Building Technology Department
Building Technology & Urban Systems Division
Lawrence Berkeley National Laboratory

Factors Influencing Building Demand Flexibility

Jingjing Liu, Lili Yu, Rongxin Yin, Mary Ann Piette, Marco Pritoni, Armando Casillas, Peter Schwartz

Energy Technology Area
December 2023

doi:10.20357/B71882

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

Factors Influencing Building Demand Flexibility

Principal Authors
Jingjing Liu
Lili Yu
Rongxin Yin
Mary Ann Piette
Marco Pritoni
Armando Casillas
Peter Schwartz

Lawrence Berkeley National Laboratory
1 Cyclotron Road, MS 90R3147
Berkeley CA 94720-8136

September 2023

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 Mr. Amir Roth for funding this research work. We also wish to acknowledge the owner organization of the retail buildings included in this study for providing valuable DR performance data and sharing their experiences and insights to support our DF benchmarking research. FLEXLAB® at Lawrence Berkeley National Lab's component of 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 thank the following experts for reviewing this report (affiliations do not imply that those organizations support or endorse this work):

Tianzhen Hong	Lawrence Berkeley National Laboratory
Eliot Crowe	Lawrence Berkeley National Laboratory
Hannah Kramer	Lawrence Berkeley National Laboratory
Robert Nirenberg	EduControls LLC
Theodore Wilinski	Milwaukee Area Technical College

Table of Contents

Table of Figures

List of Tables

Acronyms and Abbreviations

DDI	Demand Decrease Intensity
DDP	Demand Decrease Percentage
DDSD	Demand Decrease Intensity Standard Deviation
DERs	Distributed Energy Resources
DF	Demand Flexibility
DII	Demand Increase Intensity
DIP	Demand Increase Percentage
DISD	Demand Decrease Intensity Standard Deviation
DOE	U.S. Department of Energy
DR	Demand Response
EE	Energy Efficiency
EF	Energy Flexibility
EUI	Energy Use Intensity
EV	Electric Vehicle
GEB	Grid-Interactive Efficient Buildings
GTA	Global Temperature Adjustment
HVAC	Heating, Ventilation, and Air Conditioning
LF	Load Flexibility
OAT	Outside Air Temperature
PV	Photovoltaic
RTU	Rooftop Unit
SPDI	Summer Peak Demand Intensity
VAV	Variable Air Volume
WPDI	Winter Peak Demand Intensity

Executive Summary

The U.S. Department of Energy's (DOE) published A National Roadmap for Grid-interactive Efficient Buildings (GEB) in 2021, which acknowledged that building demand flexibility (DF) is both an important strategy to decarbonizing the buildings sector and an important resource for meeting the changing needs of the electrical grid such as improving grid reliability. Large-scale simulation research done as part of the DOE GEB Roadmap has shown that there is a tremendous amount of environmental and economic value associated with adopting GEB technologies and demand management strategies including demand flexibility such as load shedding and shifting. However, understanding the complexity and uncertainties in real building field performance of DF strategies is a large gap hindering stakeholders on both grid and buildings side to make investments on deploying such strategies. This gap exists despite the fact that, if well implemented, building DF can be a more economical resource than other strategies such as behind-the-meter batteries. The GEB Roadmap pointed out the need for metrics and significant data sets to support benchmarking of the field performance of DF technologies and strategies.

Acknowledging this important gap and potential, the "Framework to Define Flexible Loads in Grid-interactive Efficient Buildings" project funded by the DOE conducted research to advance understanding of the variability and influential factors in building demand flexibility. Adding such knowledge based on lab testing results and measured performance data from real buildings is an important and unique contribution of this project. The report uses field-measured DF datasets for two significant building groups of big-box retail and medium office buildings to present the challenge of building DF variability in multiple dimensions including differences among similar buildings and performance variations of the same building over time. This demonstrated the need for establishing standardized metrics and methods to quantify real world DF performance and understand correlations with key factors. We present the concept of "building demand flexibility benchmarking" and have previously published the benchmarking framework, which is summarized and employed in this report. A unique contribution of this framework is its ability to accommodate the varying grid needs across geographic and time dimensions such that it allows the users to extract the most value from their datasets.

This report presents findings related to how several key factors influence building demand flexibility from implementing a common, cost-effective DF control strategy (i.e., adjusting zone temperatures). The findings are supported by full-scale lab testing, field data analysis and simulation research. Three categories were used to orient the readers around these key factors that are static, varying by event, or stochastic. The factors listed below are not meant to represent an exhaustive list of all factors influencing DF, which is a complex topic. The complexity in creating accurate, practical counterfactual baselines in itself is an example of the challenges, which is a basis for understanding correlations in DF performance and influential factors. Through our research work, we summarize a preliminary understanding of the key factors shown in **Table ES-1** below.

Table-ES 1. Summary of Findings on Key Factors' Influence on Building Demand Flexibility

Weather	Event Timing	HVAC System Type	Building Type
Temperature: positive correlation with shed but with significant noise; Humidity: no clear correlation found; Solar Radiation: cloudy condition reduces shed.	Duration: shed decreases significantly for longer events; Start time: evening DF events yield lower shed compared to afternoon.	Single-zone RTU systems tend to yield higher shed compared to multi-zone VAV systems.	Many parameters at play: building size, construction, space use, window-to-wall ratio, internal load, HVAC type, zoning, equipment efficiency, operating hours, and more.
Building Vintage & Efficiency	Thermal Mass	Operations & Commissioning	
Many parameters are at play. Majority of efficiency measures reduce shed but some can increase it while others are neutral.	Load shift with pre-cooling can be energy efficient and improves comfort but may not significantly increase shed. Warm room temperature overnight may reduce shed significantly.	Can have profound impact; case-by-case. E.g., HVAC equipment not functioning properly or commissioning issues related to hardware or controls.	

In presenting the above findings from our research, we gave attention to real world implications and applications around building demand flexibility and demand response programs. Therefore, this report intends to benefit not only building researchers and consultants but also other stakeholders in these programs. The report provided application-oriented recommendations to stakeholders such as building aggregators, utility program design professionals, sophisticated building portfolio owners, and more. These recommendations span topics such as the following:

- The order of magnitude of load flexibility in certain building and HVAC system types;
- Potential synergies between energy efficiency and DF strategies;
- Circumstances under which DF performance may reduces and possible mitigation for under-performance; and
- Considerations in applying pre-cooling.

Although the underlying research is the same, different stakeholders may find different real world values from these recommendation themes. The findings presented in the report demonstrated the breadth and depth of this topic as well as the need for future research work.

1. Demand Flexibility is a High-potential Grid Resource.

1.1 Purpose

This report fulfills the final report of the "Framework to Define Flexible Loads in Grid-interactive Efficient Buildings" project (hereinafter "project") funded by the U.S. Department of Energy (DOE). This project conducted research work advancing the understanding of the variability and influential factors in building demand flexibility. Adding knowledge based on lab testing results and measured performance data from real buildings is an important and unique contribution of this project. In presenting the key findings from this research work, we gave attention to real world implications and applications around building demand flexibility and demand response programs. Therefore, this report intends to benefit not only building researchers and consultants but also stakeholders in this type of programs such as building resource aggregators, utility program design professionals, sophisticated building portfolio owners, and more.

1.2 DOE's GEB Initiative

Buildings account for more than 70% of U.S. electricity use and at least one-third of U.S. economy-wide CO2 emissions. Therefore, improving the way that electricity is consumed in buildings will be essential to a decarbonized economy [1]. In addition, the increasing penetration of renewable energy in the generation mix and proliferation of behind-the-meter distributed energy resources (DERs) has significantly increased dynamics and uncertainty to grid balancing [2], which underlined the importance of research on managing building demand as a grid resource. In 2019, the DOE launched a Grid-Interactive Efficient Buildings (GEB)[1] initiative that aims to optimize DERs for the benefit of building owners, occupants, and the electric grid [3]. Energy- efficient and flexible are two important characteristics of GEBs in addition to connected and smart. The GEB initiative defines *"demand flexibility"* (or "*energy flexibility"* and *"load flexibility")* as *"the capability of DERs to adjust a building's load profile across different timescales"* [4]. As illustrated in **Figure 1**, the GEB initiative also identifies four demand-side management strategies covered under demand flexibility (DF) (aside from energy efficiency, EE) – load shedding (*Shed*), load shifting (*Shift*), modulating, and generation. Each of these strategies are mapped to a range of grid services[2] such as contingency reserves, transmission and distribution capacity service, and renewable curtailment mitigation as potential value streams [4]. *Shed* and *Shift* have been identified as priority DF strategies in the DOE GEB initiative and are also focused on in this report.

1 A GEB is defined as "An energy-efficient building that uses smart technologies and on-site DERs to provide demand flexibility while co-optimizing for energy cost, grid services, and occupant needs and preferences, in a continuous and integrated way. [4]"

2 "Grid services refer to services that support the generation, transmission, and distribution of electricity and provide value through avoided electricity system costs (generation and/or delivery costs). [4]"

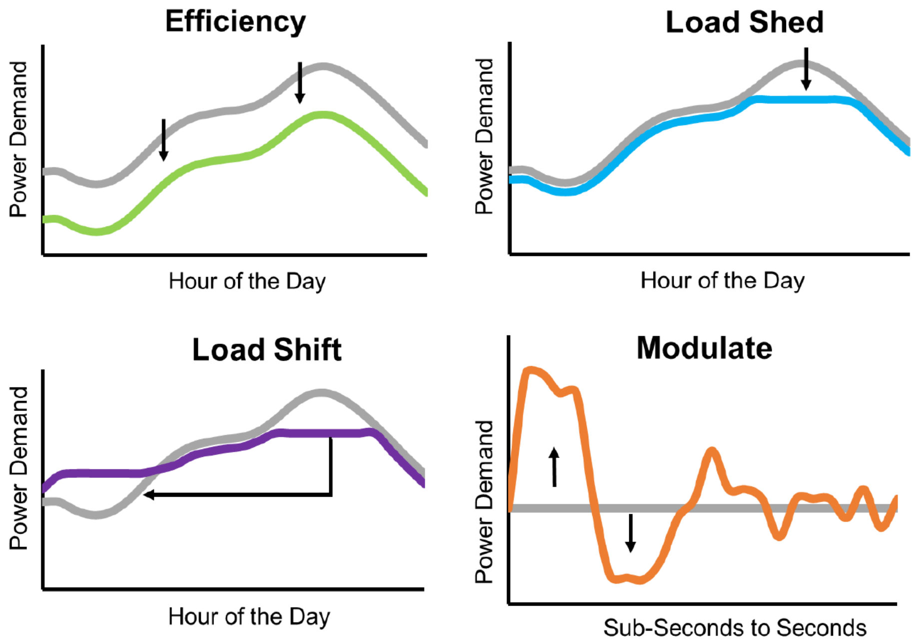

Figure 1. Building Load Curves Incorporating Four Demand-side Management Strategies

1.3 The Need for Demand Flexibility

There have been multiple important drivers for building demand flexibility research in recent years. First, the increasing penetration of renewable energy generation sources on the grid has increased the dynamics and uncertainties of electricity supply, which has made balancing electricity supply and demand more challenging [2]. Second, an important premise of DF is that, because solar and wind generation availability is subject to weather, demand needs to be flexible and able to follow power supply in order to take advantage of the renewable energy generation for grid decarbonization. Therefore, DF is used to shift building energy consumption (load) away from high greenhouse gas emission (i.e., “dirty”) periods to renewable energy abundant (i.e., “clean”) periods. The California grid is a prominent example of this with its “duck curve” phenomenon[3]. Another challenge that calls for DF resources is the capacity constraints at both bulk power system and distribution system levels and the associated high costs for capacity upgrade, which can be potentially avoided or deferred by leveraging DF and other distributed energy resources (DERs) (often referred to as “non-wire alternatives” in the utilities industry). Last, but not least, shedding load also has a unique role in maintaining grid reliability. In fact, for decades grid operators have been using Shed as an important lever when the grid is under peak or emergency conditions, which is also known as Demand Response (DR). Therefore, there is a strong connection between DF and DR.[4]

3 A description of the “duck curve” phenomenon can be found in this EIA source

4 A group of GEB researchers articulates that “like DR, DF is characterized by active load management on timescales consistent with utility system and grid needs. Unlike EE and DR, DF is not a resource in the traditional sense, but a potential that the utility or system operator can utilize to provide reliable electricity service. From the system operator’s perspective, EE and DR are what you have in your portfolio and DF is what you can do with the resources you have [5].”

One of the recent real-world events that best demonstrated the value for Shed DF was the grid emergency that happened in California on September 6, 2022 (see **Figure 2**). Around 4pm, the CA system operator had forecasted a record-high 52,000 MW system peak load during a multi-day heat wave. The system operator was ready to initiate emergency blackout procedures if DR resources could not shed enough load. In addition to the registered DR resources in wholesale and retail programs, the CA Governor's Office of Emergency Services also issued a statewide text message alert at 5:48pm calling for immediate energy conservation actions. Ten minutes after this mobile emergency text alert was sent, the system load declined by about 2,000 MW compared to the forecasted peak load. This event revealed the criticalness and high potential of DF resources in grid reliability.

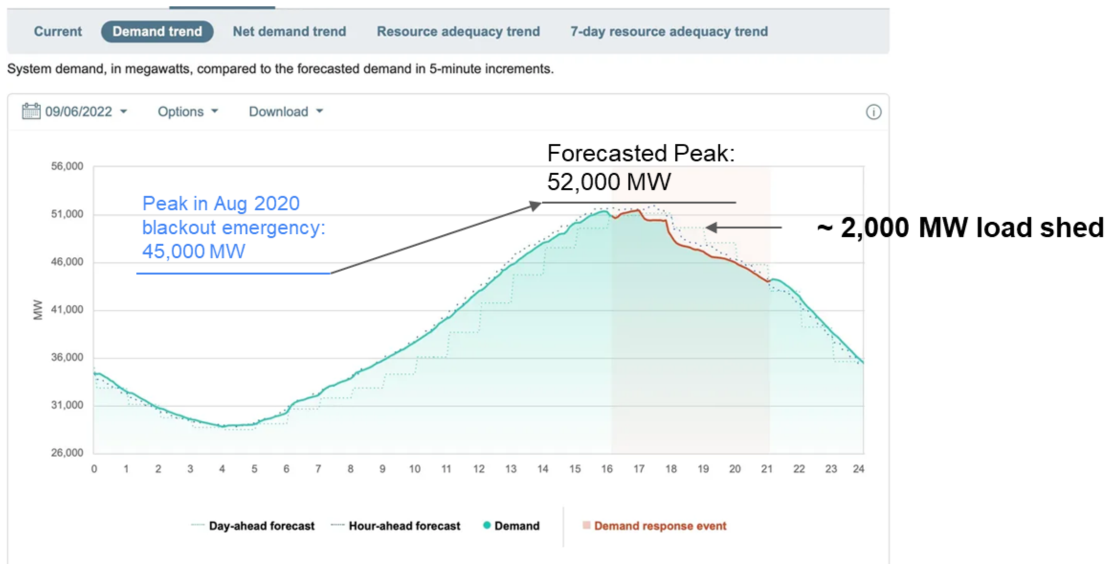

Figure 2. California Electrical Grid System Load Curve on September 6, 2022 Showing Demand Resources Effectively Reduced Peak Load by ~2,000 MW Under Emergency

1.4 High Potential Shown in Simulation

To better understand the potential value that may be provided by building DF for grid decarbonization, there has been research aimed at quantifying the DF resource size at a national or state level [6,7]. For example, the DOE GEB Roadmap [1] has identified that by reducing and shifting the timing of electricity consumption (through both EE and DF), GEBs could decrease CO_2 emissions by 80 million tons per year by 2030, or 6% of total power sector CO_2 emissions. That is more than the annual emissions of 50 medium-sized coal plants, or 17 million cars. In addition, it has also been identified that over the next two decades, national adoption of GEBs could be worth between $100–200 billion in U.S. electric power system generation and transmission cost savings. Although the Roadmap did not provide a separate estimate for emission and cost savings from DF alone, it noted that roughly one quarter of total system benefits come from avoided or deferred generation and transmission capacity costs due to EE- and demand-flexibility-induced reductions in system peak demand. In addition, the

Roadmap estimated that the commercial buildings sector will be contributing 9.3 GW[5] of new, dispatchable peak-reduction capability from demand-flexibility-only programs coming online by 2030. All of the above DF related estimates in the Roadmap only consider DF measures among traditional building end-uses enabled by control technologies (i.e., they did not include "manual" consumer actions, onsite generation, or batteries). It was also noted in the Roadmap that, of all analyzed end uses, HVAC and envelope EE and demand flexibility measures are the most impactful in the energy and peak demand savings estimates.

1.5 Gap in Understanding Real-world Performance

Currently there has been limited research progress towards understanding how the DF performance of real buildings compares with the technical estimates based on building energy simulation to support real world applications in grid planning. Considering this gap, the *Framework to Define Flexible Loads in Grid-interactive Efficient Buildings* project team at Lawrence Berkeley National Laboratory has used field performance data analysis, full-scale lab testing, and EnergyPlus simulation to advance the understanding of such comparisons and the influential factors of building DF performance. This Primer summarizes the key findings from this work.

2. Challenge in DF Variability Among Buildings and Over Time

Building DF is a special type of resource to the grid. It has at least the following characteristics. (1) As discussed in Section 1, large-scale simulation work has revealed that the technical potential of building DF resources is large in absolute scale. (2) There are DF strategies associated with building loads that are both high-impact and low-cost to implement, which makes building DF resources significantly more economical compared to other resources such as behind-the-meter batteries. The zone-level cooling temperature based strategy, "global temperature adjustment (GTA)" (as illustrated in **Figure 3**)[6] discussed in depth in this report is such a strategy. (3) In monetizing building DF for DR and other grid services, resource aggregation across multiple building sites is often either required or preferred to meet a minimum resource size threshold or counteract uncertainties in individual buildings. (4) Building DF carries a lot more variability and uncertainty in its performance compared to many other behind-the-meter grid resources (e.g., onsite generation, batteries), which creates a challenge not only for grid planners and operators but also for the aggregators and building operators.

Because of the large potential and lower cost, building DF resources cannot be ignored; it is imperative for stakeholders involved in DF and DR programs to understand its variability in order to harness the value for grid support. There are at least three dimensions to the variability in building DF:

1. DF performance difference across similar buildings;

5 Currently, about 10 GW of peak demand reduction capability in the U.S. is known to come from technology-enabled demand flexibility in residential and commercial buildings.

6 GTA is raising zone temperature setpoints "globally" across a number of zones in the building during a Shed event. It is common for buildings to couple it with "pre-cooling" by lowering the zone temperature setpoint for some time before the Shed in order to achieve deeper demand shed and provide better comfort during the event.

2. variation in DF performance of the same building across different events; and
3. fluctuation of real-time performance (i.e., shape of load changes during an event).

In this Primer, we primarily focus on (1) and (2) because event-average performance is currently a primary focus of utility DR programs.

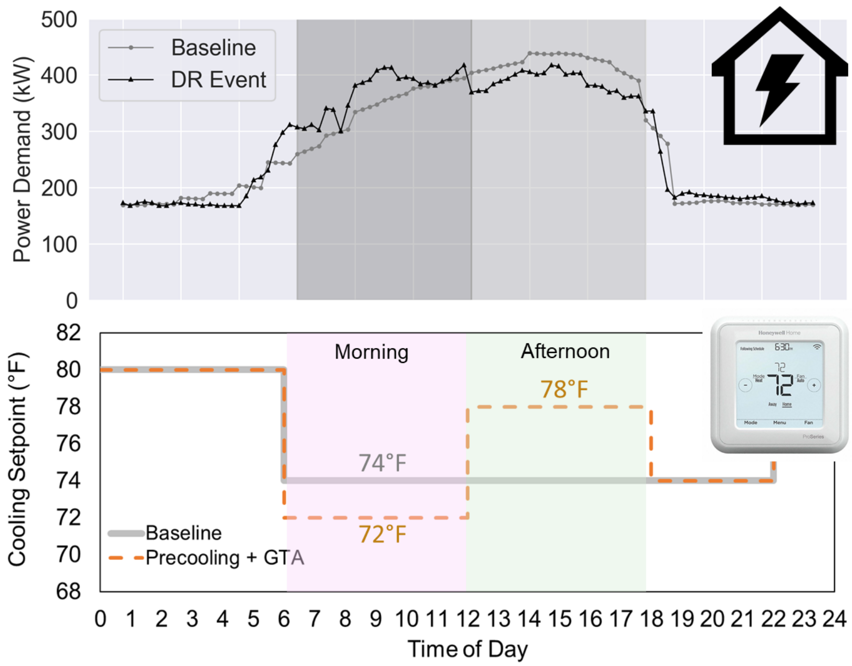

Figure 3. Global Temperature Adjustment (GTA) Strategy (Top: Load Shape of Baseline vs DR Response; Bottom: Setpoint of Baseline vs DR Response)

Standardized metrics and methods are required to quantify and support understanding of these variations. Only if we can quantify real world DF performance using standardized metrics and utilize systematic analysis methods, can we begin to understand correlations between performance and key influential factors. Such metrics and methods [2] (Section 3 provides a summary; more details found in Appendix A) have been applied in benchmarking the DF performance of two groups of different building types: 11 medium office buildings and 121 big-box retail buildings. More information about these two building groups (e.g., locations, DF strategies, data periods, etc.) are found in **Figure 4**.

	Medium Office Buildings	Big-box Retail Buildings
Sites	11 sites in California	121 sites in 11 states, West to East coasts
HVAC System	RTUs; VAV system	Single-zone RTU, VFD fan
DF Strategy	**Shed/Shift**: GTA (2-3 ºF); 5 buildings implemented precooling + GTA	**Shed**: GTA (3 ºF) for non-critical RTUs; Turn off ~50% lights in sales area
DR Event	12 - 6 pm (6 hours)	Vary by site and event, 1-6 hours long
Data Period	2008 (15 min whole building power)	2018-2021 (15 min whole building power)

Figure 4. Basic Information of Two Building Groups Used for DF Benchmarking - Medium Office Buildings and Big-box Retail Buildings

We use some high-level results from the above benchmarking analyses to introduce variabilities in building DF performance. **Figure 5** shows the load shed performance in W/ft2 and percentage measured in 94[7] of the 121 similar, big-box retail buildings that had sufficient data to calculate benchmarking metrics for 2-hour events. The metrics of this group of similar buildings spread between -0.2 W/ft2 (-8%) and 1.0 W/ft2 (40%). This is an example of "(a) DF performance difference across similar buildings".

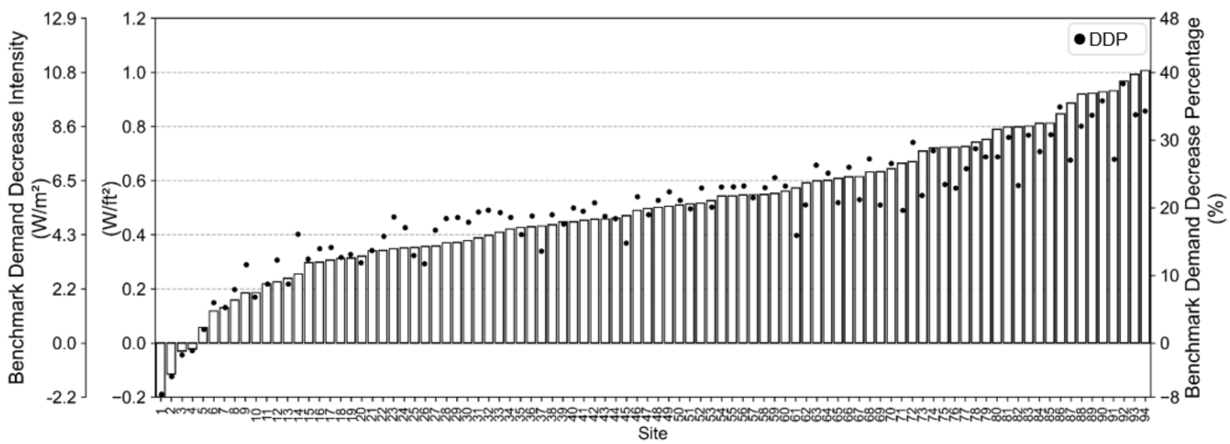

Figure 5. Benchmarking Metrics for 94 Big-box Retail Buildings That Participated in Two-hour DR Events in 2021

Figure 6 shows the event-level shed performance of the 94 retail buildings in **Figure 5** above.

[7] 94 of the 121 retail buildings had five or more 2-hour events in 2021 in order to support calculating the benchmarking metrics

As it reveals, for a given building, the shed performance across different events often spreads beyond 0.2 W/ft2 and sometimes as much as 0.8 W/ft2 or more. This is an example of “(b) variation in DF performance of the same building across different events”. A similar example of the medium office buildings group is shown in **Figure 7.** One can observe that the event-to-event variability in the six buildings that implemented additional precooling is significantly smaller compared to the five buildings that only implemented load shed. This preliminary finding is worth further exploration to understand if precooling indeed can improve the consistency of GTA load shed performance.

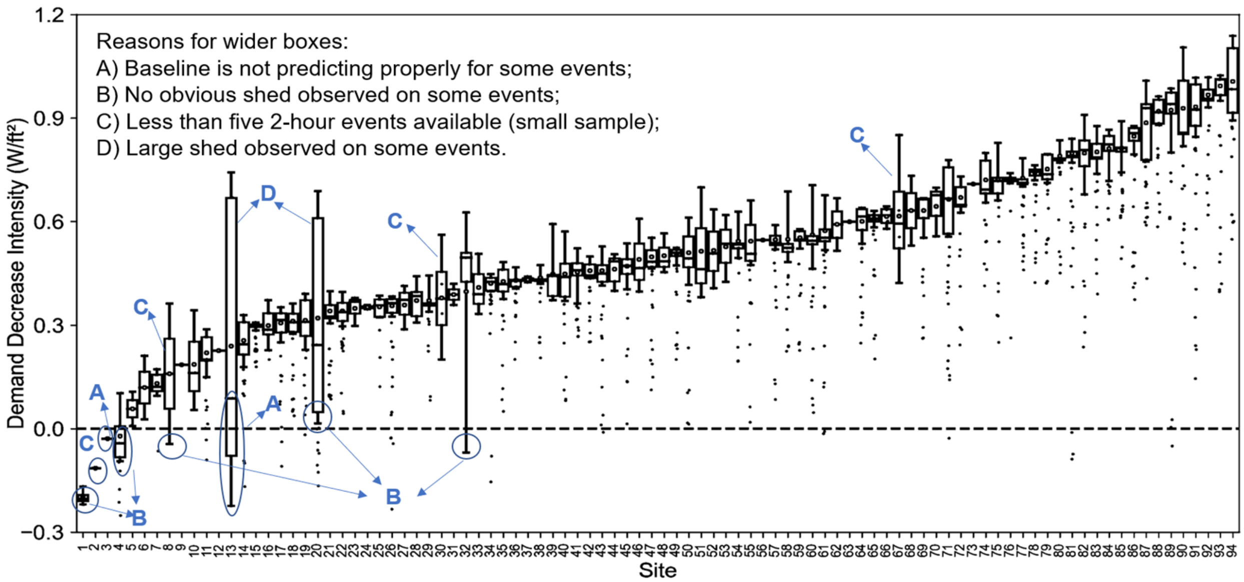

Figure 6. Boxplot of Single-event Demand Decrease Intensity Metric for 94 Big-box Retail Buildings That Participated in DR Events, 2021 Data (The boxes are drawn based on five DR events that yielded the highest values. The dots represent other DR events.)

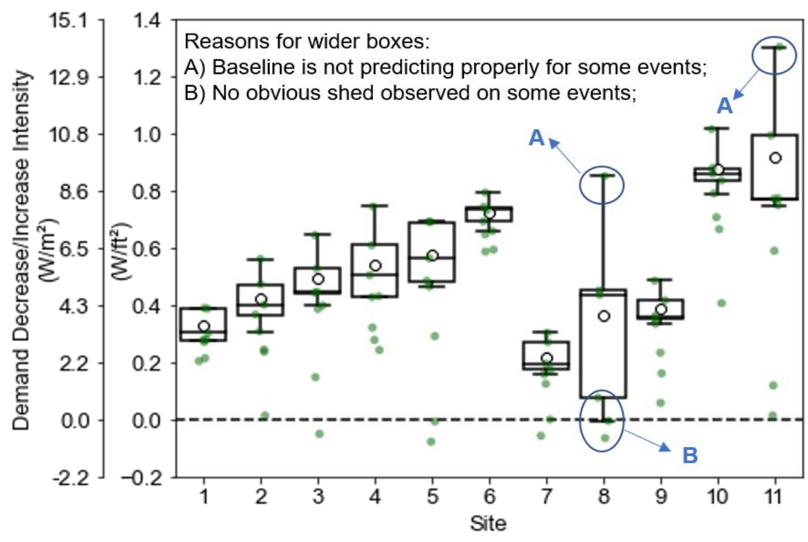

Figure 7. Boxplot of Single-event Demand Decrease Intensity Metric for 11 Medium Office Buildings. (Each dot represents a DR event; the boxes are based on five DR events that yielded the highest values.)

Many factors can contribute to the building-to-building and event-to-event variations shown in **Figure 5** to **Figure 7** including climate and weather, building characteristics, when the event was called, and building operations. Section 4 explores how several key factors influence DF performance.

3. The Need for DF Metrics and Benchmarking

As discussed earlier, there is a significant gap in understanding buildings' real-world DF performance, and we need standardized metrics and methods in order to quantify and support understanding of DF performance variations in the field. We not only need metrics to tell us how to measure a building's DF performance in snapshots (e.g., during a specific event), we also need a holistic framework and significant datasets to facilitate understanding how the performance of similar buildings compares and what key factors drive building performance in what way. The DOE GEB Roadmap stated (under Pillar 1, Recommendation 3): "Develop standard metrics and methods for data collection, data analysis, and measurement and verification (M&V) of demand flexibility technologies and strategies." The Roadmap further stated: "GEB field performance assessments and metrics are needed to enable grid operators to trust the ability of demand flexibility to reliably deliver grid services… Also, building owners and operators are unwilling to invest in technology without a clear value proposition based on proven technology benefits. Demand flexibility benchmark data sets, load shapes, and metrics are needed across all building sectors to provide relevant, comprehensive data for GEB technology performance evaluation. To draw meaningful conclusions from the data that can be relied upon by grid operators, utilities, and customers, there is a need for statistically significant data sets at scale and across different dimensions of building type and time (e.g., hourly, daily, annually)."

To address this gap, we propose a set of foundational metrics (see Appendix A) and a tested framework for benchmarking field-measured building DF performance. (Full documentation of the framework, metrics calculation procedures, and application examples used for testing this framework are found in [2].) It is important to describe some key challenges that inspired the innovative design of this benchmarking framework. The load flexibility needs for the grid vary significantly across the country and over time. We illustrate this multi-dimensional variability in **Figure 8** where we aggregated DR event timing information during 2018-2021 for 203 retail stores located across 11 states coast-to-coast [8]. DR event timing including event duration, start hour, and month of the year can vary geographically and also have changed over the four years. This presented a challenge in benchmarking DF as an effective framework should (1) allow comparing DF performance across buildings serving different grid needs, represented by different event timing, meaningfully on a level-playing field using existing field-measured data; (2) enable using empirical approaches to achieve a rich understanding of DF performance influential factors by extracting maximum information from individual buildings' field performance data.

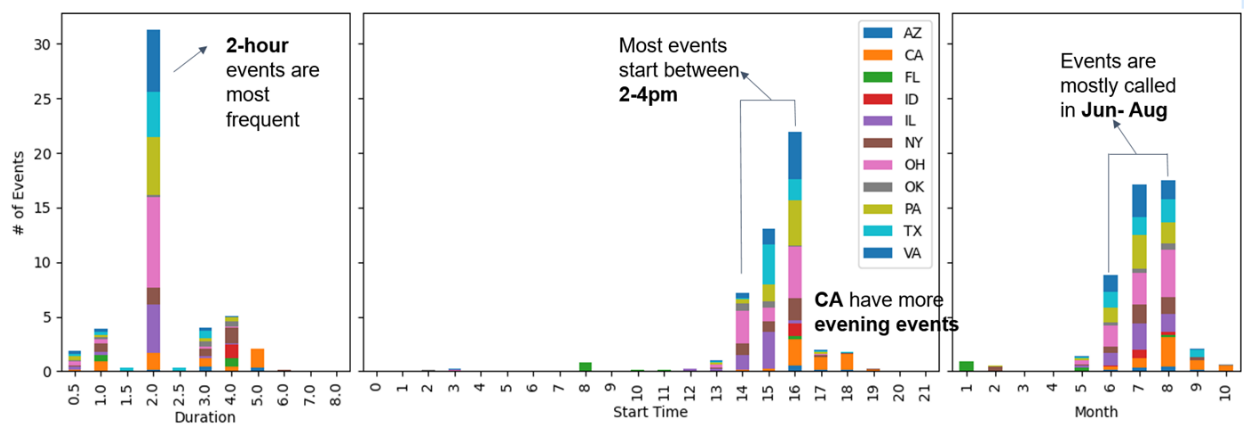

Figure 8. Average Number of DR Events Called per Site in 2018-2021 for a Group of 203 Retail Stores (Left: Event Duration; Middle: Event Start Time; Right: Month) (Note: the number of events have been normalized by the number of sites in each state)

The above two criteria inspired the concept that each field-measured DF performance data point (e.g., demand decrease intensity in W/ft2) should be associated with a set of “attributes”, which will describe under what conditions the performance data point was recorded. This is the foundation for a dimensional approach to benchmarking building DF using one or more attributes as dimensions. Such a dimensional benchmarking framework enables identifying trends and dependencies of DF performance relative to one or more influential factors using graphical analysis approaches (as illustrated in Figure 9 and described in detail in [2]).

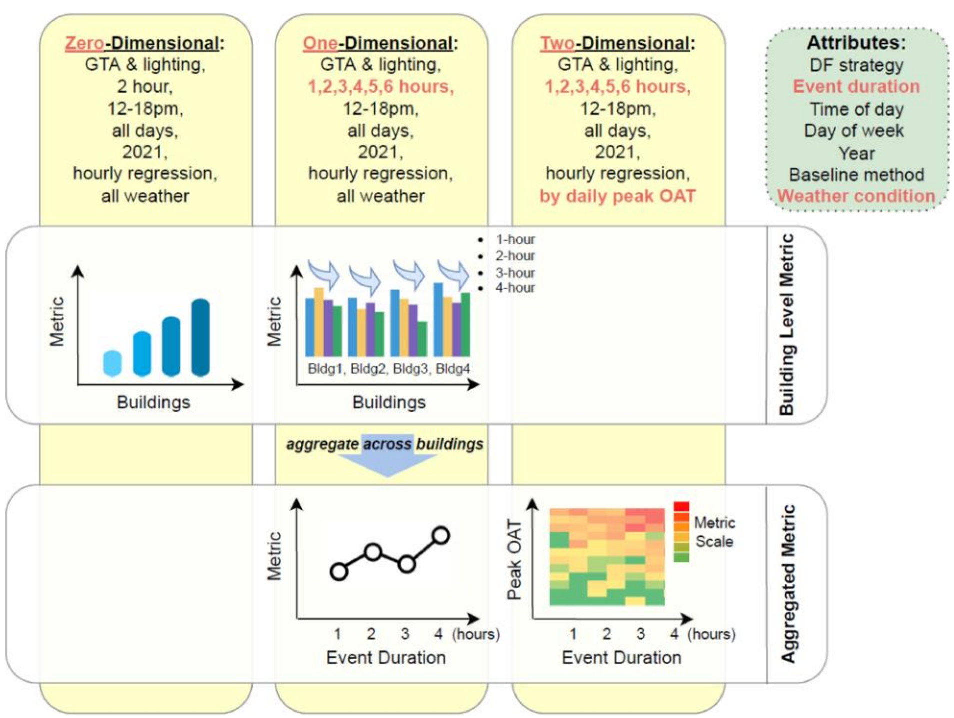

Figure 9. An Example of DF Benchmarking Metrics in Different Dimensions (Note: one or two attributes can be used as dimensions for graphing)

When it comes to benchmarking DF using field-measured performance, creating reliable counterfactual baselines is a challenge. Currently, there is no proven baseline method that works well for all commercial buildings all the time; existing baseline methods all have their pros and cons. For the same buildings, using different baseline methods can results in significant differences in the metrics. Therefore, the accuracy of selected baseline can have a significant impact on benchmarking and on how buildings' DF performance is interpreted, which may or may not match the reality. A more detailed discussion of progress and challenges around baseline methods for load shed and shift is found in [8]. The benchmarking results of the two building groups shown in this Primer are based on using a single-variable (outdoor air temperature), hourly linear regression model baseline method with 14 days training period. This method was tested to be suitable for the two building groups mentioned in Section 2 and more advantageous than the conventional "10-in-10"[8] type (with or without adjustment) of baselines for load shifting [8] although its accuracy has not been more widely validated. Clearly, there is still a research need for creating and validating reliable baseline methods for DF, which is also identified in the GEB Roadmap.

For buildings with DERs such as onsite generation, battery storage or EV charging, it is often imperative to apply the above mentioned baseline methods to the building total load rather than the utility net load. The building total load can be calculated by taking the sum of the net load and the submetered energy generation; for battery storage and EV charging, discharged energy will be equivalent to energy generation and charging will be equivalent to energy consumption. **Figure 10** provides an example of creating a counterfactual baseline for a building with PV generation by adding PV energy generation to the building net load. After that, the steps for calculating DF metrics will be identical to those for buildings without such DERs.

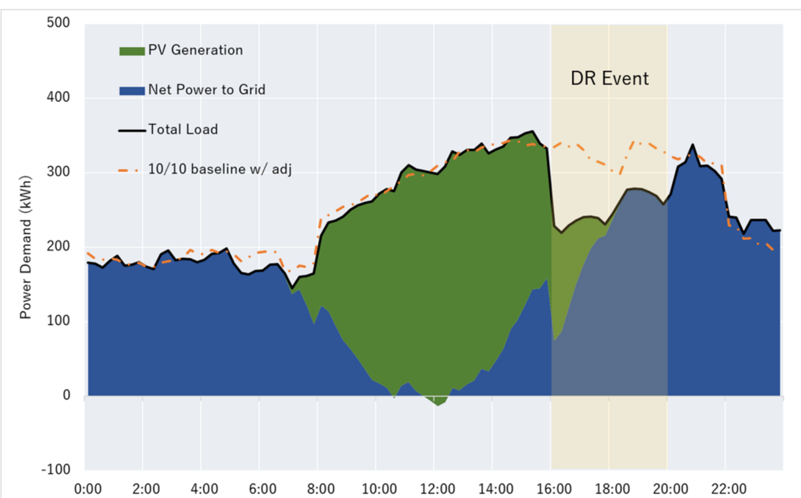

Figure 10. An Example of Creating Baseline Load Shape for Buildings with Distributed Energy Resources (DERs) for Calculating Demand Flexibility Metrics

[8] "10-in-10" baseline: a type of day-matching baseline method which takes the average demand during DR event hours over the previous 10 eligible baseline days (excluding weekends, holidays, DR event days, and none-operation days)

4. What Influences DF

As discussed in Section 2, a big challenge with harnessing building DF as a grid resource is in the multiple dimensions of performance variability. A wide range of factors may have significant influence on a building's DF. Organizing them into the following three categories (**Figure 11**) can help us understand their influences systematically.

1. **Building characteristics,** such as building type, vintage, operating hours, space use, climate zone, and state[9]. These factors are generally static, as opposed to varying event by event, and they are useful in identifying the right cohort of buildings for benchmarking.
2. **Single-event attributes,** such as DF strategy, event duration, time-of-day, day-of-week, year, baseline method, and weather condition. Significant performance variations across DR events in the same building are common. These are important conditional parameters to specify for each single-event metric value. **Table 1** summarizes considerations behind highlighting these factors' influence.
3. **Stochastic factors,** such as stochastic phenomenon in occupancy, internal loads, and business operations, as well as commissioning issues in equipment and controls, which can have profound impacts on DF and yet are often overlooked.

Figure 11. Three Categories of Factors That Influence Demand Flexibility

We used a combination of three complementary methods (**Figure 12**) – EnergyPlus simulation, full-scale lab testing, field data analysis – to explore how several key factors in the above categories influence building DF. Each of these three methods have their pros and cons. Among these three methods, FLEXLAB®[10] testing offers the greatest accuracy in terms of

---

[9] For some datasets and applications, independent system operators or utilities may be suitable alternatives to the state.
[10] FLEXLAB® (https://flexlab.lbl.gov/) is a completely customizable and configurable whole -building integrated

creating reliable baselines and isolating influential factors although this method is more expensive. Field data analysis of a significant building group is critical and irreplaceable in understanding the range and issues of real buildings' performances; it is also less expensive to analyze available existing field data than conducting new lab testing. However, the availability of desirable datasets is currently limited due to data ownership issues. It is also often unrealistic to obtain insights into causes of building performance variations on a building-by-building, event-by-event level, which makes it difficult to pin down the impact of each influential factor. EnergyPlus simulation allows users to change the values of individual influential factors to cover a wide range of scenarios quickly through batch runs at a low cost. Although there are limitations[11] to how accurately EnergyPlus can represent the true dynamic behavior of different buildings carrying out DF strategies, simulation does offer a useful tool in creating hypotheses before engaging other more expensive methods.

Table 1. Considerations for "Single-event Metric Attributes" Influencing Demand Flexibility

Single-event metric attributes	Consideration
DF strategy	A building may support multiple control strategies to shed or shift load, such as adjusting thermostat setpoint to reduce HVAC load, dimming lights, curtailing plug loads, or discharging thermal or electrical storage. Not all of the available strategies may be used in all Shed/Shift events, as they can be prioritized based on the building's utility tariff, utility program rules, and the impact on building services. For example, a building may choose to deploy a single strategy for an economic program versus several strategies in an emergency DR program event.
Event duration	Electrical load shed from some building loads is easier to sustain than from others. For example, it is easier to dim lights for a few hours than to cycle off HVAC for hours because the space may get uncomfortable for the occupants when the HVAC is off.
Time of day	Building loads (e.g., lighting, plug load, HVAC) and their ability to shed or shift vary throughout the day and week as occupancy and operation mode change (e.g., HVAC setback during unoccupied periods).
Year	The same building's DF performance from the same DF strategy can change significantly over time due to operational changes, equipment conditions, and other factors. This may include both improvement and degradation.
Baseline method	DF metric value can vary significantly depending on the chosen baseline method.
Weather conditions	Some building loads, such as HVAC, are often dependent on weather conditions, and therefore can influence load shed or shift from these loads.

systems test facility that was designed to study, develop and validate systems level solutions, tools and processes for the commercial building market.

[11] Some areas of limitations in using EnergyPlus to understand DF may include representation of building thermal mass, HVAC equipment cycling, and control latency.

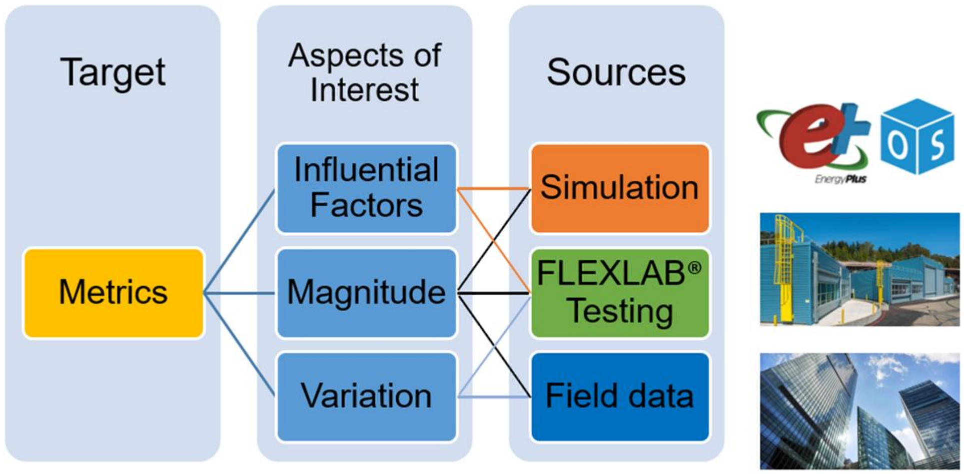

Figure 12. Three Complementary Methods Used in Exploring the Impacts of Demand Flexibility Influential Factors

In this section, we summarize the key findings from using the above three methods in understanding key influential factors' role in DF performance obtainable from the pre-cooling and GTA strategy. Again, we focused on this strategy in our work considering it is a practical and scalable DF strategy with low implementation cost and commercialized technology solutions and real-world performance data available. The following subsections each dive into the complexity and preliminary understanding of a common influential factor, including weather, event timing, HVAC system type, building type, building vintage and energy efficiency, thermal mass, and operations and commissioning issues. The key findings from these subsections are summarized in **Table 2**.

Table 2. Summary of Findings on Key Factors' Influence on Building Demand Flexibility

Weather	**Event Timing**	**HVAC System Type**	**Building Type**
Temperature: positive correlation with shed but with significant noise; **Humidity:** no clear correlation found; **Solar:** cloudy condition reduces shed.	**Duration:** shed decreases significantly for longer events; **Start time:** evening events yield lower shed compared to afternoon.	Single-zone RTU systems tend to yield higher shed compared to multi-zone VAV systems.	Many parameters at play: building size, construction, space use, window-to-wall ratio, internal load, HVAC type, zoning, equipment efficiency, operating hours, and more.
Building Vintage & Efficiency	**Thermal Mass**	**Operations & Commissioning**	
Many parameters are at play. Majority of efficiency measures reduce shed but some can increase it while others are neutral.	Load shift with pre-cooling can be energy efficient and improves comfort but may not significantly increase shed. Warm room temperature overnight may reduce shed significantly.	Can have profound impact; case-by-case. E.g., HVAC equipment not functioning properly or commissioning issues related to hardware or controls.	

4.1 How does weather influence DF?

Since building HVAC loads are typically sensitive to weather conditions, the performance of HVAC DF strategies is also expected to be influenced by weather. Here we discuss the influence from three key weather parameters - daily peak air temperature, solar radiation (considering cloudiness), and humidity - based on our study using various analysis and testing tools. The discussion below is not meant to be comprehensive considering the complexity and dynamic nature of how weather may influence building loads.

4.1.1 Outdoor temperature and solar radiation

Previous DR simulation and field studies have shown that building HVAC load shed performance is sometimes correlated to outdoor temperature (an example shown in **Figure 13**). EnergyPlus based simulation studies [9,10] have shown strong correlations between GTA load shed and outdoor temperature although how realistically the simulation tool represents certain weather parameters (e.g., cloudiness) and building thermal mass has not been well understood. In field studies, the results are less clear as such correlations may be subject to significant noise from other stochastic factors [11]. Lab testing combines the advantage of capturing the real physical system's response while being able to limit the number of variables and control their influence. We used the FLEXLAB® testing facility at LBNL to examine the correlation between load shed / take from two strategies ("GTA-only" and "Pre-cooling + GTA") and daily peak temperature[12].

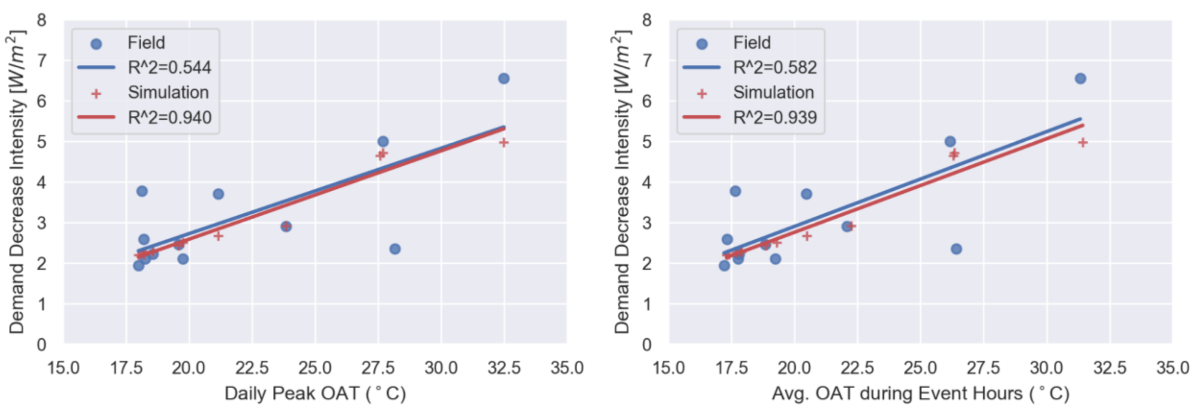

Figure 13. Relationship between Demand Shed Intensity [W/m2] and Daily Peak OAT (left) and Average OAT During Events (right) in a Large Office in California, Calculated from Field DR Event Data and Prototype Simulation [10]

Two separate lab tests were conducted in the fall of 2020 and 2021 for two different HVAC system types respectively, i.e., single-zone constant air volume rooftop unit (RTU) system and single-duct variable air volume (VAV) system[13]. The results are shown in **Figure 14**. A positive

[12] In our analysis, both daily peak temperature and average outdoor air temperature during a shed event were found to be effective independent variables. Daily peak temperature is selected because it requires less granular weather data and is easier to access.

[13] The VAV test was conducted with 3 mock-up zones divided by partition walls with one of them facing the South with windows. The other two zones are considered similar to internal zones. The two lab tests used the same envelope and internal load configurations.

correlation can be observed between load shed and daily peak temperature for both system types although there are outliers in both. The observed slope is significantly greater with the single-zone RTU system than the VAV system. No clear correlation is observed between load take and the daily peak temperature. Key contributors to the outliers in load shed correlations were found to have to do with cloudy conditions and warmer temperatures overnight[14]. Solar radiation is typically a main contributor to cooling load in commercial buildings, and therefore is influential to load shed intensity. As revealed in **Figure 14**, when daily peak temperatures are similar on different days, cloudy conditions have resulted in significantly lower shed intensity. The influence of warmer temperatures overnight is discussed later in Section "How does thermal mass influence DF?".

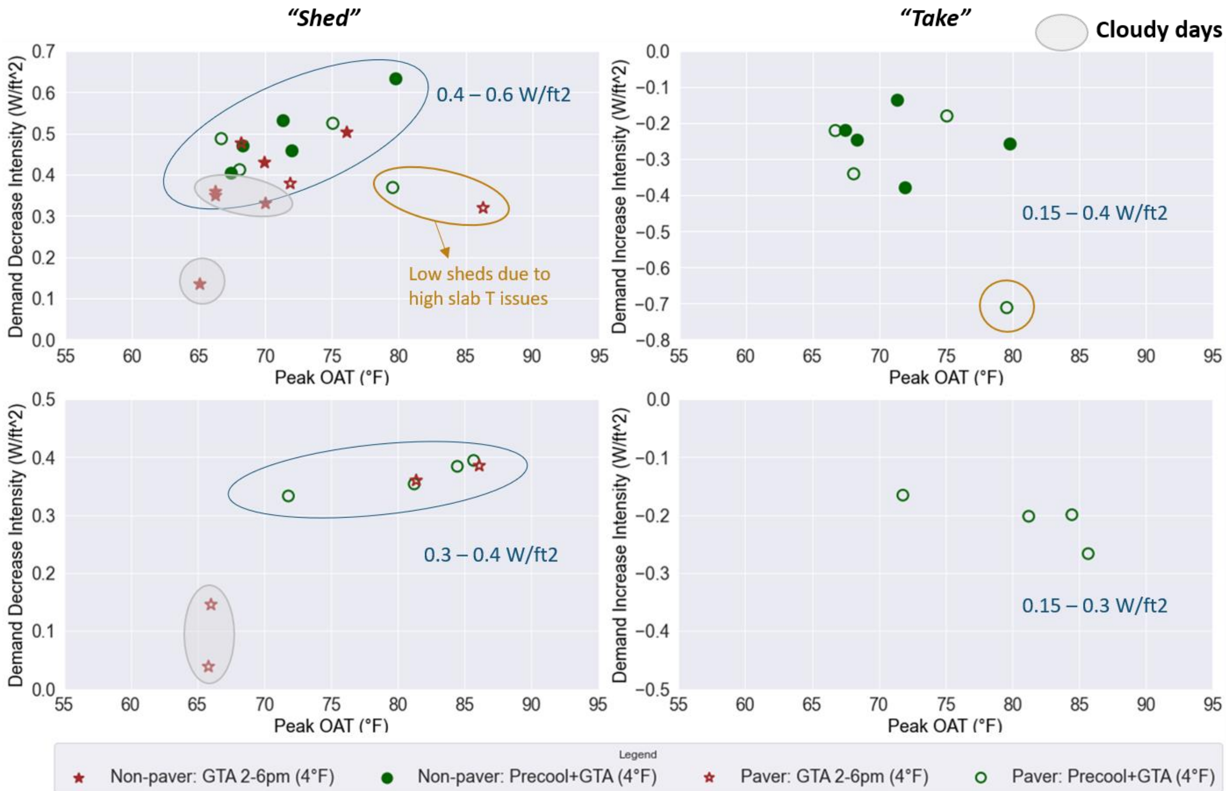

Figure 14. Relationship between DF and Daily Peak Outdoor Temperature from FLEXLAB® Testing (Top: Single-zone RTU System Test; Bottom: VAV System Test)

Well-controlled lab testing has revealed complexity in how weather influences demand flexibility - manifested as noise and outliers in main correlations. In real buildings, the influence of weather parameters is further diluted by many other factors, and is therefore not as clearly observed as in lab tests. Furthermore, the HVAC cooling load in some building types is significantly less sensitive to weather because there are fewer windows and the HVAC load is internal load dominated. Big-box retail buildings are an example.

---

[14] These are considered preliminary conclusions to be further validated in future research. Robust conclusions could not be drawn due to the limitations in data points available.

4.1.2 Humidity

Humidity is sometimes perceived as influential to building cooling load and the related demand flexibility performance. We investigated this issue using both EnergyPlus simulation and real building field performance data.

We used a DOE prototype model for a medium office building of the ASHRAE 90.1-2004 vintage as an example to explore influence from humidity. Two climate zones[15] on similar latitudes with dramatically different humidity, i.e., warm-dry (3B) vs. warm-humid (3A), were selected to run simulation for comparison. **Figure 15** shows the correlation of simulated shed intensity vs. outdoor temperature and relative humidity. Strong positive correlation is observed between shed intensity and outdoor temperature in both climate zones, and no significant difference is observed in such correlation patterns between the warm-dry vs. warm-humid climates. No clear correlation is observed between shed intensity (DDI) and relative humidity.

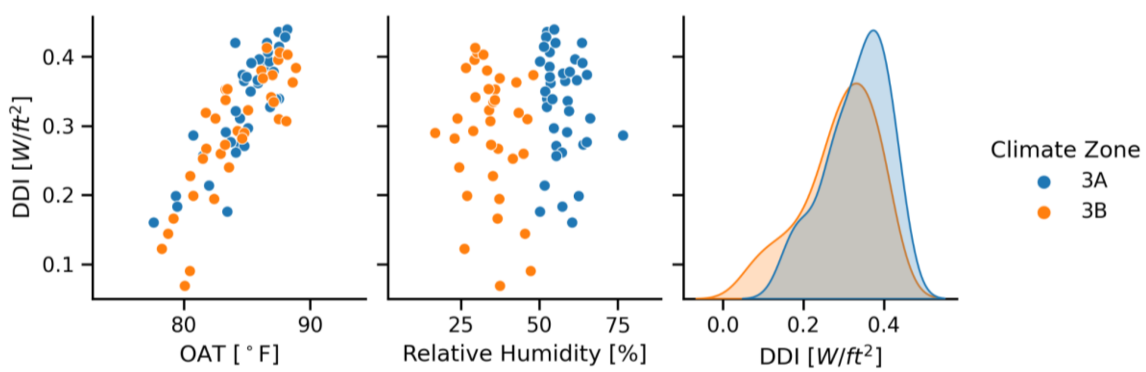

Figure 15. Compare the Correlation between Demand Decrease Intensity and Outdoor Air Temperature or Relative Humidity Using a Simulated Prototype Medium Office Building in Two Different Climates: 3A vs 3B

To better understand the underlying reasons for why humidity levels did not make a significant difference in shed performance, we examined the load shed components from sensible and latent cooling loads in **Figure 16**. On a representative hot day, the GTA load shed from sensible cooling load is predominant in both the more humid climate and the dryer climate despite their different baseline load levels. This is because the latent load is determined by the dew point of the return air vs. the cooling coil temperature, and it is not directly impacted by the room temperature coasting process resulting from the GTA strategy.

[15] ASHRAE standard

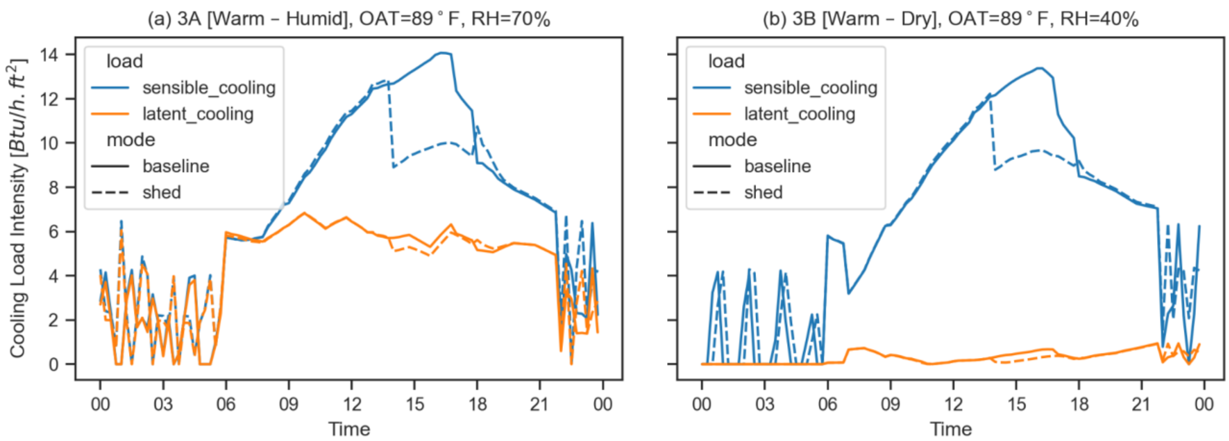

Figure 16. Sensible and Latent Cooling Load of a Simulated Prototype Medium Office Building on an Example Hot Day in Two Different Climates: 3A (Left) vs 3B (Right)

In addition to simulation, we also used field measured DR performance data to understand how outdoor air humidity potentially influences load shed. **Figure 17** shows shed intensity measured in each DR event day from around a hundred big-box retail buildings across seven states with significant data points[16]. We experimented with creating scatter plots of shed intensity against multiple humidity parameters including relative humidity, wet-bulb temperature, and dew point temperature. The shed intensity against wet-bulb temperature distribution is tighter compared to using the other two parameters tested. However, no clear correlation is observed across the seven states. For the same group of retail buildings, attempts to find correlations between shed and dry-bulb temperature also did not yield any clear patterns. This means in big-box retail buildings, which may be expanded to other commercial buildings as well, there are other factors (e.g., events that significantly change internal load) that are more influential than weather including humidity on the buildings' load shed performance. These other factors were discussed in the beginning of **Section 4**.

[16] The number of sites in each state vary between a few to a couple dozen.

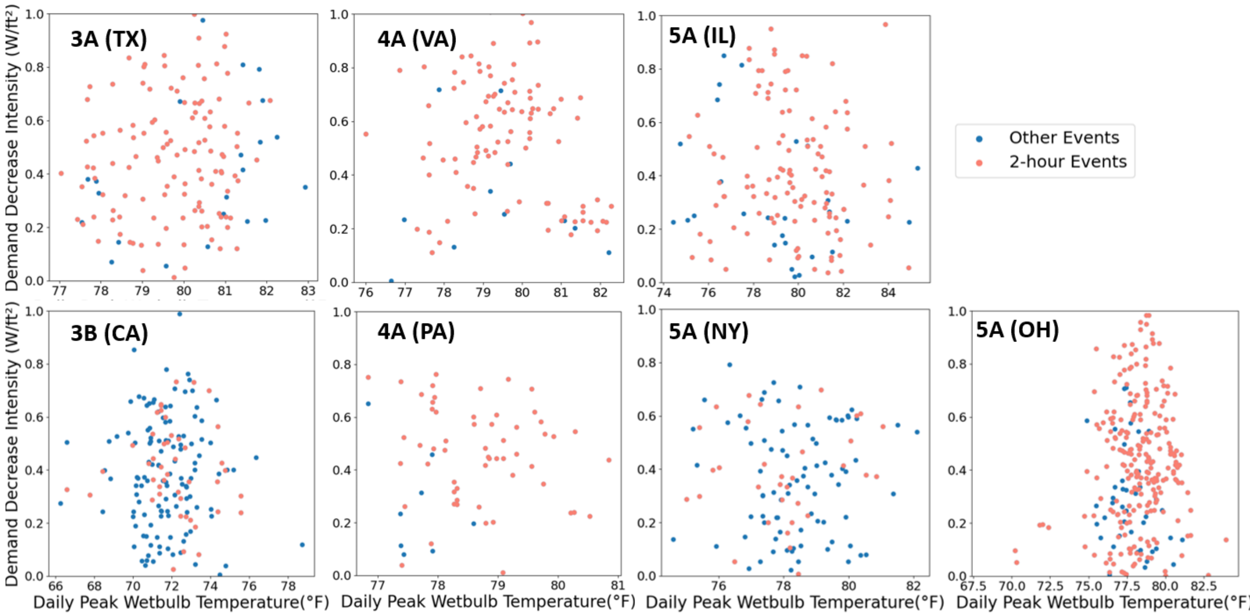

Figure 17. Demand Decrease Intensity vs Outdoor Wetbulb Temperature for Big-box Retail Buildings in Seven States

4.2 How does event timing influence DF?

The timing of the load shed / shift events are defined by the event start time and the duration. The event timing has influence over DF performance because building loads are often subject to operation schedules. In addition, solar radiation presents a typical daily profile which in turn influences outdoor temperature; both solar radiation and temperature drives building cooling and heating baseline loads and its potential for shedding/shifting.

We examined the trend of how load shed changes as a function of event duration in benchmarking DF for the aforementioned group of big-box retail buildings. By aggregating load shed performance results across all buildings over shed events of two to six hours long, we observe a clear descending trend of shed intensity as the event duration increases. Given the morning through evening operating hours of these retail buildings, we were able to conclude that the GTA strategy (rather than the lighting strategy) was primarily responsible for this descending pattern. First, when the thermostat room temperature reading rises to the new GTA setpoint, the rooftop unit compressor will restart and the load shed will decrease after this initial "coasting" period. Second, the longer events in this group of buildings often started in the afternoon and ended in the evening; therefore, the baseline cooling load for part of the duration was lower to begin with. Note that the amount of load shed intensity reduction observed in **Figure 18** is not necessarily generalizable beyond this building group although we expect that the general declining trend will also be applicable to other buildings that employ the GTA strategy.

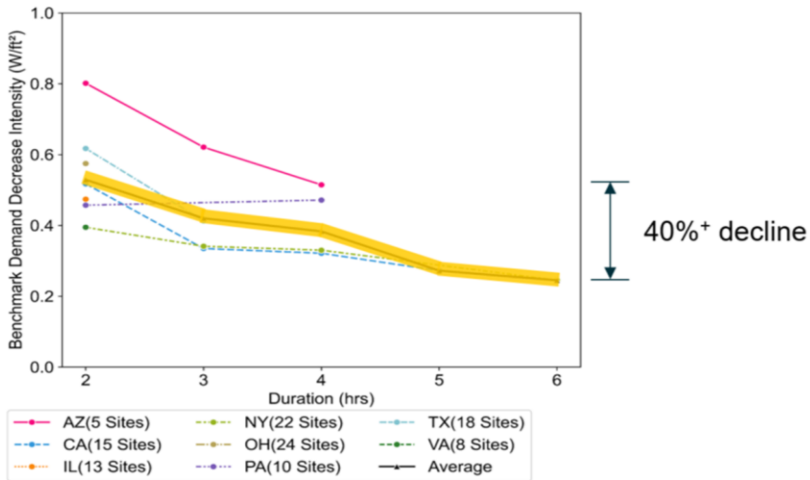

Figure 18. Demand Decrease Intensity vs Shed Event Duration Trends Found in 115 Big-box Retail Buildings across Eight States

We have also examined how event start time may influence load shed using the same building dataset although the number of retail buildings that had both afternoon and evening events were very small. Therefore, the number results in **Figure 19** are considered illustrative. It shows that big-box retail buildings that operate into evenings can still offer significant load shed potential in evening shed events although the performance is expected to be significantly lower than that in afternoon events because the baseline load decreases in evenings. It is worth noting that the influence from the event duration and start time can be difficult to separate because both may influence the baseline load level; therefore, they should not be seen as independent variables in interpreting results.

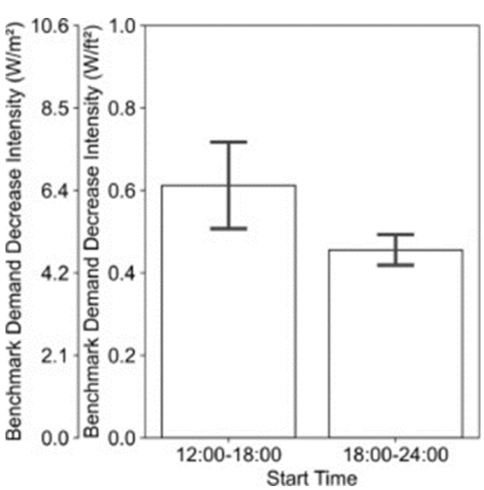

Figure 19. An Example of Demand Decrease Intensity Being Lower in Evening Events Compared to Afternoon Events in a Big-box Retail Building in California

4.3 How does HVAC system type influence DF?

For HVAC system related load shed and load shift strategies, and the GTA strategy in particular, the building's HVAC system type(s) can have a profound influence on expected DF performance. This is because the mechanism through which load shed and shift is achieved differs in different types of HVAC systems. We do not intend to cover all HVAC system types here but will contrast how the GTA strategy works in two popular system types found in commercial buildings - single-zone CAV RTU system (hereinafter "RTU systems") and single-duct multi-zone VAV system (hereinafter "VAV systems").

In single-zone RTU systems, after the GTA strategy is triggered, the RTU cooling stages off, and as a result, room temperature rises (known as the "coasting" effect) towards the new setpoint. This coasting process happens more quickly in single-zone systems than it does in VAV systems under similar conditions because there is no minimum cooling being supplied to the room in this type of system. In RTU systems, it is common to have the supply fan and compressor operate simultaneously, in which case they will be both off during the coasting process and yield significant load shed. Once the new setpoint is reached, RTU cooling stages back on and then cycles on/off after that to maintain the room temperature to stay at or below the new setpoint. This process is reflected in **Figure 20**, which shows room temperature reading in a FLEXLAB® test of this HVAC system type. During this second part of the shed period, the cooling load is also reduced compared to baseline because the indoor-outdoor temperature difference is smaller with GTA activated. However, load shed during this period with RTU cycling on/off is not as significant as it was during the coasting period.

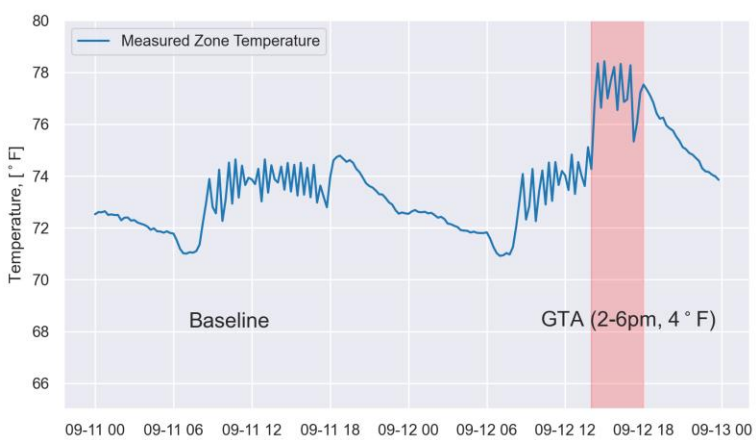

Figure 20. Room Temperature Trend from Testing DF in a RTU System at FLEXLAB® Test (Baseline vs. 4ºF GTA)

In VAV systems, there is also a similar room temperature coasting process in each zone. When the GTA strategy is triggered, the new setpoint goes above zone temperature and therefore the zone airflow drops to the minimum airflow setting. The room temperature coasting process in each zone can happen at significantly different rates depending on the cooling load relative to the zone's minimum airflow setting. If the minimum airflow is relatively high and the cooling load is relatively low for a given zone, then zone temperature coasting speed can be slow and it may or may not reach the new GTA setpoint before the shed event ends. This can often be the case for core zones but can also apply to exterior zones. **Figure 21** shows the zone temperature coasting process during a shed event using prototype building simulation. For example, the cooling load in the zone facing East typically peaks in the morning; therefore, the coasting process was slow during an afternoon shed event. Once the room temperature has reached the new setpoint, the zone supply airflow will increase above the minimum airflow setting to meet the cooling load.

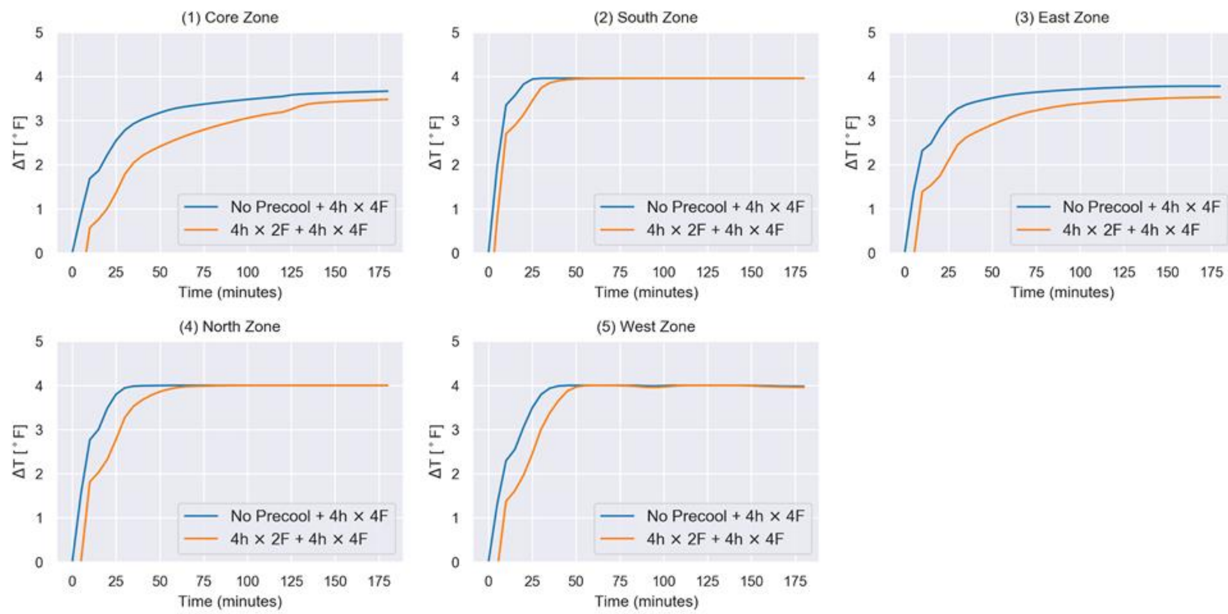

Figure 21. Zone Temperature Coasting under 4ºF GTA for 4 hours without Pre-cooling vs. with 2ºF Pre-cooling for 4 hours Simulated Using a Prototype Medium Office (90.1-2004 Vintage) in 3B Climate (Average Results during the 12 Hottest Weekdays)

In VAV systems, the load shed may also be expected from both fan power and at the cooling source (e.g., RTU compressor or chiller) during the shed event. When the zones go through temperature coasting at various paces, fan power decreases as one or multiple zones drop to minimum airflow, and cooling source power demand may decrease as the indoor-outdoor temperature difference is reduced. Because of the continued cooling requirements associated with minimum airflow supply to each zone, neither the supply fan or the cooling source will turn off during a shed event that happens within a building's operating hours. As a result, the amount of load shed that can be expected in a VAV system from the same GTA strategy is generally lower than it is in a single-zone RTU system. This difference can be observed in the FLEXLAB® test results shown in **Figure 14** – 0.3-0.4 W/ft^2 (VAV) vs. 0.4-0.6 W/ft^2 (RTU) – where the testing of these two HVAC systems were configured to be the same with respect to envelope, internal loads and HVAC setpoints.

4.4 How does building type influence DF?

The various types in commercial buildings can influence DF in various ways. For example, different building types may imply significant differences in building size, space use, building construction type, envelope-to-volume ratio, window-to-wall ratio, internal load density, occupancy schedules, HVAC system type, zoning, equipment efficiency, operating hours, etc. These factors can also vary significantly even within a given building type considering the differences in individual building designs. Therefore, "building type" should not be treated as a single parameter in assessing how it influences DF given the number of key factors at play.

Below we discuss how each of the following groups of key factors could potentially influence DF. These should be considered common examples rather than a comprehensive analysis.

- Building size, construction type and envelope-to-volume ratio can dictate the amount of thermal mass in a building's structure. Taller buildings with concrete structure tend to have larger thermal mass.
- Window-to-wall ratio has significant influence over window solar heat gain, which is often the single largest cooling/heating load component in perimeter zones.
- Space use, internal load density, and occupancy schedules largely determine the internal load profile. Therefore, they can influence DF strategies associated with lighting systems and plug loads. In addition, internal loads typically dominate the cooling load in interior zones and are often significant load components in exterior zones; therefore, they can influence HVAC related DF strategies as well.
- HVAC system type has been discussed in an earlier section. The mechanism through which HVAC load shed or shift is generated is important to understand and determines how much DF can be potentially expected. The zoning in VAV systems can also influence DF because, as explained in the earlier section, interior vs exterior zones and the zone orientation makes a difference in the GTA coasting process. In addition, HVAC equipment and systems in larger buildings (e.g., chilled water systems) tend to have a higher efficiency than those in smaller buildings (e.g., direct expansion based systems).
- A building's operating hours can influence the end-uses and amount of loads available for shedding and shifting at certain times of the day and week. For example, grocery stores often operate in the evenings and on weekends whereas K-8 schools typically close in the afternoon and operate only partially during summer.

4.5 How does building vintage and energy efficiency influence DF?

There are similarities in the approach to understanding how building vintage influences DF and how building type influences DF. This is because building vintage influences DF through many factors related to the energy efficiency (EE) of the building envelope, internal loads, HVAC equipment, control strategies and beyond. Therefore, just like building type, building vintage also should not be treated as a single parameter in assessing how it influences DF.

Of particular importance is understanding how EE influences a building's DF because both EE and DF are critical strategies for decarbonization and also fundamental characteristics of GEBs. There is a myth in the industry that EE will always reduce building baseline load and therefore reduce the potential for load shed and shift. While this is an important observation and applies in many cases, it overlooks the time-varying nature of load change from many EE measures. At least conceptually, some EE measures can be less countervailing to DF than others, or even complementary to DF. Therefore, whether EE improves or reduces DF should be evaluated on an individual measure basis accounting for weather dependencies and EE-DF interactions.

We have explored how various EE measures[17] (see **Table 3**) influence DF through EnergyPlus whole-building energy simulation using a prototype medium office building of ASHRAE 90.1-2004 vintage in one climate zone (3B Warm-Dry) as an example. The impact of each EE measure on load shed intensity from the GTA strategy was shown in **Table 3**. While the internal load and window solar heat gain related EE measures significantly reduce load shed, several other envelope and outdoor air related EE measures were neutral or only have slight negative effect on load shed. Two HVAC control related EE measures even significantly increased load shed in contrary to the aforementioned perception in industry.

Table 3. Energy Efficiency Parameters Included in DF Impact Simulation Study

Category	EE Parameter	Impact on Shed
Envelope	Wall assembly R-value (higher)	Decrease (slightly)
	Roof assembly R-value (higher)	Decrease (slightly)
	Window SHGC (lower)	Decrease
	Infiltration rate (lower)	Decrease (slightly)
Internal Loads	Lighting power density (lower)	Decrease
	Equipment power density (lower)	Decrease
HVAC Systems	**System sizing factor (not over- or under- sized)**	**Non-monotonic**
	Ventilation rate (lower)	Decrease (slightly)
	VAV min damper position (lower)	**Increase**
	Enable supply air temperature (SAT) reset (on vs off)	**Neutral (hot) / Increase (mild)**
	Enable economizer (on vs off)	Neutral
	Cooling setpoint (higher)	**Increase**

[17] These measures are not intended to be comprehensive but are selected to demonstrate common EE measures associated with different aspects of a building.

These two HVAC control EE measures complementing DF are increasing zone cooling setpoints and reducing zone minimum airflow in VAV systems. As shown in **Figure 22**, varying parameters from an inefficient but common value (e.g., 70 ºF cooling set point, 50% minimum airflow rate) to a more efficient, close to best practice value (e.g., 75 ºF cooling set point, 10-20% minimum airflow rate) in both measures can increase shed intensity by ~0.8 W/ft^2, which is as significant as ~20% of an office building's typical peak demand intensity. A VAV system's ability to shed load through a GTA strategy is dependent on the airflow rate to decrease to the minimum set level during zone temperature coasting; the zone temperature may not be able to reach the reset temperature setpoint if the minimum airflow is set high and therefore limiting load shed. Increasing cooling setpoint not only saves cooling energy but also extends the period it takes for zone temperature to reach the new GTA setpoint and increases the average load shed during the event. More detailed methodology and explanation of results of this simulation study are found in an ACEEE conference paper [12].

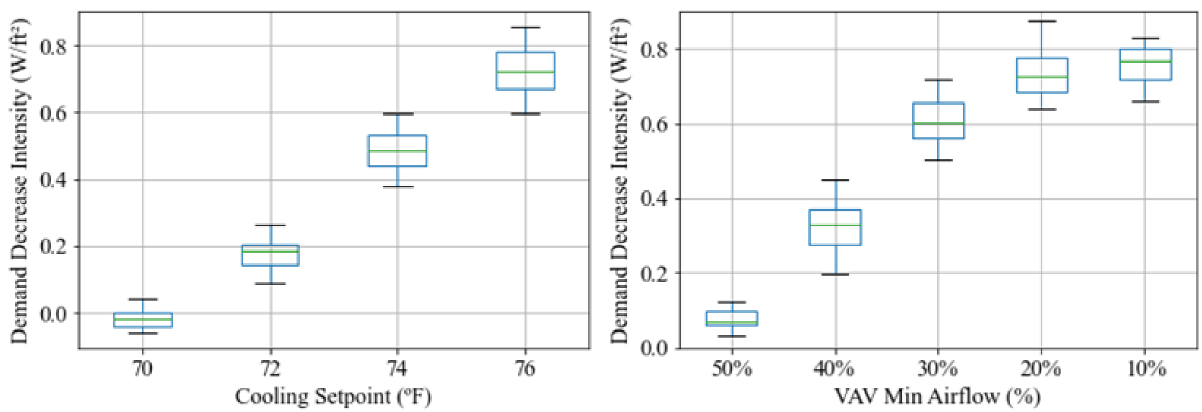

Figure 22. Raising Zone Cooling Setpoint and Reducing VAV Minimum Airflow Rate Significantly Increase Demand Decrease Intensity In a Simulated Prototype Medium Office Building in 3B Climate [12]

4.6 How does thermal mass influence DF?

Thermal mass in buildings buffers against ambient temperature fluctuations and reduces building peak thermal loads by absorbing and releasing heat in line with the daily heating and cooling cycles, thus smoothing out thermal energy flows throughout the day. With respect to using building thermal mass for cooling load control, there have been both simulation and field studies that demonstrated precooling building thermal mass can effectively reduce total and peak cooling demand while maintaining thermal comfort within an acceptable range [13, 14, 15, 16]. One concept of particular interest in GEBs and DF research is using building thermal mass as a "thermal battery" (i.e., energy storage) to actively shift cooling load away from a high electricity price and/or high emission period. This would often involve coupling a GTA strategy to shed load during the high-price/emission period with a pre-cooling period. Although coupling pre-cooling and GTA has been tested in limited existing field studies [13,14], there are two outstanding research questions: (1) does pre-cooling increase load shed over implementing GTA alone? and (2) does a higher level of thermal mass increase load shift over the coupled strategy for the same building? In this section, we explore these questions.

The influence of building thermal mass on DF is challenging to evaluate for a few reasons. First, building thermal mass can be broadly seen as including the mass in the building envelope, floors, and furniture. It is difficult to quantify the mass in an existing building even when detailed design documentation is available. Second, the way the thermal mass in the building wall vs. floor and furniture absorb and release heat can be significantly different (i.e., primarily through radiation as opposed to convection) depending on the construction type, building geometry, orientation, window shading, and many other factors. How thermal mass and pre-cooling together influence DF in a coupled load shift strategy still has to be better understood. There are at least three key aspects of interest:

(a) efficiency of energy storage in thermal mass as reflected in net energy consumption;
(b) influence on load shed during the high-price, high-emission period; and
(c) impact on thermal comfort.

To add to our understanding of thermal mass, we have three research methods available - lab testing, field study, and simulation - each has its advantages and limitations in evaluating the influence of thermal mass. We discuss our exploration and findings using each method below, which should be seen as up to date learning rather than definitive conclusions.

FLEXLAB® test offers an advantage of learning realistic thermal mass effects while isolating other stochastic factors found in real occupied buildings. We tested two strategies ("GTA-only" and "Pre-cooling + GTA") under a range of weather conditions; during part of the test period, an extra layer of garden pavers was added on the floor as a higher thermal mass scenario for comparison. As shown in **Figure 23**, the daily net kWh change from 2ºF pre-cooling and 4ºF GTA is between -4% and 0 showing that this load shift strategy does not increase electricity consumption and is energy efficient. One hypothesis associated with the concept of "thermal battery" is that pre-cooling can increase load shed from GTA alone. While the limited number of lab test data points in **Figure 23** is not sufficient for drawing conclusions, it indicates that such increases are likely to be marginal rather than significant. In addition, the load shed in the higher mass scenario (with pavers) is similar to the no-paver scenario under similar outdoor weather conditions as shown in **Figure 24**.

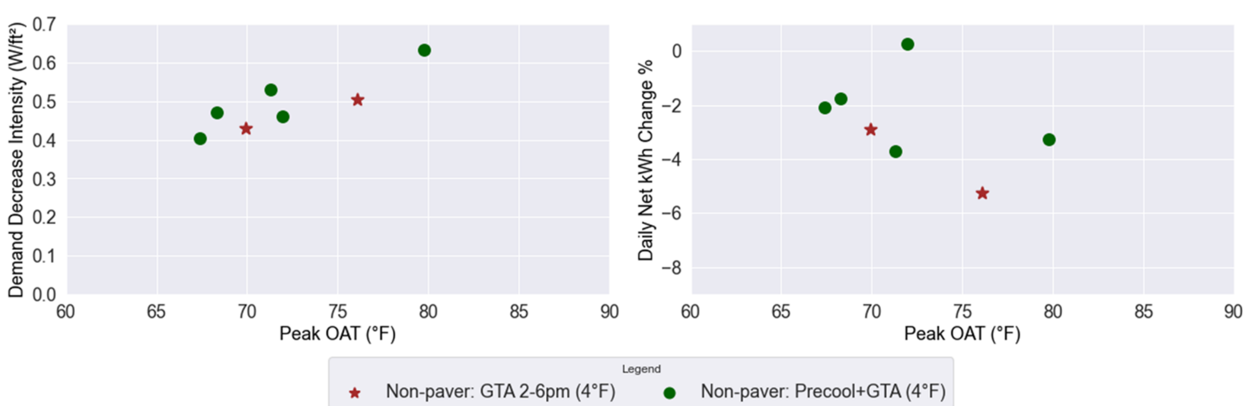

Figure 23. RTU System FLEXLAB® Test Results (without Pavers): Comparing GTA 2-6PM (4°F) Strategy with and without Pre-cooling 10AM-2PM (2°F) (Note: cloudy days and days with high slab temperature issue have been removed from this plot)

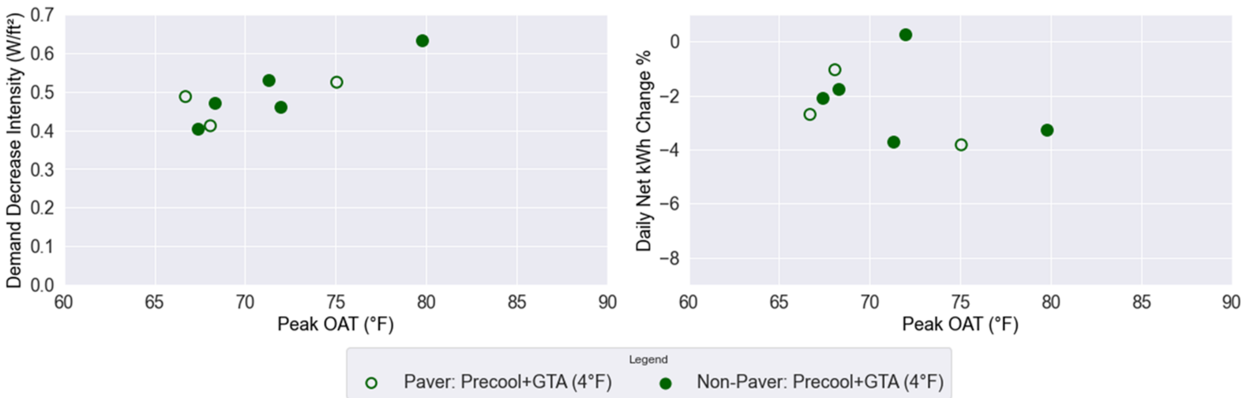

Figure 24. RTU System FLEXLAB® Test for Pre-cooling + GTA Strategy: Comparing Results with and without Pavers (Note: cloudy days and days with high slab temperature issue have been removed from this plot)

As shown earlier, there is a general positive correlation between load shed intensity and outdoor temperature. However, an outlier in this trend revealed another aspect of how thermal mass influences load shift. In **Figure 25**, the load shed intensity on 9/8/2021 was significantly lower and the load take intensity was significantly higher compared to the results on 9/4/2021; the same Pre-cooling + GTA strategy was tested on these two days and their peak temperatures were similar (< 5°F).

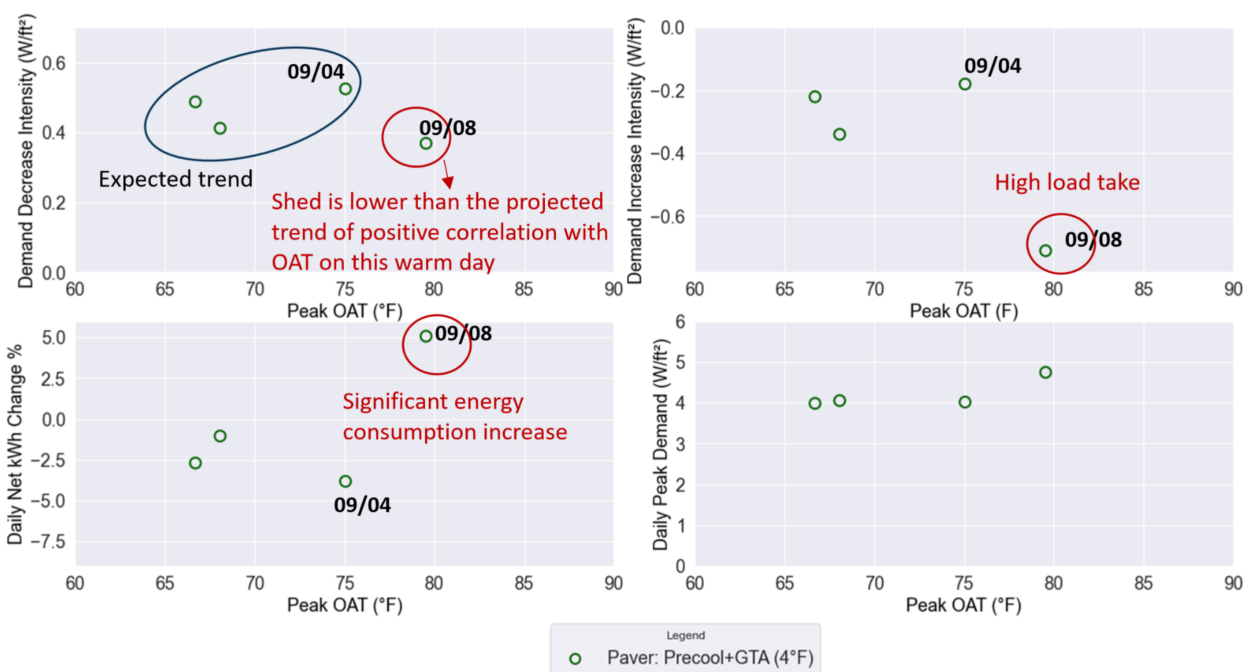

Figure 25. An Outlier (9/8/2021) from the Projected Trend in RTU System FLEXLAB® Test for Pre-cooling + GTA Strategy

When investigating this unexpected result, it was revealed that the previous night was warm (i.e., small diurnal temperature difference) and the rooftop heat pump had also stopped working for more than half a day before that. These two factors led to warm temperatures in the concrete slab of the test cell as shown in **Figure 26** (~ 4°F warmer on 9/8/2021). The results suggest that the slab's inability to cool down overnight had yielded a much higher pre-cooling energy requirement and a more limited ability to subsequently shed load during GTA. This finding was very significant because it implies that non-24/7 operation commercial buildings will likely have significantly less DF on days following building closures such as on Mondays and after holidays. It also implies that a building's DF may be significantly dependent on outdoor temperatures during unoccupied hours if the HVAC system turns off or has aggressive reset schedules.

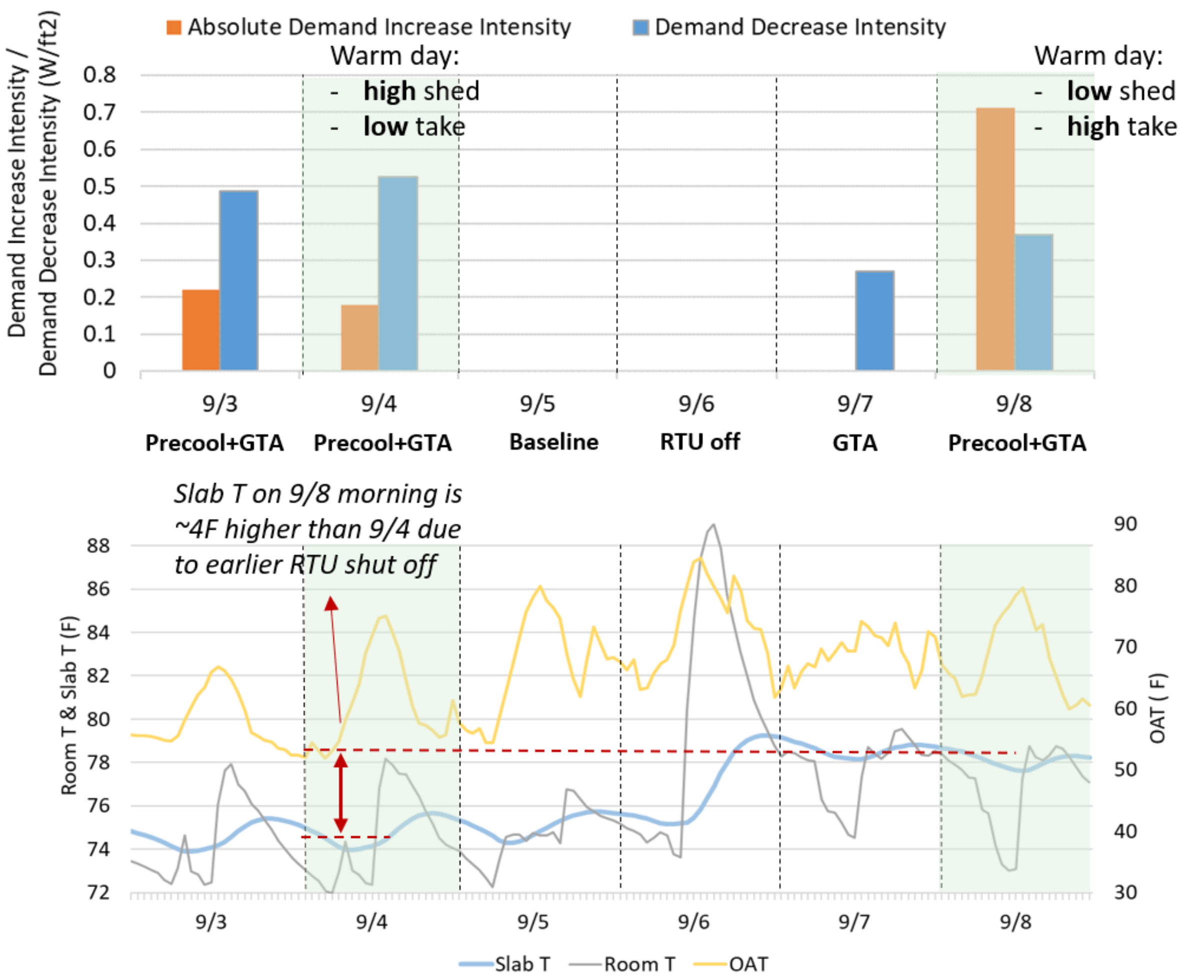

Figure 26. Comparing the Slab Temperature on Two Days with Similar Outdoor Temperatures in a RTU System FLEXLAB® Test for Pre-cooling + GTA Strategy

Field studies are also valuable for understanding energy impact from using building thermal mass to shift cooling load. Currently there are very limited field studies and publicly available data points. In addition, creating counterfactual baselines is a known challenge and real buildings' load profiles are subject to many stochastic factors. They make evaluating thermal mass effects challenging. In the aforementioned 11 office buildings example, each building has implemented 2-3°F GTA while five of them also implemented additional pre-cooling. **Figure 27** shows that in the group of five buildings with pre-cooling (i.e., load shift), the absolute value of load take intensity was much smaller than the load shed intensity; in other words, load shift does not increase electricity consumption and is energy efficient. No significant difference in load shed intensity was observed between the two groups of buildings that implemented pre-cooling vs. not. Although the dataset size is small and cannot be used to draw conclusions, these two observations are consistent with the above discussions on FLEXLAB® test results. Another interesting observation from this field study was that the standard deviation of load shed intensity across multiple events for each building (i.e., DDSD in Figure 26) is smaller and more consistent in the shift group than the shed group. This implies that pre-cooling may be able to improve a building's load shed performance consistency although the data sample is not large enough to be conclusive.

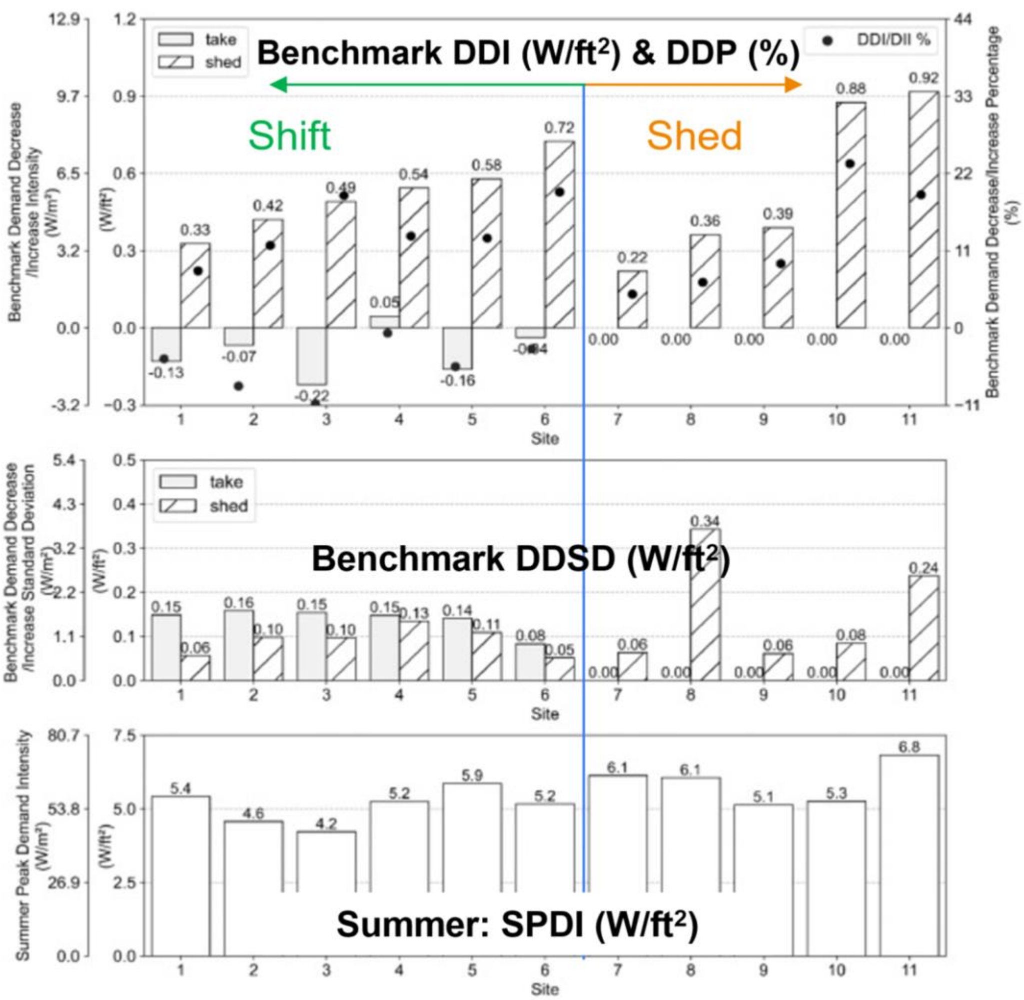

Figure 27. Benchmark Metrics of a Group of 11 Medium Office Buildings That Implemented a GTA Strategy, Divided into Two Subgroups of Six Buildings with Pre-cooling and Five without Pre-cooling

Simulation is another way of understanding the influence of thermal mass. EnergyPlus has been a primary simulation tool used in our study. Compared to FLEXLAB® test and field study, it is much easier to construct comparison scenarios in simulation and to avoid inaccuracies associated with baselines. However, it is not well understood how realistically EnergyPlus models the real-world behavior of thermal mass embedded in different parts of a building. For example, the potential limitations associated with using a single node to uniformly represent the temperatures of both air and furniture in a room and its influence on DF results have not been previously studied. In addition, the internal thermal mass multiplier in EnergyPlus is tested to be a sensitive parameter in DF simulation results. **Figure 28** shows that a higher multiplier value can increase the absolute values of both load shed and load take intensities simulated in single-zone RTU systems. In VAV system simulations, zone temperature coasting rate is sensitive to the multiplier value; zone temperature often does not reach the GTA setpoint when a higher multiplier value is used (default value =1), as shown in **Figure 29**. Choosing an internal thermal mass multiplier value that will yield realistic DF results can be challenging and currently there lacks such guidance.

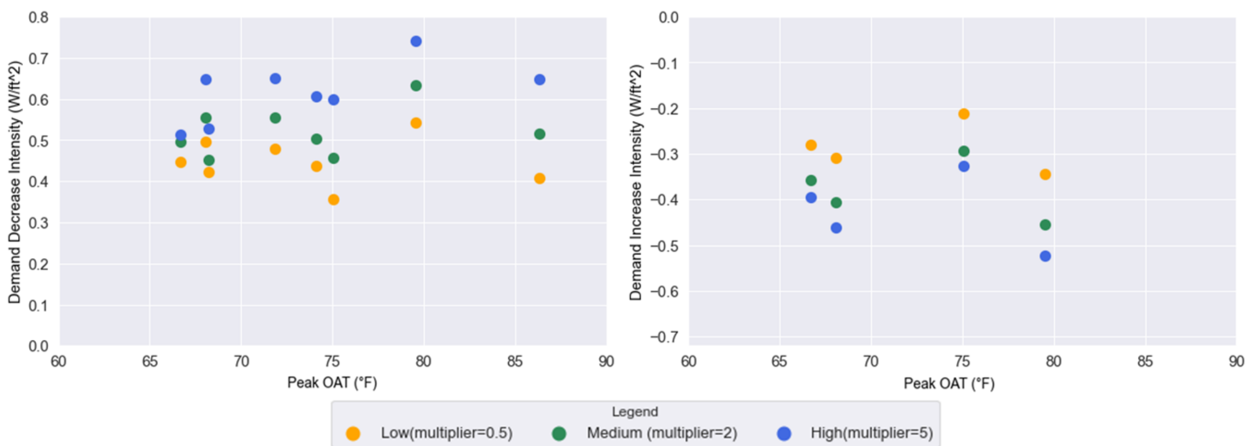

Figure 28. Simulated Sensitivity of Demand Decrease/Increase Intensity of a Single-zone RTU System in FLEXLAB® to Thermal Mass Levels Using a Calibrated Model

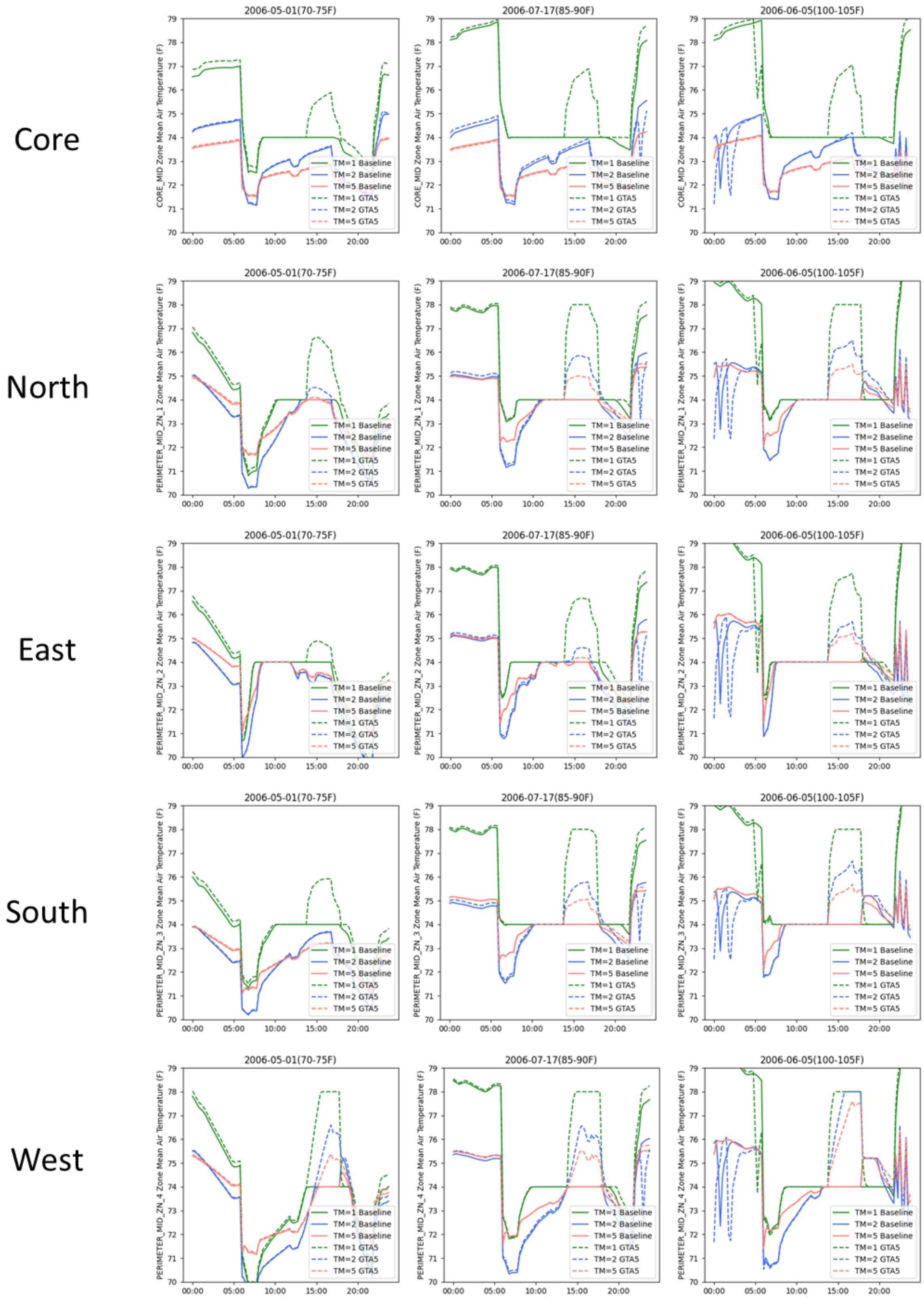

Figure 29. Simulated Different Zone Temperature Coasting Trajectories from Implementing 4°F GTA in a Prototype Medium Office Building in 3B Climate by Varying Thermal Mass Levels

In summary, while thermal mass improves thermal comfort on hot days is a general knowledge, how building thermal mass influences DF in real buildings is a complex topic that requires future research. Our FLEXLAB® test results indicate that adding pre-cooling may marginally increase load shed over GTA strategy, and increasing floor thermal mass may not significantly change load shed. This may have to do with the slow process of relying on surface convection to store energy. Of course, the lab test scenario only represents one set of building characteristics in one climate so the results cannot be extrapolated. For example, building envelope insulation and air tightness levels can play an important role - i.e., intuitively, better insulation and air tightness can help keep the energy stored in thermal mass from pre-cooling for longer. In addition, for non-24/7 operation buildings, we found that not only the current day weather influences DF, the weather of the previous days may also play into the thermal mass temperature and therefore affect the thermal mass' ability to store and release energy.

4.7 "Wild cards" in building DF: operations and commissioning

We have so far discussed how several categories of most common factors may impact a building's DF. It is important to acknowledge that although weather, event timing, HVAC system type, building vintage and efficiency, and thermal mass have covered most of the key influential factors, there are other factors associated with a building's operation and equipment conditions that can also have profound impacts on DF performance. For example, an event that significantly changes the occupancy or internal load in a building can make a significant impact on DF. When HVAC equipment is not functioning properly or the system needs commissioning, it can impact DF significantly. For example, research has shown that supply air temperature and zone temperature sensor faults are among the top three most reported faults for HVAC systems in large buildings; both of these fault types could impact the effectiveness of global temperature adjustment strategies for DF [17]. These operations and commissioning related factors often need to be evaluated on a case-by-case basis, and therefore are not discussed in depth here.

5. Practical Recommendations and Path Forward

Building DF is a special type of grid resource. It has at least the following characteristics. First, as multiple DF potential studies have shown, the technical potential of DF resources is large in absolute scale. Second, there are building DF strategies that are both high-impact and low-cost to implement, which makes building load DF resources more economical compared to other DF resources such as behind-the-meter batteries. The pre-cooling and GTA strategy discussed in depth in this report is one example. Third, resource aggregation across multiple building sites is often either required or preferable for small and medium buildings, and sometimes even for larger buildings, in order to monetize DF for DR and other grid services. Last, but not least, building DF carries a lot more variability and uncertainty in its performance compared to many other grid resources (e.g., onsite generation, batteries), which creates a challenge not only for grid planners and operators but also for the aggregators and building operators.

In this report, we have discussed our findings of how several major categories of influential factors can contribute to the variability and uncertainty in buildings' DF performance from the pre-cooling and GTA strategy based on lab testing, field data analysis, and simulation. These findings may help different stakeholders in strategy planning and performance improvement. Below we discuss the business interest of each stakeholder group and some practical recommendations based on the above research findings.

5.1 Building customers and practitioners

Depending on the building size and ownership type, the party financially responsible for electricity utility cost and DF program participation could be the building owner or the tenant organization; in either case, there may or may not be an assigned building operator to execute DF strategies on their behalf. Furthermore, the party responsible for making retrofits and/or operational changes to building systems that would enable DF may be different from the party responsible for utility cost. Therefore, split incentive may be an issue for some buildings and will need to be addressed. For simplicity of the discussion below, the term "building customer" is used to refer to the parties responsible for enabling DF and for ongoing participation in DF programs.

Building customers may participate in DF programs associated with a local utility or a wholesale market entity. Participation in utility programs can be either directly with the utility or through a third-party aggregator. Wholesale program participation most commonly requires a third-party aggregator. Either way, it is common for the customer to provide an estimated or committed level of load change when they sign up for the program. Some contracts may have a customer performance payment or even an under-performance penalty component; others may not. When such features are applicable to a building customer, it is in the customer's interest to understand how their building's DF performance in each dispatched event is expected to vary from their commitment level given the specifics (e.g., weather, timing, etc.) of each event. With that knowledge, they can avoid potential penalties or unnecessary sacrifice of building service (e.g., occupant comfort). In addition, there may be building practitioners who provide consulting and technical support services to the building customers and therefore have shared interests.

Recommendation C1. For new participating building customers who are expected to estimate a load shed target, it is recommended to identify the primary HVAC system type in the building because the GTA strategy works differently in different HVAC systems. 0.3-0.8 W/ft^2 of load shed is a common range although the actual performance is dependent on weather and other factors. Under similar conditions, the load shed intensity that can be expected from this strategy in single-zone systems (most common in small-medium buildings and retail buildings) is generally higher as compared to VAV systems (common in medium-large buildings).

Recommendation C2. Some energy-efficient HVAC control strategies (e.g., permanently increasing cooling setpoint, reducing zone minimum airflow setting in VAV systems) can potentially increase the load shed from the GTA strategy.

Recommendation C3. Load shed is generally expected to be larger on warmer days although there are exceptions. In buildings where the HVAC system is either shut off or has a deep reset during unoccupied hours, the building customer should expect that the building's ability to shed cooling load may be significantly reduced on those days following a particularly warm evening (i.e., small diurnal temperature difference) or a longer closure (e.g., weekend or holiday). The building customer may consider supplementing the GTA load shed by adding other load shed sources such as lighting and plug load to avoid any under-performance penalty if applicable.

Recommendation C4. When combined with GTA, pre-cooling may be used to improve thermal comfort during shed period and may have a small positive effect on load shed. A small degree of pre-cooling (2°F) coupled with GTA (4°F) often does not increase the building's energy consumption although the conditions described in Recommendation C2 above would be an exception. Building customers should be aware that using pre-cooling in such conditions runs a risk of significantly increasing energy consumption and the building's peak demand.

Recommendation C5. Three other scenarios to be aware of for expected lower performance: 1) cloudy conditions preceding or during a shed event even if the day is warm; 2) events occurring outside traditional DR event period (i.e., noon to 5pm); 3) long events (4 hours or longer). The building customer may consider supplementing the GTA load shed to avoid any under-performance penalty if applicable. Commissioning issues of the HVAC system and equipment should be addressed to ensure proper DF performance.

5.2 Aggregators

Building DF resource aggregators typically have a financial relationship with both the grid entities and the building customers. Aggregators get financially compensated by the grid entities for the grid services they provide and typically pass a portion of the compensation to the contributing building customers; they are often although not necessarily a for-profit organization. Performance payment and under-performance penalty features in aggregator-customer contracts are discussed above. The contract structure between an aggregator and a grid entity can also vary significantly. However, performance payment and/or under-performance (relative to the load change commitment) penalty are often key components in these contracts. Therefore, it is key to the aggregator's business interest to predict customer's performance for each upcoming grid event and develop an optimized dispatch strategy (e.g., the number of customers to be dispatched, staggering strategy) to maximize its net revenue.

In addition, besides aggregating DF resources, some aggregators' business models may also include implementing technological solutions at customer sites. The aggregator may also use knowledge of influential factors in developing their customer recruitment and load control strategies to improve the program impact to transaction cost ratio.

Recommendation A1. If the aggregator develops programs that implement HVAC-based DF strategies, it helps to be aware that (a) the HVAC system type may play a role in the results (see Recommendation C1), (b) buildings with larger window-to-wall ratios and higher internal load densities tend to carry more DF potential, and (c) there is opportunity to increase DF by adjusting efficiency settings in HVAC systems (see Recommendation C2).

Recommendation A2. During a long event (4 hours or longer), each individual building's load shed is expected to decline after the first 1-2 hours, so it may be useful to stagger the dispatch time of various buildings in order to maintain the level of load shed throughout the entire event.

Recommendation A3. It is useful to be aware of a few scenarios where building HVAC based DF strategies are likely to under-perform and be ready to deploy other resources (e.g., batteries, lighting loads) in the aggregator's portfolio to firm up the load shed to meet program requirements. These include 1) cloudy conditions preceding or during a shed event even if the day is warm; 2) events occurring outside traditional DR event period (i.e., noon to 5pm); 3) consecutive hot days with warm evenings; and 4) events following a warm weekend or holiday (see Recommendation C3).

5.3 Grid entities

Currently, DF resources most commonly represent a small fraction of the grid's total resources. They may be dispatched for emergency or economic reasons. Their performance variability has often made it difficult for grid operators to value DF resources with confidence. Therefore, identifying how influential factors may impact buildings' DF performance will help address grid operator's concern and inform better dispatch strategies when DF resources are predicted to over-perform or under-perform based on known factors in advance of an event. In general, grid entities interact with building DF resources through aggregators, so understanding those scenarios stated in Recommendation A3 can help grid entities more accurately anticipate DF resources' performance and align other types of resources to supplement as needed in response to grid emergencies and other needs.

6. References

[1] Satchwell, A., Piette, M.A., Khandekar, A., Granderson, J., Frick, N.M., Hledik, R., Faruqui, A., Lam, L., Ross, S., Cohen, J. and Wang, K., (2021). A national roadmap for grid-interactive efficient buildings. Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States) https://doi.org/10.2172/1784302

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

[3] Li, H., Wang, Z., Hong, T., & Piette, M. A. (2021). Energy flexibility of residential buildings: A systematic review of characterization and quantification methods and applications. *Advances in Applied Energy* 3 (Aug.): 100054. https://doi.org/10.1016/j.adapen.2021.100054.

[4] Neukomm, M., Nubbe, V., & Fares, R. (2019). *Grid-interactive Efficient Buildings Technical Report Series - Overview of Research Challenges and Gaps*. (December). Retrieved on April 8, 2022 from https://www1.eere.energy.gov/buildings/pdfs/75470.pdf.

[5] Gerke, B., Zhang, C., Satchwell, A., Murthy, S., Piette, M., Present, E., Wilson, E., Speake, A., & Adhikari, R. (2020). Modeling the Interaction Between Energy Efficiency and Demand Response on Regional Grid Scales. 2020 ACEEE Summer Study on Energy Efficiency in Buildings (Virtual Conference). https://www.nrel.gov/docs/fy20osti/77423.pdf

[6] Langevin, J., Harris, C.B., Satre-Meloy, A., Chandra-Putra, H., Speake, A., Present, E., Adhikari, R., Wilson, E. J., & Satchwell, A. J. (2021). US building energy efficiency and flexibility as an electric grid resource. *Joule* 5: 1–27. https://doi.org/10.1016/j.joule.2021.06.002.

[7] Gerke, B., Gallo, G., Smith, S., Liu, J., Alstone, P., Raghavan, S., Schwartz, P., Piette, M. A., Yin, R., & Stensson, S. 2020. *The California Demand Response Potential Study, Phase 3: Final Report on the Shift Resource through 2030.* https://doi.org/10.20357/B7MS40.

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

[9] Yin, R., Kara, E. C., Li, Y., DeForest, N., Wang, K., Yong, T., & Stadler, M. (2016). Quantifying flexibility of commercial and residential loads for demand response using setpoint changes. *Applied Energy* 177: 149–164. https://doi.org/10.1016/j.apenergy.2016.05.090.

[10] Liu, J., Yin, R., Pritoni, M., Piette, M. A., & Neukomm, M. (2020). Developing and Evaluating Metrics for Demand Flexibility in Buildings: Comparing Simulations and Field Data.

2020 ACEEE Summer Study on Energy Efficiency in Buildings (Virtual Conference). https://doi.org/10.20357/B7WW34.

[11] Motegi, N., Piette, M. A., Watson, D. S., Kiliccote, S., & Xu, P. (2007). Introduction to Commercial Building Control Strategies and Techniques for Demand Response--Appendices (No. LBNL-59975). Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States). https://doi.org/10.2172/1004169

[12] Yu, L., Liu, J., Yin, R., Piette, M. A., Pritoni, M., Casillas, A., ... & Roth, A. (2022). When Do Efficiency and Demand Flexibility Go Hand-in-hand?. 2022 ACEEE Summer Study on Energy Efficiency in Buildings. https://doi.org/10.20357/B7GG6D

[13] Xu, P., & Zagreus, L. (2009). Demand shifting with thermal mass in light and heavy mass commercial buildings . Proceedings of 2009 ASHRAE Annual Conference, Louisville, KY, June 20-24, 2009. https://doi.org/10.2172/988082

[14] Xu, P., Haves, P., Piette, M. A., & Braun, James. (2006). Peak Demand Reduction from Pre-Cooling with Zone Temperature Reset in an Office Building. 2006 ACEEE Summer Study on Energy Efficiency in Buildings. https://www.osti.gov/servlets/purl/980751.

[15] Braun, J. E., Lawrence, T. M., Klaassen, C.J., & House, J. M. (2002). Demonstration of load shifting and peak load reduction with control of building thermal mass. 2002 ACEEE Buildings vol. 3 55–68. https://www.eceee.org/media/uploads/site-2/library/conference_proceedings/ACEEE_buildings/2002/Panel_3/p3_5/paper.pdf

[16] Xu, P. (2006). Evaluation of Demand Shifting Strategies with Thermal Mass in Two Large Commercial Buildings. Proceedings of SimBuild Conference 2006, Cambridge, Massachusetts (United States), August 02-04, 2006. https://escholarship.org/uc/item/5f1914k0

[17] Crowe, E., Chen, Y., Reeve, H., Yuill, D., Ebrahimifakhar, A., Chen, Y., Troup, L., Smith, A., Granderson, J. (2023). Empirical Analysis of the Prevalence of HVAC Faults in Commercial Buildings. Science & Technology for the Built Environment.

Appendix A. Metrics Definitions for Shed and Shift

The tables and figures in this appendix are selected form [2] to serve as background information on DF metrics definitions and calculation procedures

Table A- 1. Single-event Metrics for Demand Decrease (Load Shed)

Metrics	Formula / Definition	Unit
E1: Net Building Consumption Change Percentage (24 hours)	= Net daily kWh change / Baseline daily kWh consumption x 100%	%
D1: Demand Decrease Intensity (DDI)	= D2 / Floor Area	W/ft^2 (or W/m^2)
D2: Demand Decrease	= Average demand decrease during a single "shed" period	kW
D3: Demand Decrease Percentage (DDP)	= D2 / Baseline average demand during "shed" period	%

Table A- 2. Single-event Metrics for Demand Increase (Load Take)

Metrics	Formula / Definition	Unit
I1: Demand Increase Intensity (DII)	= I2 / Floor area	W/ft^2 (or W/m^2)
I2: Demand Increase	Average demand increase during a single "take" period	kW
I3: Demand Increase Percentage (DIP)	= I2 / Baseline average demand during "take" period	%

Table A- 3. Benchmarking Metrics for Shed and Shift

	Benchmarking Metrics	Formula / Definition	Unit
Peak (for reference)	**SPDI:** Summer Peak Demand Intensity	= Summer hourly (or 15-min) peak demand / Floor area	W/ft^2 (W/m^2)
	WPDI: Winter Peak Demand Intensity	= Winter hourly (or 15-min) peak demand / Floor area	W/ft^2 (W/m^2)
Shed	**Benchmark DDI**[1] (Benchmark Demand Decrease Intensity)	= Average of *the five highest per-event D1*[2] *values* ("top-5 shed")	W/ft^2 (W/m^2)
	Benchmark DDP[1] (Benchmark Demand Decrease Percentage)	= Average of the per-event D3[2] values of the "top-5 shed"	%
	Benchmark DDSD[1] (Benchmark Demand Decrease Intensity Standard	= Standard Deviation of the "top-5 shed"	W/ft^2 (W/m^2)

	Deviation)		
Take (apply to Shift only)	**Benchmark DII**[1] (Benchmark Demand Increase Intensity)	= Average of the I1[3] values associated with the "top-5 shed"	W/ft^2 (W/m^2)
	Benchmark DIP[1] (Benchmark Demand Increase Percentage)	= Average of the I3[3] values associated with the "top-5 shed"	%
	Benchmark DISD[1] (Benchmark Demand Increase Intensity Standard Deviation)	= Standard Deviation of the I1[3] values associated with "top-5 shed"	W/ft^2 (W/m^2)

[1] These benchmarking metrics values are associated with a given set of *Attributes* values such as DF strategy used, event duration, time of day, day of week, year, baseline method, and weather condition. A separate set of metrics may also be used for the cooling and heating seasons.

[2] See Table A1.

[3] See Table A2.

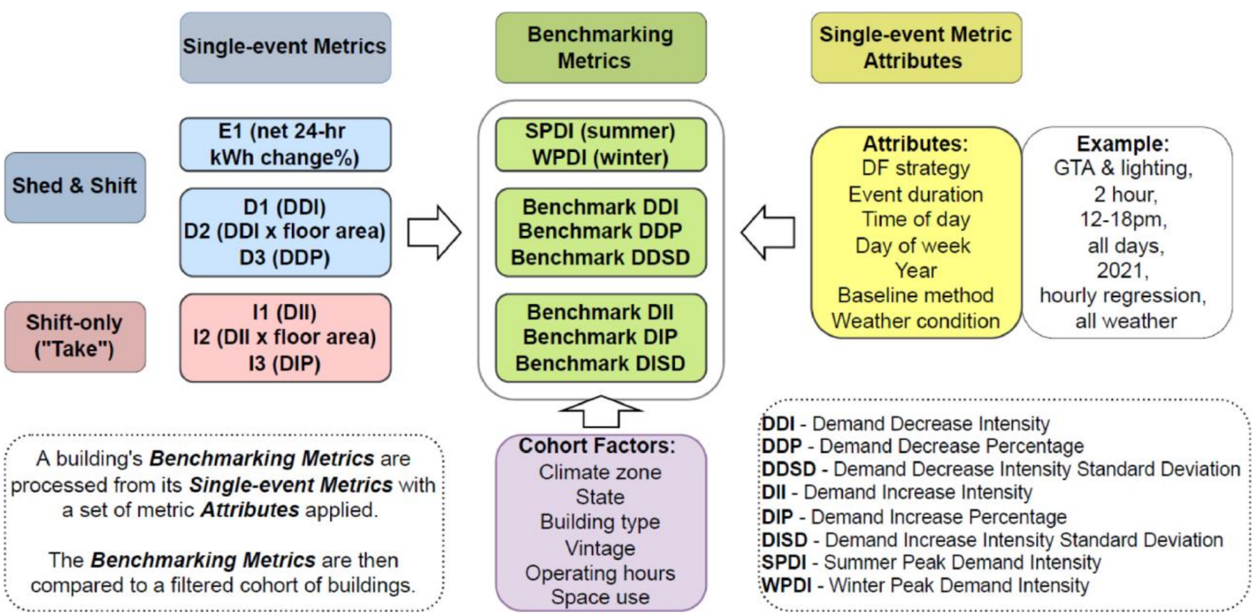

Figure A- 1. DF Benchmarking Metrics Framework with Three Building Blocks: Single-event Metrics, Metrics Attributes, and Benchmarking Metrics.

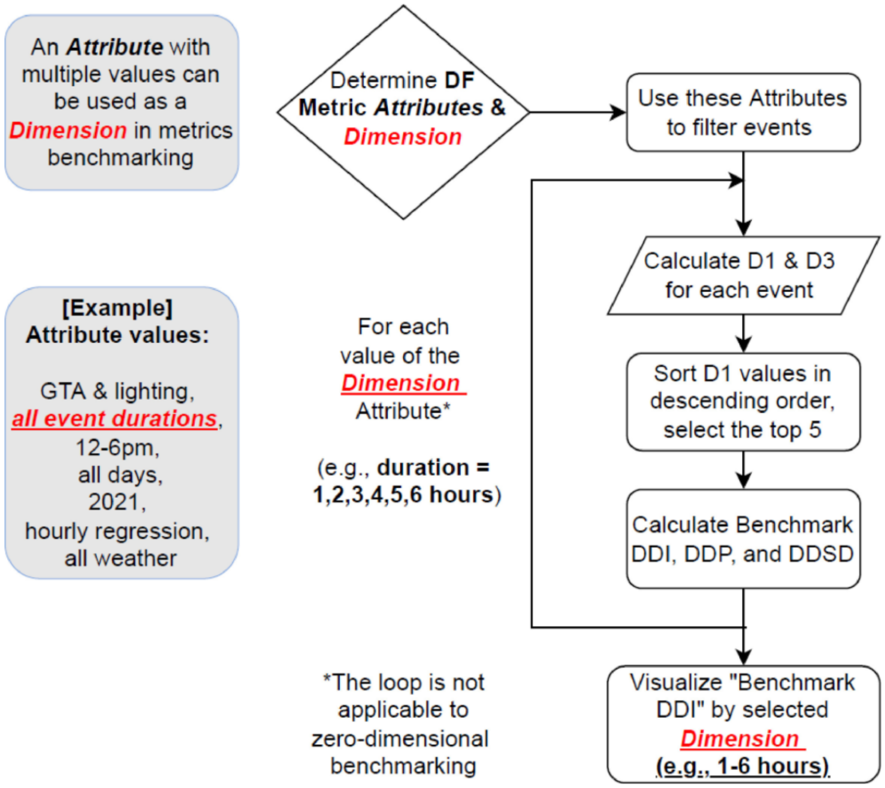

Figure A- 2. DF Benchmarking Metrics Calculation Procedure Diagram.