikar, V. Ganti, N. Matson, M. A. Piette, Demand Response Opportunities and Enabling Technologies for Data Centers: Findings from Field Studies, Tech. rep., LBNL-5763E. Lawrence Berkeley National Laboratory (2012).
URL http://eta-publications.lbl.gov/sites/default/files/LBNL-5763E.pdf

[49] P. Xu, R. Yin, C. Brown, D. E. Kim, Demand Shifting with Thermal Mass in Large Commercial Buildings in a California Hot Climate Zone, Tech. rep., LBNL-3898E. Lawrence Berkeley National Laboratory (2009).
URL http://eta-publications.lbl.gov/sites/default/files/lbnl-3898e.pdf

[50] P. Xu, Evaluation of Demand Shifting with Thermal Mass in Two Large Commercial Buildings, in: Proceedings of SimBuild 2006, LBNL-2907E, Lawrence Berkeley National Laboratory, Cambridge, MA, 2006.
URL http://eta-publications.lbl.gov/sites/default/files/lbnl-2907e.pdf

[51] A. Schwarzenegger, Automation of Capacity Bidding with an Aggregator Using Open Automated Demand Response, Tech. rep., LBNL-4383E. Lawrence Berkeley National Laboratory (2008).
URL http://eta-publications.lbl.gov/sites/default/files/cec-500-2008-059.pdf

[52] J. Granderson, J. H. Dudley, S. Kiliccote, M. A. Piette, Chilled Water Storage System and Demand Response at the University of California at Merced, in: 9th International Conference for Enhanced Building Operations, LBNL-2753E. Lawrence Berkeley National Laboratory, Austin, TX, 2009.
URL http://eta-publications.lbl.gov/sites/default/files/lbnl-2753e.pdf

[53] S. Kiliccote, M. A. Piette, G. Ghatikar, E. Koch, D. Hennage, A. J. Hernandez, A. Chiu, O. Sezgen, J. Goodin, S. Kiliccote, M. A. Piette, G. Ghatikar, L. Berkeley, E. Koch, D. H. Akuacom, J. Hernandez, A. Chiu, O. Sezgen, J. Goodin California, Open Automated Demand Response Communications in Demand Response for Wholesale Ancillary Services, in: Grid-Interop Forum 2009, LBNL-2945E, Denver, CO, 2009. URL http://eta-publications.lbl.gov/sites/default/files/lbnl-2945e.pdf

[54] S. Kiliccote, S. Lanzisera, A. Liao, O. Schetrit, M. A. Piette, Fast DR: Controlling Small Loads over the Internet, in: ACEEE Summer Study on Energy Efficiency Buildings, LBNL-185726, Pacific Grove, CA, 2014. URL http://eta-publications.lbl.gov/sites/default/files/fast$_dr_controlling_smallo$

[55] J. J. Kim, R. Yin, S. Kiliccote, Automated Demand Response Technologies and Demonstration in New York City using OpenADR, Tech. rep., LBNL-6470E. Lawrence Berkeley National Laboratory (2013). URL http://eta-publications.lbl.gov/sites/default/files/lbnl-6470e.pdf

[56] T. Hotchi, A. T. Hodgson, W. J. Fisk, Indoor Air Quality Impacts of a Peak Load Shedding Strategy for a Large Retail Building, Tech. rep., LBNL-59293, Lawrence Berkeley National Laboratory (2006). URL http://www.energy.ca.gov/research/index.html

[57] K.-H. Lee, J. E. Braun, An Experimental Evaluation of Demand Limiting Using Building Thermal Mass in a Small Commercial Building, ASHRAE Transactions 112 (1) (2006) 559–571.

[58] K. ho Lee, J. E. Braun, Model-based demand-limiting control of building thermal mass, Building and Environment 43 (10) (2008) 1633–1646. doi:10.1016/j.buildenv.2007.10.009.

[59] J. Han, M. A. Piette, S. Kiliccote, Field Test Results of Automated Demand Response in a Large Office Building, in: 8th International Conference on EcoBalance, LBNL-1131e, Tokyo, Japan, 2008. URL http://eta-publications.lbl.gov/sites/default/files/lbnl-1131e.pdf

[60] M.-a. Leduc, K. Lavigne, A. Poulin, Demand Reponse in Commercial Buildings : a Cold Climate Field Study (2015) 1–8.

[61] S. Prabhu, S. Trikande, O. Sheik, R. Gadh, M. Fung, F. Pina, R. P. Oglesby, Implementation of Demand Response in a University Campus,

Tech. rep., California Energy Commission. CEC-500-2015-087. (2015). URL http://smartgrid.ucla.edu/

[62] A. Keskar, S. Lei, T. Webb, S. Nagy, I. A. Hiskens, J. L. Mathieu, J. X. Johnson, Assessing the performance of global thermostat adjustment in commercial buildings for load shifting demand response, Environmental Research: Infrastructure and Sustainability 2 (1) (2022) 015003. doi:10.1088/2634-4505/ac51c5.

[63] U.S.EIA, 2012 commercial buildings energy consumption survey (cbecs) data (2016).
URL https://www.eia.gov/consumption/commercial/data/2012/

[64] M. L. Goldberg, G. Kennedy, Measurement and Verification for Demand Response, Tech. rep., Federal Energy Regulatory Commission (2013).
URL https://www.ferc.gov/industries/electric/indus-act/demand-response/dr-pot

[65] ANSI/ASHRAE/IES, ANSI/ASHRAE/IES 90.1-2004, Energy Standard for Buildings Except Low-Rise Residential Buildings, Tech. rep., American Society of Heating, Refrigerating and Air-Conditioning Engineers, Atlanta, Georgia (2004).

[66] ANSI/ASHRAE/IES, ANSI/ASHRAE/IES 90.1-2013, Energy Standard for Buildings Except Low-Rise Residential Buildings, Tech. rep., American Society of Heating, Refrigerating and Air-Conditioning Engineers, Atlanta, Georgia (2013).

[67] P. A. Mathew, L. N. Dunn, M. D. Sohn, A. Mercado, C. Custudio, T. Walter, Big-data for building energy performance: Lessons from assembling a very large national database of building energy use, Applied Energy 140 (2015) 85–93. doi:10.1016/J.APENERGY.2014.11.042.

[68] J. Liu, L. Yu, R. Yin, M. A. Piette, M. Pritoni, A. Casillas, M. Neukomm, A. Roth, Benchmarking Demand Flexibility in Commercial Buildings and Flattening the Duck – Addressing Baseline and Commissioning Challenges, in: ACEEE 2022 Summer Study on Energy Efficiency in Buildings, 2022. doi:10.20357/B7M89Q.