Quantifying Demand Flexibility Technical Potential in Commercial Buildings: A Scalable Approach for Grid Stability, Energy Savings, and Emissions Reduction

Tarun Arasu, Abigail Andrews, Jennifer Senick
Center for Urban Policy Research, Rutgers University, New Brunswick, NJ

ABSTRACT

The electric grid is under increased stress due to aging infrastructure, extreme weather, electrification, the growth of data centers, and the integration of intermittent renewables. Non-wire alternatives may be a cost-effective mechanism to reduce grid strain. Buildings hold a unique opportunity to serve as a grid resource as they are large electricity consumers. This study quantifies peak load reduction and the energy savings potential achievable through strategic demand flexibility (DF) in the commercial building stock of a Mid-Atlantic State and provides a step-by-step framework for other jurisdictions to leverage building stock data to measure their building stock's technical DF potential.

We use building energy modeling (BEM) to evaluate various DF measures (e.g., peak-hour load shedding and thermostat adjustments) across the most common commercial building type-code combinations of the Mid-Atlantic state for a bottom-up building stock approach to assess the technical potential of DF. Our results indicate that DF strategies can reduce peak summer demand in the modeled state by up to 2.3%, contributing to a total reduction of 239 MW under specified DF scenarios.

This approach to quantifying the DF potential of a commercial building stock, utilizing publicly available information and desktop data analysis, provides actionable information to guide future program and policy design towards accomplishing long-term energy resilience, grid management, and decarbonization strategies. Furthermore, it can improve the targeting of conventional, fielded baseline and potential studies by identifying areas where public data is not robust enough or missing for specific building types, in turn, improving the accuracy of building stock information.

INTRODUCTION

The buildings sector is a focal point of decarbonization as buildings use over 75% of all electricity consumed in the United States (U.S.) and are responsible for 35% of the U.S.'s carbon dioxide emissions (EIA 2024). To address this many individual states are pursuing building decarbonization goals. States across the Northeast and Mid-Atlantic have enacted policies to phase out fossil fuel use in buildings and vehicles and reach over 80% reduction in economy-wide emissions by 2050. These targets will require large-scale electrification of the built environment. Notably, building decarbonization efforts are accelerating electricity demand growth at the same time that the share of variable renewable energy on the grid is increasing. As solar and wind penetration rises, so does the challenge of matching supply and demand in real time.

Demand flexibility (DF) is the ability to shift, shed, or modulate electricity use in response to grid signals and has emerged as a complementary strategy to energy efficiency and infrastructure upgrades for ensuring grid reliability, integrating renewable energy, and lowering emissions (Satchwell et al. 2021). In this paper, we primarily focus on DF. While earlier literature predominantly discusses Demand Response (DR), which generally implies targeted event-based load shifting. DF encompasses these DR strategies but also includes continuous or automated actions such as preconditioning, thermal storage, and other load-shaping methods.

Commercial buildings are of particular interest in DF efforts due to their substantial, predictable, and controllable energy consumption. The U.S. Department of Energy (DOE) has identified Grid-Interactive Buildings (GEBs) as essential for achieving these goals, estimating that GEBs could provide up to 80 GW of demand-side flexibility by 2030 (Satchwell et al. 2021).

Early estimates of DF potential relied heavily on regression-based techniques to infer building responsiveness from observed load profiles and pricing events. Mathieu et al. (2011, 2013) initially introduced the use of load variability models to evaluate DR baselines and customer responsiveness using historical data from industrial and commercial facilities. Other statistical models have extended these methods to estimate DF potential at a macroscale. Hledik et al. (2019) developed a national-level modeling framework using econometric methods to forecast DR enrollment and dispatch potential under various pricing scenarios. This paper estimated that the U.S. has approximately 200 GW of economically viable flexible load, emphasizing residential and commercial sectors with advanced metering infrastructure. National assessments such as the Federal Energy Regulatory Commission's (FERC) 2009 DR Potential Report estimate a peak reduction potential of 150 GW from DR programs across sectors. EPRI (2017) estimated 175 GW of technical summer peak reduction from DR by 2030. But it is important to note that these methods may overlook building specific details and technical or operational constraints, making it harder to estimate the actual flexibility available in different building types.

More recently, several state-level assessments have emerged to quantify demand flexibility potential at higher resolution than national studies. California's Demand Response Potential Study (LBNL, 2022-24) estimates 12–21 GW of shift and shed flexibility by 2035, based on bottom-up load-shape clustering of smart-meter data calibrated to projected electrification. In New York, Brattle's Grid Flexibility Potential Study (2025) projects 10 GW of achievable summer peak reduction by 2040, combining Advanced Metering Infrastructure (AMI) based load profiles with scenario-driven device adoption modeling (Brattle 2025). Massachusetts' 2025–27 Energy Efficiency and Demand Response Potential Study uses multi-source aggregated data to estimate ~1.4 GW of economic DR potential by 2030 (Guidehouse 2024). However, these studies often rely on utility data that may not be easily available and on field surveys, which although accurate, require a high capital investment.

Another major stream of research uses building energy modeling to estimate the technical potential of DF. Cai and Braun (2019) performed a nationwide parametric simulation study using DOE reference building models for seven commercial building types in five climates. They evaluated different strategies (e.g. adjusting thermostat schedules, dimming lights) under various utility rate structures. Their simulation-driven approach enabled mapping of which regions and building types are most promising for DF notably finding that thermostat setpoint scheduling was the most effective measure, whereas lighting and shading control alone yielded only marginal peak reduction. More recently, Jiang et al. (2023) applied EnergyPlus simulations across various commercial building prototypes to evaluate strategies such as thermostat setbacks, plug load controls, and pre-cooling. Their results showed peak demand reductions between 10–20% depending on the building type and strategy.

To extrapolate individual building results to broader geographies, large-scale building stock models like National Renewable Energy Labortory's (NREL) ComStock and ResStock have become central tools. These models combine bottom-up simulations of thousands of representative building prototypes with stock data on regional construction practices, climate zones, and end-use saturation. NREL (2021) developed the ComStock dataset to characterize hourly end-use load profiles across U.S. commercial buildings. Langevin et al. (2021) extended this by simulating flexibility measures such as load shifting, HVAC control, and lighting dimming along with energy efficiency measures on the national commercial building stock to assess hourly grid impacts under high renewables scenarios. Their findings suggest that up to 200 GW of peak load reduction is technically achievable by 2030 from efficiency and flexibility measures combined. These approaches provide high-resolution, reproducible estimates and can account for regional heterogeneity. However, they are computationally intensive and depend heavily on input assumptions about technology performance, occupant behavior, and future adoption trajectories.

These studies collectively and consistently find that commercial buildings have substantial and controllable load flexibility, particularly in HVAC systems, lighting, and plug loads. While prior research has provided important insights through pilot studies, field studies, and AMI based analyses, these approaches often require extensive data collection efforts, utility partnerships, or proprietary modeling platforms For example, large-scale pilot and field studies such as LBNL's California Phase 4 Demand Response Potential Study took over four years from meter installation (2018–2019) to final public release (April 2024) (Gerke et al. 2024). Studies based on AMI often require proprietary utility data and face integration challenges, limiting transparency and replicability (LBNL 2024). In econometric or regression-based models driven by price elasticity, analysts must apply subjective judgment when extrapolating from historical pilot customer participation to broader populations (FERC 2009). It is also important to note that these approaches tend to yield highly accurate estimates because of extensive on-ground research and real-world world assumptions.

In contrast, our approach, as discussed in this paper, leverages publicly available data sources and tools to develop a scalable, replicable framework that jurisdictions can employ quickly and cost-effectively. Rather than replacing field studies, this simulation-based framework can help inform and refine them identifying where detailed fieldwork may be most valuable, filling data gaps, and providing preliminary estimates that support program design and targeting. In doing so, this complementary approach offers a practical tool for policymakers and utilities to better understand demand flexibility potential at the state level while preserving opportunities for more detailed validation through field-based studies.

METHODS AND FRAMEWORK

To estimate the DF potential from a model Mid-Atlantic state's commercial building stock, we employed a bottom-up modeling framework integrating building energy simulations with regional building stock and grid-level data. The presented methodology and associated case study may be relevant to other states to provide a low-cost first estimate of the technical demand flexibility potential of the commercial building stock. A state may take the following steps to do so (Figure 4). We opted to focus on the Mid-Atlantic as the region has balanced thermal load demands for summer cooling and winter heating, and in a highly electrified future, this balance could give rise to dual seasonal peak loads, presenting unique challenges for grid planning and demand management. Further, this region is highly urbanized, with the commercial sector forming a significant part of both the economy and electricity demand. In 2024, the commercial sector accounted for approximately 40.9% of total electricity consumption in the Middle Atlantic region (145,510 million kWh out of 355,526 million kWh), compared to 36.2% nationally (1,434,007 million kWh out of 3,961,890 million kWh) (EIA, 2024). Peak electricity demand in these states typically aligns with commercial operating hours, especially during summer, due to the current prevalence of cooling loads (NYISO Gold Book 2023). However, the modeling framework itself is generalizable and can be readily applied to other states using publicly available building stock and grid data.

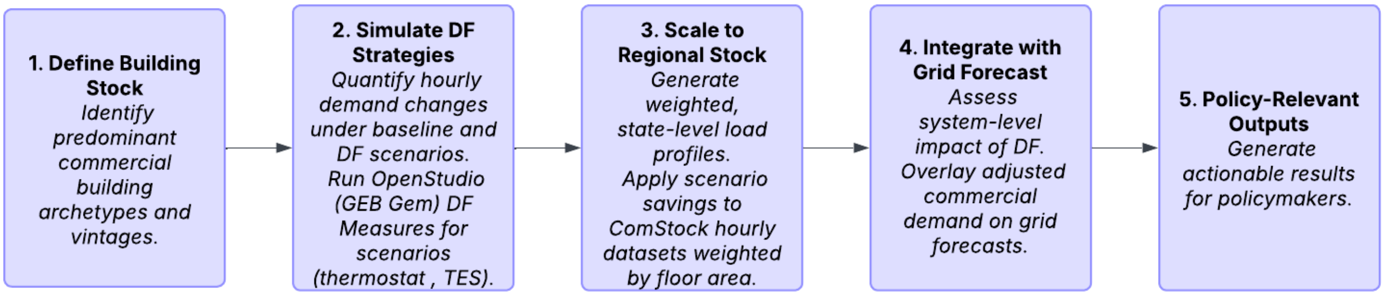

Figure 1. Bottom Demand Flexibility Quantification Framework

STEP 1: DEFINE BUILDING STOCK

We first identify the most common building use types in the modeled state using the National Renewable Energy Laboratory's (NREL) ComStock dataset. ComStock provides energy models that represent 61% of the U.S. commercial building floor area. We use ComStock data at the state level to identify the 25 most common combinations of building use type and associated energy code vintage (energy-code "vintage" refers to the code version associated with each building's last major retrofit). To ensure the analysis captures the majority of the building stock, we identified and selected type-code combinations that represent 1% or more of the state's total commercial floor area. Type-code combinations are defined by both building use type (e.g., full-service restaurant, office) and energy code vintage (e.g., ASHRAE 90.1-2004, 2007, 2013).

After identifying these type-code combinations, we used EIA's 2018 Commercial Building Energy Consumption Survey (CBECS) microdata for the Mid-Atlantic to determine the percentage of floor area by property type that has one- or two-way grid communication. This helped us assess which building types are likely to have the capability to participate in DF programs.

Table 1. Step 1 Methodss

Goal	Identify predominant commercial building type-code combinations and vintages
Process	- Use ComStock to filter and rank building-type/code combinations by floor area contribution - Select those representing ≥1% of the state's total commercial floor area. Use CBECS to estimate grid communication readiness for each property type. - Use CBECS to estimate BAS penetration and grid communication readiness for each property type.
Data/Tools	ComStock metadata with State Specific filters, EIA 2018 CBECS microdata
Output	List of most common type-code combinations covering majority of floor area , with associated DF communication and automation-readiness indicators

STEP 2: SIMULATE DF STRATEGIES

We used OpenStudio to simulate energy use in each of the 25 selected building type-code combinations under four DF strategies, in addition to a baseline case. OpenStudio uses EnergyPlus, which calculates hourly building energy demand by modeling heat transfer through the building envelope, internal gains, HVAC system response, and outdoor weather conditions. In our case, we used state-specific EnergyPlus Weather (EPW) files to represent local hourly temperature, humidity, and solar radiation, corresponding to the ASHRAE climate zones (e.g., 4A or 5A) present in the modeled state. We implemented the DF scenarios by applying pre-built OpenStudio "measures" which are scripts that modify building model parameters to simulate various energy management strategies. We utilized the Grid-interactive Efficient Buildings (GEB) measure gem developed by NREL to systematically apply four DF strategies as summarized below:

Table 2. Demand Flexibility Scenarios Modeled

Scenario	**Demand Flexibility Strategy**	**Description**
S1	Peak Hour Thermostat Adjustment	Thermostat setpoint increased by 4°F in summer and decreased by 2°F in winter during 2–8 PM peak window (June–September for cooling, November – March for heating).
S2	Precooling and Preheating Only	Thermostat adjusted by ±4°F for four hours before 2–8 PM, without any change during the peak itself.
S3	Peak Thermostat + Precooling/Preheating	Same as S1, with additional preconditioning: thermostat adjusted by ±4°F for four hours before the 2–8 PM peak period.
S4	Thermal Energy Storage (TES)	Cooling load stored during 8 PM–8 AM using idealized TES, discharged between 8 AM–8 PM to meet building loads (applied to Large Office buildings only).

We chose the 2-8 PM slot based on the Model Mid-Atlantic State's commercial sector load profiles which exhibit higher electricity demand during summers relative to other times of the day and year. These strategies were selected based on findings from prior literature showing that thermostat adjustments, preconditioning, and thermal energy storage (TES) are among the most effective and widely studied methods for reducing electricity use in commercial buildings during peak periods (Hledik et al. 2019; Kiliccote et al. 2006; Poolla et al. 2020). They specifically target HVAC loads, which represent a large and controllable share of electricity consumption in these buildings. In addition, these strategies can be automated through existing building automation systems (BAS) or enabled via two-way communication technologies, making them suitable for both utility-driven demand response and building-level energy management programs (DOE 2019; Piette et al. 2007).

Post the simulations, we used the simulated electric hourly load profiles to calculate the relative hourly electricity savings for each scenario and building-code type:

$$\text{Relative Saving}_{b,t}^{(s)} = \frac{D_{b,t}^{\text{baseline}} - D_{b,t}^{(s)}}{D_{b,t}^{\text{baseline}}} \quad \text{....(1)}$$

Where $D_{b,t}^{(s)}$ is the simulated demand for building type b, time t, and scenario s, and $D_{b,t}^{\text{baseline}}$ is the baseline demand.

Table 3. Step 2 Methods

Goal	Quantify hourly demand changes under baseline and DF scenarios
Process	-Run OpenStudio simulations for 4 DF Strategies using GEB DF Measures from NREL
Data/Tools	OpenStudio, GEB Measure Gem, EnergyPlus Weather (EPW) files, ASHRAE climate zones
Output	Hourly relative savings (%) for each building-type and scenario combination

STEP 3: SCALE TO REGIONAL STOCK

The idea is to aggregate building-level results to develop a weighted hourly demand profile representative of the entire state. We applied the relative savings by scenario and building type to baseline hourly electricity consumption profiles from ComStock. For each building type b, code vintage v, and hour t, the load $L_{b,v,t}$ represents the total modeled demand for that segment, scaled up using floor area estimates and building counts from Comstock data.

We then aggregated across vintages for each building using the below formula:

$$L_{b,t}^{(s)} = \sum_{v} \left(L_{b,v,t} \cdot \left(1 - \text{Relative Savings}_{b,v,t}^{(s)} \right) \right) \quad \text{....(2)}$$

This provided a stock-adjusted, scenario-specific hourly load profile for each building type under each DF strategy.

Table 4. Step 3 Methods

Goal	Generate weighted, state-level load profiles
Process	- Match simulated buildings to ComStock type-code combinations - Apply floor area weights to time series
Data/Tools	ComStock time-series data, building counts, Python for aggregation
Output	Scenario-specific hourly demand profiles for the entire commercial stock in model state.

STEP 4: INTEGRATE WITH GRID FORECAST

To assess the grid-level impacts of commercial building DF, we applied the absolute hourly electricity savings derived from prototype simulations to NREL's Cambium 2022 Mid-Case hourly load forecast for the modeled state in 2024. Cambium is produced annually by the National Renewable Energy Laboratory's Standard Scenarios team, which couples the ReEDS capacity expansion model with hourly dispatch simulations to generate state-level 8,760-hour load, generation, and marginal emission profiles (Gagnon, Cowiestoll, & Schwarz 2023). Specifically, for each hour, we calculated the total commercial sector savings by summing the difference between the baseline and scenario-adjusted electricity demand across all building types. These hourly savings were then subtracted from Cambium's projected commercial load for that hour.

This method allowed us to isolate the impact of DF on the commercial sector without altering assumptions related to other sectors, like residential or industrial. By directly applying the simulation informed absolute savings to Cambium's load forecasts, we maintained consistency in all non-commercial components of the grid forecast and avoided introducing structural changes beyond those modeled in the building simulations. This integration offers a clear estimate of how Commercial Building DF strategies could shape the future load curve at the state level.

This approach enables granular, scalable estimation of commercial building DF at the state level, based on simulation derived savings that utilize real-world building stock data.

Table 5. Step 4 Methods

Goal	Assess system level impact of DF
Process	- Overlay regional load curves on ISO/Cambium forecasts - Compare baseline and DF scenarios
Data/Tools	Cambium grid projections, Python (pandas/plotly), time-aligned hourly datasets
Output	Hourly grid load curves with/without DF effects for the commercial sector

STEP 5: GENERATE POLICY-RELEVANT OUTPUTS

Translate technical modeling results into actionable insights for program design and policymaking. Summarize demand savings, peak demand reductions and load shifting potential. This data can support TOU rate design, incentive planning, and load shaping programs.

Table 6. Step 5 Methods

Goal	Generate actionable insights for policymakers
Process	- Calculate % peak savings - Determine shifted kWh
Data/Tools	Python, R, Excel/Power BI for visualization
Output	Tables and graphs for peak reduction, total savings, timing of demand shifts

RESULTS AND DISCUSSION

BUILDING-LEVEL RESULTS

We simulated energy use for the 25 commercial type-code combinations in the modeled state, including offices, retail stores, schools, warehouses, and restaurants, under four DF strategies designed to reduce or shift electricity use during the 2–8 PM peak window in summers (June-September). These included thermostat adjustments, pre-cooling/heating, and TES for large buildings with central chilled-water systems.

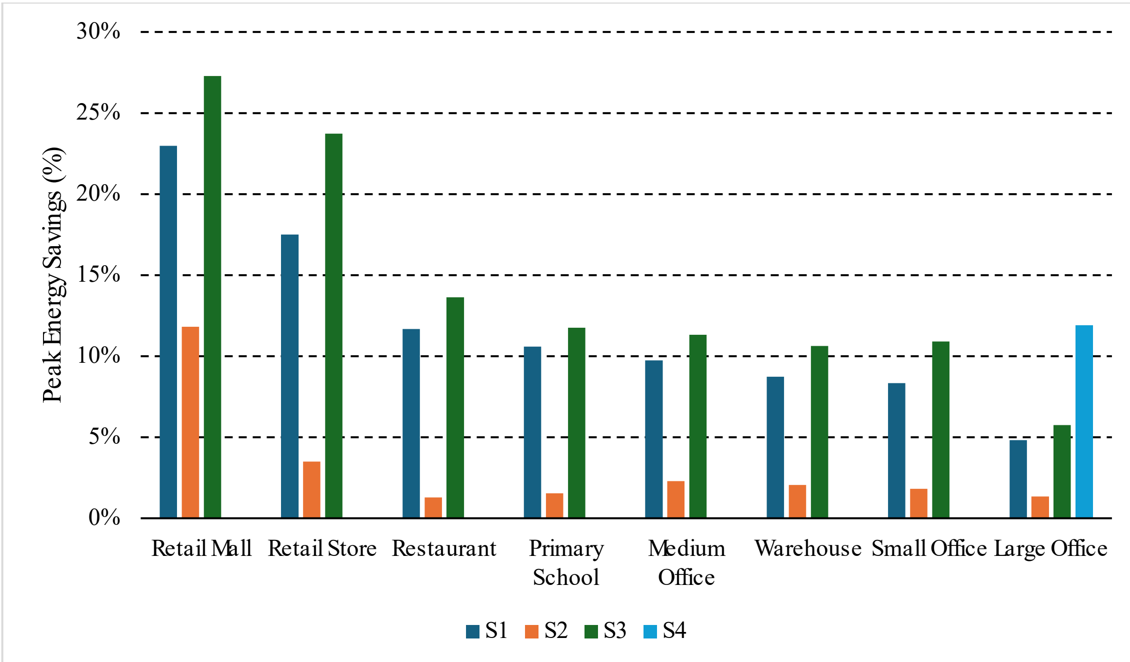

Figure 2. Summer Peak Electricity Savings by Scenario and Building Type

Figure 2 presents the percentage reduction in electricity use during the 2–8 PM window , averaged across all building code vintages. The Peak Hours Thermostat Adjustment (S1) strategy led to consistent savings across most building types, ranging from 4% to 20%. These savings were highest in Restaurants, Primary Schools, and Retail buildings. The Precooling and Preheating (S2) strategy, which shifts HVAC activity earlier in the day without modifying peak-hour settings, produced more modest savings, generally in the 1–10% range. When combined (S3), these two approaches yielded larger peak reductions, exceeding 10% in all building types except for Large Offices, reflecting the additive effect of pre-conditioning and peak-time adjustments. The TES strategy (S4), modeled only for Large Office buildings with central chilled-water systems, delivered the highest peak-period savings, above 10%, by shifting cooling load entirely to off-peak hours.

Strategy S1 produced little to no increase in off-peak usage, making it attractive for buildings with shorter daily schedules. In contrast, the S2 and S3 strategies led to small but visible increases in off-peak consumption (often under 5%) as HVAC systems ran earlier to pre-condition spaces. The TES strategy (S4) showed the largest off-peak increases (10%) during early morning hours, when cooling loads were shifted to charge the thermal storage systems. This load shifting remains beneficial from a grid perspective, as it reduces stress during peak hours and aligns better with low-cost, off-peak generation.

Given our focus on HVAC-based strategies, buildings with larger HVAC loads and longer daily operation, such as offices and retail spaces, consistently offered higher DF potential. Small Office, Retail, and Warehouse buildings showed notable reductions with basic thermostat changes, while Large Office buildings benefited most from TES. Buildings with simpler systems (e.g., Primary School) demonstrated more limited flexibility, though some savings were still observed under thermostat adjustments.

GRID LEVEL IMPACT

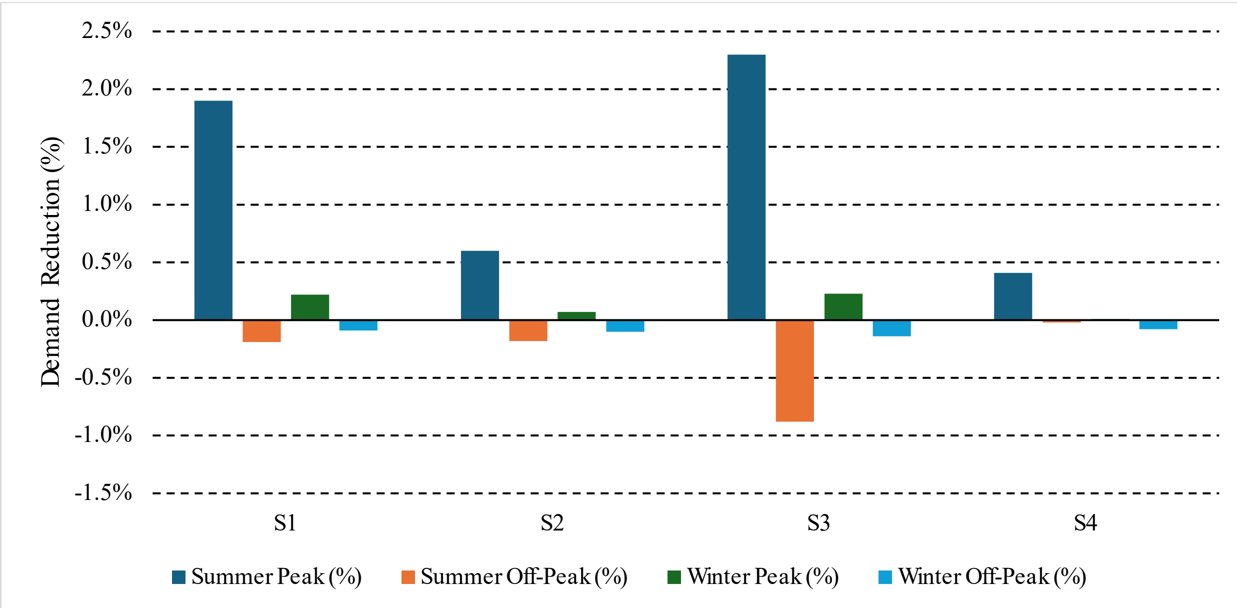

Figure 3: Change in Average Electricity Demand by Scenario and Period

The Peak Thermostat + Precooling/Preheating Strategy delivers the largest reduction in peak demand (−2.3%), though it also results in an increase in off-peak usage (+0.8%). This suggests a load-shifting dynamic that could be beneficial under appropriate rate structures or grid conditions (Figure 3).

S1 strategy, in comparison, provides a nearly equivalent peak demand reduction (−1.9%) but with a smaller off-peak increase (+0.19%), indicating it may be more suitable for buildings with tighter operational constraints or limited thermal mass.

The S2 strategy shows lower peak savings but redistributes cooling demand into earlier hours, often resulting in elevated off-peak loads. This strategy may inadvertently introduce new system peaks in the late afternoon or early evening, depending on the diversity of building response. The S4 Strategy is only applied to Large Office prototypes with chilled-water systems and, while effective in those buildings, has limited system-level impact due to its narrow applicability. As expected, winter season effects across all strategies are negligible, as the modeled flexibility strategies primarily affect cooling loads.

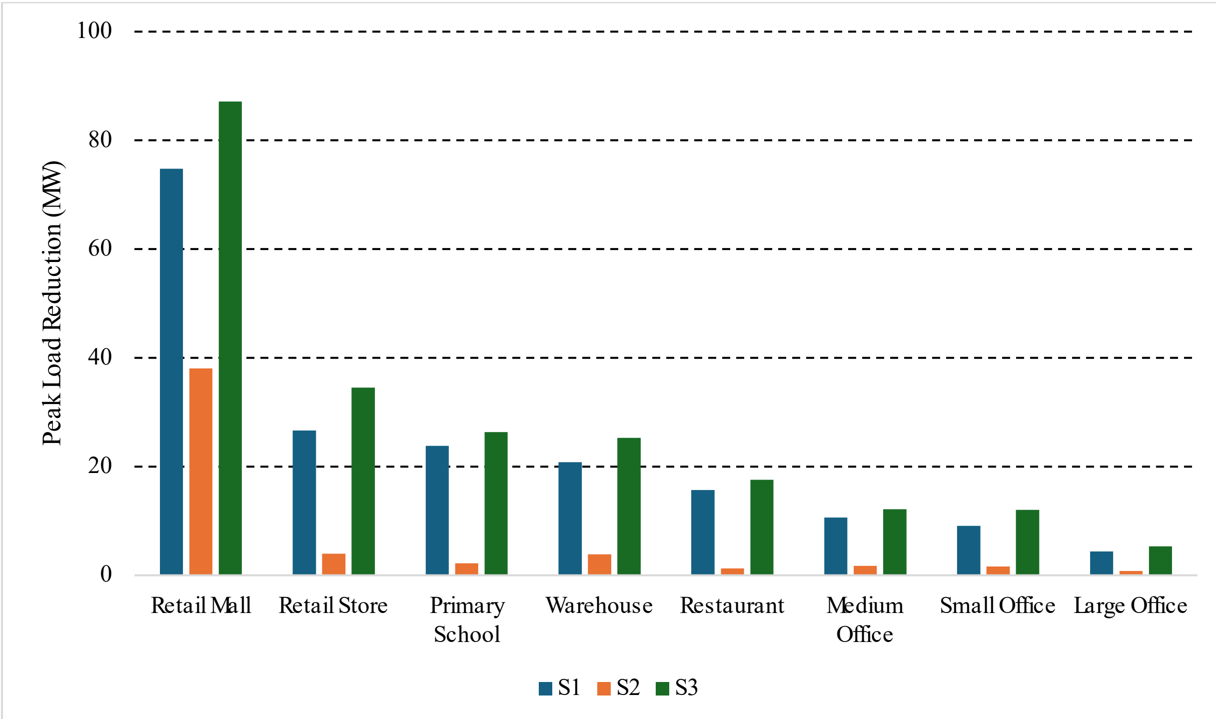

Figure 4: Absolute Peak Reduction by Building Type

The contribution of each building type to system wide peak reduction is closely tied to its baseline demand and its responsiveness to DF strategies (Figure 4). Retail Malls demonstrate the highest impact, with a baseline demand of approximately 372 MW and an S1 strategy reduction of 86 MW. Under the more aggressive S3 strategy, this rises to 114 MW, the largest absolute reduction observed, underscoring their role in load shifting potential.

Retail Stores, Primary Schools and Warehouses demonstrate the next highest impacts after Retail Malls , with reductions of 28–37 MW, 24-28 MW and 22–29 MW respectively, due to large HVAC loads and high baseline demand. Primary School buildings, though smaller in aggregate demand (approximately 255 MW), respond well to thermostat-based control strategies and show reductions of 30–34 MW comparable to much larger retail buildings. Full-Service Restaurants, despite having a modest baseline (approximately 142 MW), exhibit strong potential flexibility with 17–21 MW reductions.

Conversely, Small, Medium, and Large Offices exhibit more limited savings, partly due to lower baselines but also reflecting operational constraints and building configurations that limit DF potential. Notably, TES benefits are only feasible for Large Offices in our Openstudio based modeling approach where chilled-water based HVAC systems allow for meaningful load shifting, but these impacts remain localized.

Together, the results show that buildings with both high electricity demand and responsiveness to thermostat and storage interventions, particularly large-format retail and warehouse spaces, hold the greatest potential to shape aggregate load profiles. The effectiveness of flexibility strategies varies by building type, scenario, and operational characteristic, emphasizing the need for tailored approaches when designing state-level or utility-scale demand-side programs. These insights offer a foundation for incorporating targeted commercial DF into broader grid planning and policy strategies in regions with similar building stock and load profiles.Building automation systems (BAS) and smart thermostats play a critical role in enabling automated demand response. These technologies are commonly deployed in larger commercial facilities such as large and medium offices (greater than 80% penetration) and schools (76% penetration). In contrast, only 18% of small office buildings have BAS systems therefore cannot participate in automated demand response so, price signals such as time-of-use (TOU) rates, critical peak pricing, or real-time pricing can incentivize building operators to shift or curtail loads during high-cost periods without automation. Additionally, direct financial incentives, such as rebates for installing smart controls or participation payments for demand response events, can help overcome upfront cost barriers and encourage broader adoption of demand flexibility measures among small and mid-sized commercial customers.

Our results are broadly consistent with academic and industry state-level assessments in literature. At the state level, California's Demand Response Potential Study (LBNL, 2024), New York's Grid Flexibility Potential Study (Brattle, 2025), and Massachusetts' EEAC Potential Study (Guidehouse 2024) estimate commercial-sector thermostat adjustments could reduce overall system peak demand by approximately 3–6%, depending on building stock composition and enrollment assumptions.

At the individual building level, our findings also align with national-scale assessments. For example, studies show retail spaces can achieve 30–50% individual peak reductions due to their large HVAC and lighting loads (RMI 2020). Other studies support our findings on the TES potential of large office buildings and seasonal DF targeting (Poolla et al. 2020, Afroz et al. 2023, Fu et al. 2021).

POLICY AND PROGRAM DESIGN CONSIDERATIONS FOR COMMERCIAL DEMAND FLEXIBILITY

The results of this study demonstrate the value of commercial building DF in reducing peak load and supporting grid reliability in the modeled jurisdiction. Drawing from established strategies in commercial DR programs, we highlight several common approaches and describe how our results inform their potential effectiveness:

- **Tailor program design to building automation readiness:** Our results indicate, larger commercial buildings such as offices, schools, and warehouses exhibit higher penetration of BAS and advanced energy management controls, making them well-positioned for automated demand response programs. For these segments, policies can leverage automation by offering performance-based incentives, and technical assistance to optimize flexible load participation. On the other hand, smaller commercial buildings often lack automation infrastructure and require behavior-based demand response programs enabled by price signals (e.g., TOU pricing) combined with targeted financial incentives to support adoption of enabling technologies such as smart thermostats and simplified load controllers. This differentiated strategy ensures both scalability and cost-effectiveness across the commercial sector.
- **Standardize flexibility measures through building codes and program requirements**: While our study does not explicitly test the impact of standardized building code requirements, our simulations do demonstrate the consistent effectiveness of thermostat setback strategies across building types. This supports the rationale that embedding basic flexibility requirements, such as smart thermostats or automated controls, into building codes or retrofit standards can mainstream low-cost strategies like thermostat setbacks. These can be especially impactful in large or older commercial buildings with HVAC-dominated loads.

- **Align pricing and incentives with grid conditions and renewables availability:** Our simulation scenarios specifically highlight peak-period reductions between 2–8 PM, aligning with times of declining solar energy output and elevated cooling demand. This reinforces the potential value of dynamic or time-sensitive pricing structures that discourage consumption during critical peak periods and incentivize energy use during times when renewable resources, especially solar and wind, are abundant.
- **Scale participation through streamlined enrollment and targeted program design**: Although our analysis does not directly examine enrollment strategies, the variability in DF potential across building types and vintages identified in our results suggests targeted outreach and simplified enrollment procedures could greatly enhance program participation. Programs can prioritize high impact building types (e.g., retail, warehouse, schools). Using opt-out enrollment strategies, enrollment protections and providing simple automation tools can significantly increase uptake, especially among small and mid-sized businesses.
- **Strategic long-term partnerships and grid planning integration**: Our simulations underscore the feasibility and value of advanced DF strategies such as TES in large office buildings. Policymakers and utilities can collaborate with third-party service providers to deploy advanced technologies, such as AI-driven controls or thermal energy storage, while incorporating DF into long-term resource and capacity planning frameworks.

LIMITATIONS AND FUTURE WORK

This analysis offers valuable insights into the technical potential for DF in the commercial sector at a level of analysis that is appropriate to the provision of a step-by-step guide, or playbook for replication in other jurisdictions. Accordingly, the following limitations are acknowledged. First, the modeling framework assumes a fixed 2–8 PM peak window across all building types. This overlooks the diversity in load profiles. Offices often peak earlier in the afternoon (12–6 PM), whereas retail buildings may peak later in the evening (4–10 PM). A more granular, building-type-specific approach to defining flexibility windows would likely yield a higher, more tailored estimate of DF potential. Second, the results reflect an idealized technical potential, not a market-based outcome, as load shifting is assumed rather than derived from price or comfort-driven decision-making models. This means the findings may not represent the nuances of how DF would play out with the addition of real-world complexity such as the incorporation of factors such as occupant comfort and behavior, economic incentives, and operational constraints can significantly influence participation and outcomes, especially in sectors with variable occupancy or sensitive end uses (Chandra Putra et. al. 2017; Senick et. al. 2018).

Third, our analysis does not explicitly address the temperature-sensitive nature of building thermal loads. The TMY EPW files do not isolate the direct sensitivity of building thermal loads specifically to outdoor temperature variations measured in heating degree days (HDD) and cooling degree days (CDD). This limitation means our study does not disaggregate weather-sensitive loads from internal or schedule-based loads, potentially missing finer-grained insights into how buildings respond specifically to outdoor temperature fluctuations. Future research could employ regression-based load disaggregation techniques to explicitly quantify and separate these weather-sensitive load responses.The study also does not incorporate a high electrification scenario. In such a case, especially relevant for cold Northeastern states, winter peaks could become more prominent due to widespread electrification of heating systems. This would shift the focus of flexibility strategies beyond summer cooling, requiring updated modeling of winter DF and dual-peak system planning. Finally, the ComStock dataset represents approximately 61%

of the national commercial building stock as captured by CBECS. The exact coverage for the modeled state is uncertain, limiting the comprehensiveness of the results. Therefore, the estimates presented here should be interpreted as a lower-bound technical potential rather than a full system-wide DF assessment.

Future work might extend the benefits of this low-cost framework to evaluate baseline and DF scenarios under proposed time-of-use (TOU) rates offered by Electric Distribution Companies (EDCs), helping to align modeled flexibility with real-world tariffs. Further analysis is also needed to estimate the cost of shifting load, considering not just economic factors, but also comfort and emissions impacts. Incorporating high-electrification scenarios would allow for a more realistic appraisal of winter flexibility potential. Finally, integrating health impact assessments under different grid conditions could broaden the value proposition of DF beyond operational savings.

REFERENCES

Addy, N. J., Mathieu, J. L., Kiliccote, S., & Callaway, D. S. (2013). Understanding the effect of baseline modeling implementation choices on analysis of demand response performance. In Proceedings of the ASME 2013 International Mechanical Engineering Congress & Exposition (IMECE 2013), San Diego, CA (Nov 15–21). LBNL-5560E. https://eta-publications.lbl.gov/sites/default/files/lbnl-5560e.pdf

Afroz, Z., M. Goldsworthy, and S. White. 2023. "Energy Flexibility of Commercial Buildings for Demand Response Applications in Australia." *Energy and Buildings* 300: 113533. https://doi.org/10.1016/j.enbuild.2023.112873

Australia Institute and Buildings Alive. 2023. *Buildings as Batteries: How Shifting Electricity Consumption Away from Peak-Demand Times Could Save Billions.* Canberra, ACT: The Australia Institute. https://reneweconomy.com.au/buildings-as-batteries-how-shifting-electricity-consumption-away-from-peak-demand-times-could-save-billions/

California Energy Commission. 2022. *Building Energy Efficiency Standards—Title 24.* Sacramento, CA. https://www.energy.ca.gov/programs-and-topics/programs/building-energy-efficiency-standards

California Legislature. 2018. *Senate Bill No. 100: The 100 Percent Clean Energy Act of 2018.* Sacramento, CA. https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180SB100

Cai, J., and J. E. Braun. 2019. "Assessments of Demand Response Potential in Small Commercial Buildings across the United States." *Science and Technology for the Built Environment* 25 (10): 1437-55. https://doi.org/10.1080/23744731.2019.1629245

Chandra Putra, H., C. J. Andrews, and J. A. Senick. 2017. "An Agent-Based Model of Building Occupant Behavior during Load Shedding." *Building Simulation* 10 (6): 845-59. https://doi.org/10.1007/s12273-017-0384-x

Creutzig, F., J. Roy, W. F. Lamb, I. M. L. Azevedo, W. Bruine de Bruin, H. Dalkmann, O. Y. Edelenbosch, et al. 2018. "Towards Demand-Side Solutions for Mitigating Climate Change." *Nature Climate Change* 8 (4): 268-71. https://doi.org/10.1038/s41558-018-0121-1

Dyson, M., C. Goldenberg, and H. Masters. 2018. *Demand Flexibility: The Key to Enabling a Low-Cost, Low-Carbon Grid* (Insight Brief). Basalt, CO: Rocky Mountain Institute. https://rmi.org/insight/demand-flexibility

Electric Power Research Institute (EPRI). 2017. *U.S. National Electrification Assessment.* Palo Alto, CA: EPRI. https://www.epri.com/research/products/000000003002011471

Energy Information Administration (EIA). 2022. *Form EIA-861 Detailed Data.* Washington, DC: U.S. Department of Energy. https://www.eia.gov/electricity/data/eia861/

Federal Energy Regulatory Commission (FERC). 2009. *A National Assessment of Demand Response Potential.* Washington, DC. https://www.ferc.gov/sites/default/files/2020-05/06-09-demand-response_1.pdf

Fu, Y., Z. O'Neill, J. Wen, and S. T. Bushby. 2021. "Utilizing Commercial Heating, Ventilating, and Air-Conditioning Systems to Provide Grid Services: A Review." *Applied Energy* 297: 117055.

Gagnon, Pieter, Brady Cowiestoll, and Marty Schwarz. 2023. Cambium 2022 Scenario Descriptions and Documentation. Golden, CO: National Renewable Energy Laboratory. NREL/TP-6A40-84916. https://www.nrel.gov/docs/fy23osti/84916.pdf

Gerke, B., J. Kresch, A. Satchwell, M. A. Piette, and A. Stokes-DiMaggio. 2024. *The California Demand Response Potential Study, Phase 4: Report on Shed and Shift Resources through 2050.* Berkeley, CA: Lawrence Berkeley National Laboratory. https://eta-publications.lbl.gov/sites/default/files/phase_4_dr_potential_study_final_2024-05-21.pdf

Guidehouse Inc. 2024. *Massachusetts 2025-2027 Energy Efficiency and Demand Response Potential Study.* Boston, MA: Massachusetts Energy Efficiency Advisory Council. https://ma-eeac.org/wp-content/uploads/2024-02-28-2025-2027-MA-PS-Final-Report-wAppx-Unitil.pdf

Hledik, R., Faruqui, A., Lee, T., & Higham, J. (2019). *The National Potential for Load Flexibility: Value and Market Potential Through 2030.* The Brattle Group. https://www.brattle.com/wp-content/uploads/2021/05/16639_national_potential_for_load_flexibility_-_final.pdf

Hledik, R., A. Ramakrishnan, K. Peters, S. Edelman, and A. S. Brooks. 2025. *New York's Grid Flexibility Potential Study: Volume I—Summary Report.* Albany, NY: New York State Energy Research and Development Authority. https://www.brattle.com/wp-content/uploads/2025/02/New-Yorks-Grid-Flexibility-Potential-Volume-I-Summary-Report.pdf

Kiliccote, S., Piette, M. A., & Hansen, D. (2006). *Advanced Controls and Communications for Demand Response and Energy Efficiency in Commercial Buildings*. Presented at the Second Carnegie Mellon Conference in Electric Power Systems: Monitoring, Sensing, Software and Its Valuation for the Changing Electric Power Industry, Pittsburgh, PA, January 12, 2006. LBNL-59337. Lawrence Berkeley National Laboratory. https://research.ece.cmu.edu/electriconf/2006/kiliccote-piette-hansen%20final.pdf

Langevin, J., C. B. Harris, and J. Cook. 2021. *A National Roadmap for Grid-Interactive Efficient Buildings.* Washington, DC: U.S. Department of Energy. https://connectedcommunities.lbl.gov/sites/default/files/2021-07/1.%20A%20National%20Roadmap%20for%20GEBs-20210712.pdf

Langevin, J., C. B. Harris, A. Satre-Meloy, H. Chandra-Putra, A. Speake, E. Present, R. Adhikari, et al. 2021. "Flexible Building Technologies and Their Role in U.S. Decarbonization." *Joule* 5 (8): 2102-28. https://doi.org/10.1016/j.joule.2021.06.002

Massachusetts Executive Office of Energy and Environmental Affairs. 2022. *Clean Energy and Climate Plan for 2025 and 2030.* Boston, MA. https://www.mass.gov/info-details/massachusetts-clean-energy-and-climate-plan-for-2025-and-2030

Mathieu, J. L., D. S. Callaway, and S. Kiliccote. 2011. "Variability in Automated Responses of Commercial Buildings and Industrial Facilities to Dynamic Electricity Prices." *Energy and Buildings* 43 (12): 3322-30. https://doi.org/10.1016/j.enbuild.2011.08.020

Navigant Consulting. 2017. *U.S. Department of Energy Grid-Interactive Buildings Market Characterization Report.* Washington, DC: U.S. Department of Energy. https://www1.eere.energy.gov/buildings/pdfs/75470.pdf

New York Independent System Operator (NYISO). 2023. *2023 Load & Capacity Data (Gold Book).* Rensselaer, NY. https://www.nyiso.com/documents/20142/2226333/2023-Gold-Book-Public.pdf

New York State Senate. 2019. *Climate Leadership and Community Protection Act, Chapter 106, Laws of 2019.* Albany, NY. https://www.nysenate.gov/legislation/bills/2019/S6599

Piette, M. A., Watson, D., Motegi, N., & Kiliccote, S. (2007). *Automated Critical Peak Pricing Field Tests: 2006 Pilot Program Description and Results* (LBNL-62218). Lawrence Berkeley National Laboratory. https://escholarship.org/content/qt39n0h7x4/qt39n0h7x4.pdf

Poolla, C., A. K. Ishihara, D. Liddell, R. Martin, and S. Rosenberg. 2020. "Occupant Plug-Load Management for Demand Response in Commercial Buildings: Field Experimentation and Statistical Characterization." *arXiv* preprint arXiv:2004.06633.

RMI (Rocky Mountain Institute). (2020, September 24). *Cut Costs, Reduce Carbon, and Improve Health with Demand Flexibility.* https://rmi.org/cut-costs-reduce-carbon-and-improve-health-with-demand-flexibility/

Satchwell, A., M. A. Piette, A. Khandekar, J. Granderson, N. Mims Frick, R. Hledik, A. Faruqui, et al. 2021. *A National Roadmap for Grid-Interactive Efficient Buildings.* Washington, DC: U.S. Department of Energy. https://connectedcommunities.lbl.gov/sites/default/files/2021-07/1.%20A%20National%20Roadmap%20for%20GEBs-20210712.pdf

Senick, J., C. J. Andrews, H. Chandra Putra, I. Tsoulou, and M. A. Sorensen-Allacci. 2018. "Synthetic Populations of Building Office Occupants and Behaviors." In *Building Performance Evaluation*, edited by W. Preiser, A. Hardy, and U. Schramm, 89-104. Cham: Springer. https://doi.org/10.1007/978-3-319-56862-1_5

U.S. BEA (Bureau of Economic Analysis). 2021. *Gross Domestic Product by State.* Washington, DC: U.S. Department of Commerce. https://www.bea.gov/data/gdp/gdp-state

U.S. Department of Energy, Building Technologies Office (DOE BTO). 2019. *Grid-Interactive Efficient Buildings Technical Report.* Washington, DC. https://www.energy.gov/eere/buildings/grid-interactive-efficient-buildings

U.S. Energy Information Administration (EIA). (2025). *Electric Power Monthly* (Table 5.1: Sales of Electricity to Ultimate Customers: Total by End-Use Sector). U.S. Department of Energy. https://www.eia.gov/electricity/monthly/