On Data-Driven Energy Flexibility Quantification: A Framework and Case Study

Han Li, Tianzhen Hong*

Building Technology & Urban Systems Division, Lawrence Berkeley National Laboratory

*Corresponding author: thong@lbl.gov

Abstract

Building energy flexibility is an important resource for a sustainable and resilient power grid, and an important measure to reduce utility costs for building owners. Quantifying energy flexibility for existing buildings can provide critical insights in optimizing their operation. Data-driven methods for building energy modeling and analytics are gaining popularity due to the increasingly available sensor and meter infrastructure, affordable computational resources, and advanced modeling algorithms. However, their application in quantifying the energy flexibility of real buildings is still limited due to the heterogeneous data types and limited data availability. This study proposes a framework for building-level data-driven energy flexibility quantification that considers different levels of data availability and use cases. Two case studies with real building data collected at different scales were conducted to demonstrate the proposed framework for different purposes.

Keywords: Grid-interactive Efficient Buildings, Data-Driven Methods, Energy Flexibility, Demand Flexibility, Building Dataset

1. Introduction

As more countries and jurisdictions commit to aggressive decarbonization goals, grid power systems worldwide are undergoing a rapid transition to renewable sources. According to the International Renewable Energy Agency, more than 80% of newly added electricity generation capacity in 2020 was renewable, and solar and wind accounted for 91% of that capacity [1]. The growing portion of renewable energy generation also means the room for decarbonizing the supply side is reducing. Moreover, it is becoming increasingly challenging to incorporate those highly variable renewable resources into the grid. Because renewable energy generation is not always available at the location and time it is needed, behind-the-meter interventions have become very important for balancing supply and demand. As buildings account for nearly 40% of global energy consumption, and over 80% of peak power demand [2], they are becoming important assets for a sustainable and resilient power grid.

In recent years, many initiatives, including grid-interactive efficient buildings (GEBs), have aimed to remake buildings into clean and flexible energy resources by combining energy efficiency and demand flexibility technologies to enable time- and location-sensitive load shaping [3]. Among the initiatives, building energy flexibility is a trending research topic, and it has attracted research interest worldwide. According to Annex 67 under International Energy Agency's Energy in Buildings and Communities Programme, energy flexibility (interchangeably referred as *demand flexibility*) refers to the "ability of an energy system or a group of interconnected energy systems to adjust and adapt to changes in energy supply and demand, in response to different operational,

economic, and environmental conditions." Research groups have formed to address the topics; these include the International Energy Agency's Energy in Buildings and Communities Programme Annex 67, Annex 81 and Annex 82, and the IEEE's FlexGEB. Researchers have collectively looked into different aspects of building energy flexibility, such as definitions and terminologies [4], technology evaluations [5–7], policy evaluations [8], and field demonstrations [9]. Despite the diverse aspects, they are all inextricably linked with one topic, which is the quantification of energy flexibility in buildings.

Hundreds of studies have looked into quantifying building energy flexibility according to recent literature reviews [10–12]. A large portion of those studies applied to hypothetical buildings to evaluate the effectiveness of different building designs, system types, and control strategies, with various key performance indicators (KPIs). In those studies, simulations were often used to obtain load profiles under reference and flexible operation scenarios. For example, Stinner et al. [13] analyzed the influence of heat generator and thermal energy storage size on building energy flexibility in terms of time, power, and energy using a simulation environment. Arteconi et al. [14] proposed a method to label building energy flexibility at the design phase and demonstrated the method. Tang et al. [15] proposed a suite of energy flexibility indices and conducted a simulation study to quantify the impacts of different energy system designs and control parameters on those indices. Ruusu et al. [16] introduced a model predictive control (MPC) method to improve energy flexibility for a residential building with multiple energy systems and compared its performance with a baseline rule-based control method.

Although those studies shed light on how energy flexibility can be quantified using various KPIs, their methodologies can hardly be applied to real operational buildings. The main reasons are twofold. First, it is impossible to get both baseline and flexible load profiles simultaneously without a controlled experiment, which is unrealistic for existing buildings. Second, physics-based modeling of existing buildings requires detailed building characteristics and calibration if simulations are to be used to generate baseline load profiles. Given the aforementioned challenges, data-driven quantification methods are gaining popularity. As sensor and meter data became more available, more studies have started to investigate novel KPIs and modeling methods that can be applied to real measurements from existing buildings.

Depending on whether a reference scenario is needed, energy flexibility KPIs can be classified in one of two ways: as baseline-required or baseline-free [17]. Therefore, the main challenges for data-driven energy flexibility quantification include (1) selecting or developing baseline-free KPIs that can be directly used with the actual operation scenario, and (2) developing data-driven models that can generate baseline data. Research concerning those two aspects are found in recent publications. Wang et al. [18] developed a data-driven approach to explore the energy flexibility potential of building clusters, where reduced order models were developed to simulate different operation scenarios. Qi et al. [19] developed an unsupervised load decomposition method to quantify baseline energy consumption from smart meter data, which is used to quantify the demand response potential of residential buildings. Pinto et al. [20] proposed a district scale energy management using deep reinforcement learning and quantified the energy flexibility with both baseline-needed and baseline-free KPIs. Kathirgamanathan et al. [21] reviewed 115 data-driven predictive control studies for unlocking building energy flexibility. The review systematically

summarized applications, building types, data-driven models, control algorithms, optimization objectives, and other factors. A key finding of that review is there is a lack of a clear pipeline for a data-driven model and control development. While much of the literature addresses data-driven energy flexibility quantification, most are case-specific, and there is still a lack of generic guidelines for data-driven energy flexibility quantifications. This motivated us to develop a generic framework for building-level energy flexibility quantification that can adapt to different use cases and data availability scenarios.

This paper is organized as follows: Section 2 introduces the four-step data-driven building energy flexibility quantification framework. Section 3 presents two case studies with real measurements from different buildings with different data availability and quality. Section 4 discusses the main technical contributions of the proposed framework and future opportunities. Section 5 presents conclusions.

2. A Framework for Data-Driven Energy Flexibility Quantification

The proposed framework consists of seven steps, which are shown in Figure 1. Each step has different scenarios that are associated with different levels of difficulty (i.e., requirements for expertise, efforts and resources), and these are color-coded by green (easy), yellow (moderate), and red (difficult). This section will introduce the considerations and provide guidance for each step.

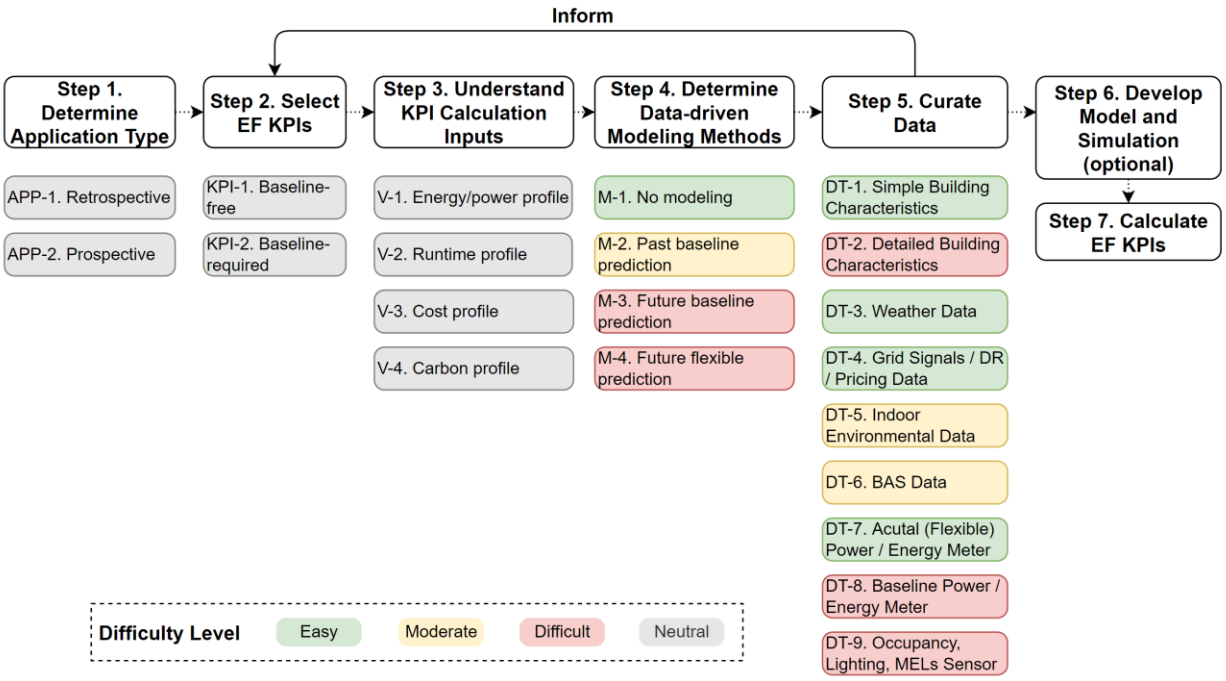


Figure 1. Data-Driven Energy Flexibility Quantification Framework

2.1. Determine the application type

As introduced in Section 1, many existing studies used detailed energy models to evaluate building energy flexibility with different design options and operation strategies. Those studies are cross-sectional because researchers can assess how different variable combinations affect a building's energy flexibility at the same time before the building even exists. On the other hand, data-driven energy flexibility quantification with existing buildings in their lifespan requires a longitudinal study. Depending on the purpose of energy flexibility quantification, there are two types of applications—retrospective and prospective. Their definitions, examples, and stakeholders are summarized in Table 1.

Table 1. Definitions and examples of retrospective and prospective applications

Application Type	Definition	Examples	Typical Stakeholders
Retrospective (past)	Historical data collected from existing buildings is used to evaluate energy flexibility in the past.	• Tracking a building's energy flexibility performance using historical data. • Benchmarking a building's energy flexibility against others.	• Building owner • Building operator • Building auditor
Prospective (future)	Historical data collected from existing buildings and future covariates are used to predict energy flexibility in the future.	• Predict a building's energy flexibility given certain interventions under future conditions (e.g., weather, grid signal). • Guide building control strategies to maximize energy flexibility.	• Building operator • Control engineer • Grid operator • Load aggregator

When dealing with an energy flexibility quantification problem, the first step is to determine whether it is retrospective (evaluating the past), or prospective (predicting the future). Retrospective studies are usually easier to do as they aim to evaluate the energy flexibility as-is and therefore do not rely on forecasting and optimization for future energy flexibility potential estimation.

2.2. Select Energy Flexibility KPIs

Once the application type is determined, different stakeholders should select KPIs that best suit their purposes. Recent literature has summarized building energy flexibility KPIs and demonstrated their usage [12,15,22]. Note that there are different energy flexibility modes, such as load shedding, load shifting, modulation, and onsite generation. Because they provide different grid services to different stakeholders, the corresponding KPIs have different temporal resolutions and evaluation durations. For example, the peak power reduction KPI is usually used by building operators to assess load shedding flexibility, which requires a 15-minute to hourly resolution and concerns only the peak hours within a day. Conversely, the fast regulation KPI [15] is usually used by grid operators to evaluate the load modulation performance, which requires a second-level data resolution and needs to be calculated over 24-hour cycles. In addition, the quantity of interest of the flexibility KPIs can vary, depending on the performance goal. While most KPIs focus on flexibility of electric power demand and energy consumption, other KPIs can also evaluate energy

flexibility with respect to carbon emissions [23,24], cost [25,26], and occupant comfort [27,28]. With the above-mentioned diversity, one should carefully consider flexibility modes and performance goals when choosing KPIs. In practice, KPI selection is also usually an iterative process that needs to consider data availability. Common data-driven KPIs along with KPI selection method can be found in [17].

2.3. Understand KPI Calculation Inputs

Depending on the selected KPIs, different input variables are needed for the calculation. The input variables can be grouped into five categories:

- Energy/power profile: Data related to historical and real-time energy consumption or power demand, energy or demand forecasts, and renewable energy generation.
- Runtime profile: Information about building energy system operating hours, equipment schedules, and on/off patterns of appliances.
- Cost profile: Data on electricity tariff rates, time-of-use pricing, peak demand charges, penalty rates, and the cost of energy storage or backup systems.
- Carbon profile: Information on emission factors, carbon intensity of the electricity grid, renewable energy certificates, carbon pricing, and carbon reduction targets or regulations.

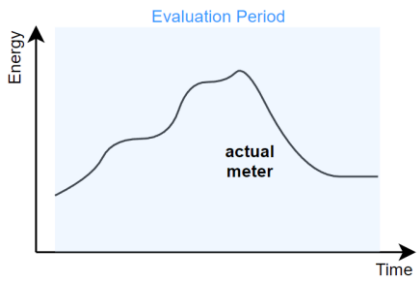
In practice, the above input variables may be derived in different ways, where some variables are available from measurement, while others are only predicted using simulations. Actual projects are constrained by the time and budget. Understanding the data requirement to derive the input variables can help reduce the amount of data curation and modeling effort in the next steps.

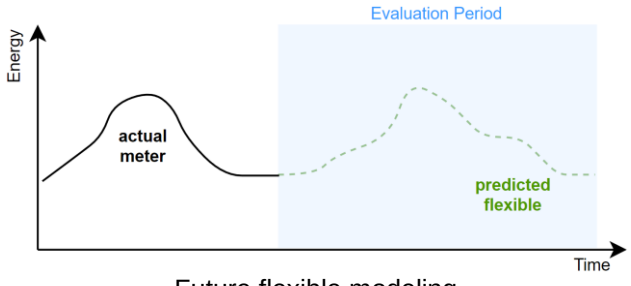
2.4. Determine Data-driven Modeling Methods

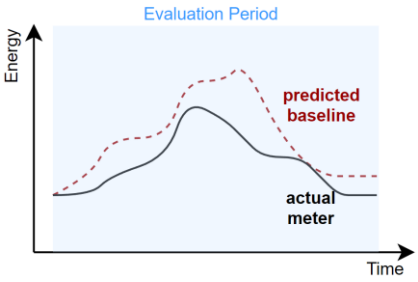
A wide variety of data-driven building energy modeling methods can be found in the existing literature, including regression models [29,30], thermal resistance-capacitance models [31], and machine learning models [32]. There is usually a trade-off between model fidelity and data requirements. For instance, detailed building and system characteristics are difficult to obtain, but can inform parameter identification for reduced-order models [33] and physics-informed neural networks [34]. A few review articles [35,36] have systematically summarized different types of data-driven models with different data requirements and their strengths and weaknesses. However, one should also consider the application type and KPIs when choosing a data-driven algorithm for energy flexibility quantification. For example, a retrospective study with baseline-free KPIs only require historical data, which needs no data-driven modeling at all; whereas, a prospective study with baseline-required KPIs needs both the baseline and flexible load profiles, which mandates baseline and flexible modeling. Table 2 shows the requirements of data-driven models for different application types and KPIs.

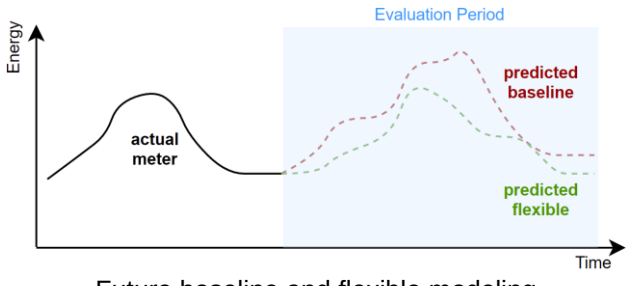
Table 2. Data-driven modeling requirements for energy flexibility quantification

	Retrospective	**Prospective**

<table>
<tr>
<td>Baseline-free KPI</td>
<td> 

No modeling is needed</td>
<td> 

Future flexible modeling</td>
</tr>
<tr>
<td>Baseline-required KPI</td>
<td> 

Past baseline modeling</td>
<td> 

Future baseline and flexible modeling</td>
</tr>
</table>

2.5. Curate Data

Once the application type and data-driven modeling method are determined, corresponding data need to be curated for subsequent modeling and simulation. Typical data types for building energy flexibility quantification with their definitions and accessibility are listed in Table 2. However, not all data types are necessary in each application. Theoretically, more KPIs can be calculated if more data is available and more advanced modeling techniques are applied. However, real studies are constrained by the budget and the investigator's expertise in data curation and modeling development. Therefore, the data curation process can also inform KPI selection in an iterative manner until an appropriate KPI, data, and model set is identified within a given project scope.

Table 2. Data type and availability for energy flexibility quantification

Data Type	Definition	Accessibility
Simple building characteristics	Building type, location, age, size, etc.	easy
Detailed building characteristics	Building geometry, envelope thermal properties, HVAC system types, etc.	difficult
Weather data	Historical and/or future outdoor weather conditions, including temperature, humidity, solar irradiance, precipitation, etc.	easy
Grid signal	Historical and/or future electricity pricing and/or CO_2 emission signals from the grid operator	easy
Indoor environmental data	Historical sensor readings of indoor environmental conditions, such as temperature, humidity, CO_2	moderate

	concentration, etc.	
Building automation system (BAS) data	Monitored data from building automation systems, including HVAC system supply temperature and flow rate, fan and compressor runtime, hot water status, photovoltaic (PV) generation, electric vehicle (EV) charging, etc.	moderate
Actual power / energy meter data	Readings of whole-building or system-level energy meters	easy
Baseline power / energy meter data	Readings of whole-building or system energy meters in the baseline scenario (applicable to controlled experiments only)	difficult
Occupancy, lighting, miscellaneous electrical loads (MELs) sensor data	Information about occupancy, lighting, and MELs operation schedules	difficult

2.6. Develop Model and Simulation

With the selected data-driven modeling method and curated data, models can now be developed to predict the counterfactual load profiles, which serve as the input variables for KPI calculations. It should be noted that modeling is not necessarily required for all cases. For example, baseline-free retrospective studies (top-left cell in **Error! Reference source not found.**) only require actual historical measurements for KPI calculations. Depending on the specific data-driven method determined in Step 4, data engineering and feature selection [39] might be required for the model development. The developed models should also be properly verified to ensure their prediction accuracy in the simulations [36].

2.7. Calculate KPIs

With the input variables ready, the final step is to calculate the KPIs. In most existing studies, the KPI calculation is carried out manually in a case-by-case manner. Recently, researchers have investigated methods to facilitate the energy flexibility KPI selection and calculation with semantic ontologies [37] and open-source software [38]. Those efforts aim to standardize and streamline the quantification process in the near future.

In reality, energy flexibility quantification is constrained by many factors throughout the project, such as the application purpose, data availability, domain expertise required for modeling and analytics, time and budget. The seven-step framework introduced above aims to provide a generic framework that can guide stakeholders when they are facing with different purposes, data availability, budget, expertise, etc. Recognizing the interplay between these factors enables stakeholders to prioritize resources, make informed choices of the modeling and analytical methods, and achieve accurate results.

3. Case Studies

In this section, we demonstrate the proposed framework with two case studies using real building operational datasets. The main goal is to show how the proposed framework could be used for different applications with different levels of data availability and quality.

3.1. Retrospective study with EcoBee DYD data

The first case study uses a large-scale smart thermostat dataset to quantify the energy flexibility of residential buildings during past demand response (DR) events, following the workflow shown in Figure 2.

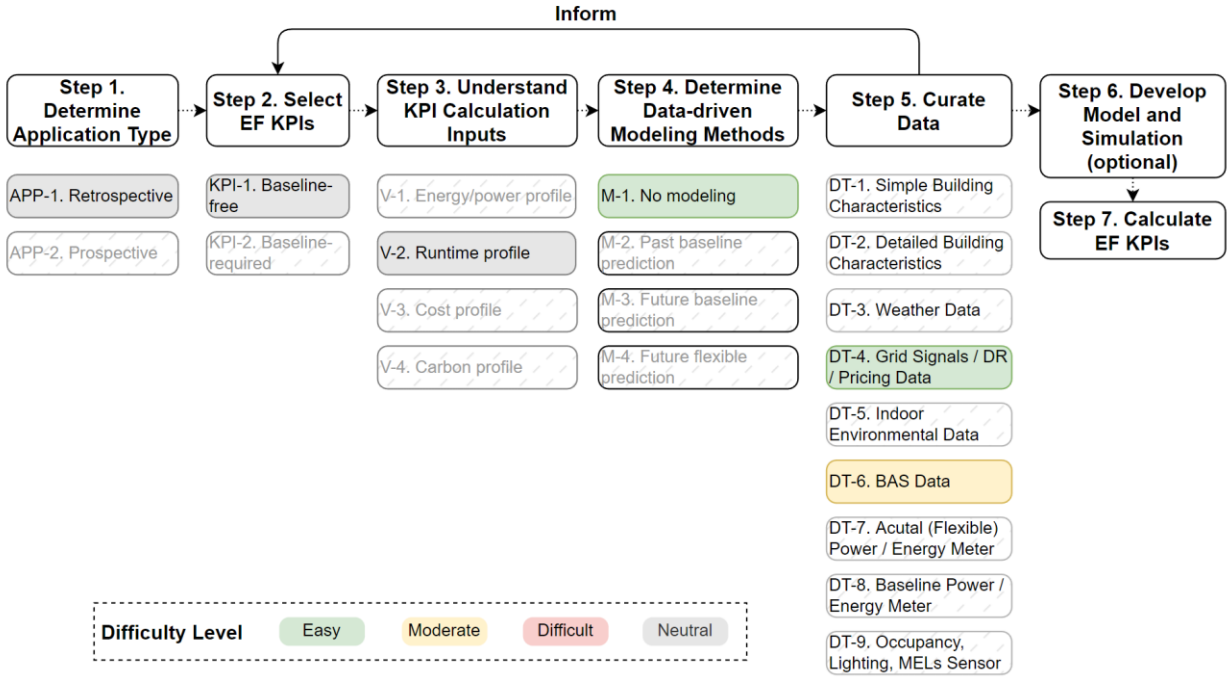


Figure 2. Data-driven energy flexibility (EF) quantification workflow for case study 1

3.1.1. Determine Application type

Quantifying and benchmarking existing buildings' historical energy flexibility can help building owners and operators understand and improve the energy flexibility in the future. The goal of this case study was to calculate and benchmark the historical energy flexibility of residential HVAC systems during DR events. Therefore, it was a retrospective study.

3.1.2. Select KPIs

The dataset was collected by EcoBee's Donate Your Data (DYD) program [40], where more than 190,000 households in the U.S. and Canada voluntarily shared data anonymously for research purposes. In this case study, we used a subset that includes a total of 3,556 households in three U.S. states (i.e., Texas: 1,469, California: 1,530, and New York: 557), which cover various

different climate zones. Given the large sample size, it is challenging to develop a data-driven model for counterfactual baseline predictions for each building. Therefore, we chose to use a baseline-free KPI named flexibility factor (FF) to assess to what extent the HVAC system can shift its operation from DR periods to non-DR periods. The FF is a KPI first introduced by Le Dréau et al. in 2016 [42] to quantify heating system energy flexibility. But its concept can be extended to other quantities of interest, including HVAC system runtime, as shown in Equation 1.

$$FF = \frac{Usage_{non\ DR} - Usage_{DR}}{Usage_{DR} + Usage_{non\ DR}} \quad \text{Equation 1}$$

3.1.3. Understand KPI Inputs

As implied by Equation 1, the FF KPI takes a generic usage profile and the DR window as the input variables. When applied to the HVAC system, its value ranges from negative one, where the system only runs during DR windows, to positive one, where the system runs entirely outside of DR periods. The usage profile variable is usually energy consumption or thermal load of the HVAC system. However, since the DYD dataset does not include energy meter data, we used system runtime as a proxy, which is a common practice recommended by the U.S. Environmental Protection Agency [41].

The flexibility factor value is influenced by multiple factors such as duration of DR events, building thermal mass, air tightness, HVAC system capacity, and operation schedules. For example, a building with good thermal mass and air tightness can better shift its HVAC system runtime outside of summer peak hours by pre-cooling. With the large-scale thermostat data, we can evaluate the distributions of flexibility factors of buildings with different types and locations.

3.1.4. Determine Data-driven Modeling Methods

As stated in Step 2, the FF does not require data-driven modeling.

3.1.5. Curate Data

The DYD dataset contains user-reported metadata about building location, space type, gross floor area, number of floors, thermostat model, and time when the thermostat was first connected. In addition to the metadata, the timeseries data include user-defined climate modes (e.g., sleep, home, away), calendar events (e.g., demand response start and end timestamps), indoor air temperature and humidity, and HVAC system (i.e., fan, compressor, heater) runtime. This case study concerns with the residential buildings' energy flexibility during past DR events. The most common strategy for smart thermostat DR participants is the so-called direct load control (DLC), where users voluntarily give control of their thermostats to utilities or other third parties. During peak hours, grid operators can override the thermostats by raising the temperature setpoint in the cooling season or lowering the temperature setpoint in the heating seasons. To evaluate the energy flexibility of residential HVAC systems during DR events, we first extracted a subset of the entire DYD dataset where the calendar events contain "DR" related keywords.

To provide some general understanding about the curated data, we first looked into the temperature and DR characteristics. Figure 3 shows the temperature distribution comparison

between DR and non-DR periods by state. On average, homes in the heating season lowered their temperature setpoint for 1°C, 1.8°C, and 1.6°C in Texas, California, and New York, respectively. In the cooling season, the cooling setpoint temperatures were raised by 1.3°C, 1.4°C, and 1.7°C in Texas, California, and New York, respectively.

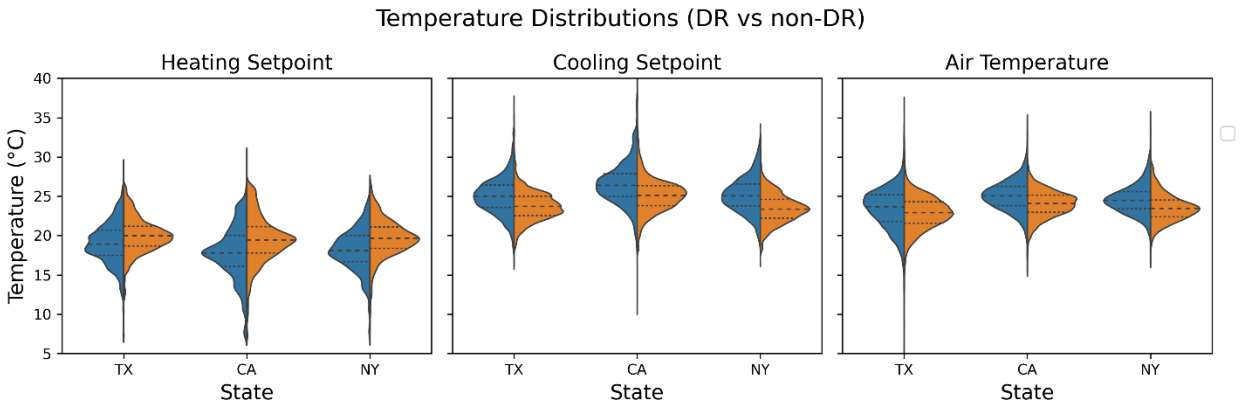


Figure 3. Temperature distribution comparisons between DR and non-DR periods

Figure 4 shows the distribution of DR events' durations. In all three states, most DR events ranged from 10 minutes to about 400 minutes. The average durations for Texas, California, and New York homes were 152 minutes, 182 minutes, and 230 minutes, respectively.

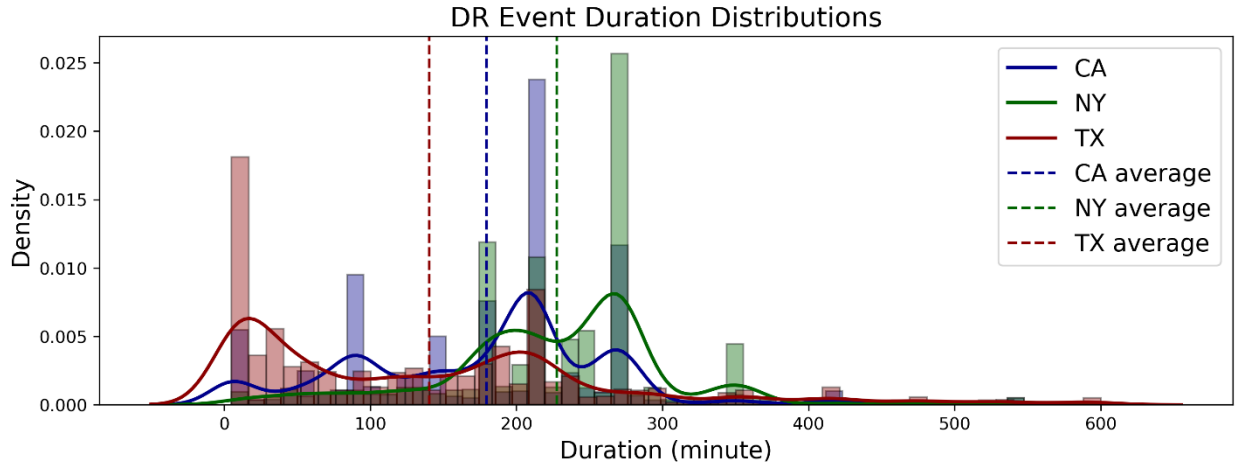


Figure 4. DR events duration comparison by state: the bars show binned durations, the curves indicate the data density, and the dash lines indicate the sample mean.

3.1.6. Develop Model and Simulation

This step is skipped because the case study does not require data-driven model and simulation to get historical baseline runtime profiles.

3.1.7. Calculate KPIs

The curated data was used to calculate the FF KPI for all homes in the sample. Specifically, an FF is calculated for each home at a daily basis. Figure 5 shows the histograms of the daily flexibility factor of fans and compressors during the summer of 2019 and 2020, where the horizontal axis ranges between -1 to 1 and the vertical axis shows the count of each interval. Each row is a U.S. state, each column is a building type, and the color distinguishes fan and compressor. All histograms have long left tails, meaning most buildings on most days have a flexibility factor towards the positive side. Also, the rightmost bins (where the KPI equals 1) have the highest count in all the histograms, which means on those days the buildings successfully shifted HVAC system operation outside of DR periods. There is no significant difference between fans and compressors because their operations are highly overlapped.

In terms of building types, apartments and townhouses tend to have higher average flexibility factors (shorter tails) than those of detached single-family buildings. The possible reason is multiplex buildings tend to have fewer exterior walls (less infiltration and heat loss) than detached buildings. As for location, houses in Texas showed higher average flexibility factors than those in California and New York. A possible explanation is that grid operators in California and New York issued longer DR events than those in Texas in summer 2019 and 2020 (shown in Figure 3), which made it more difficult to shift the HVAC system operation in the buildings.

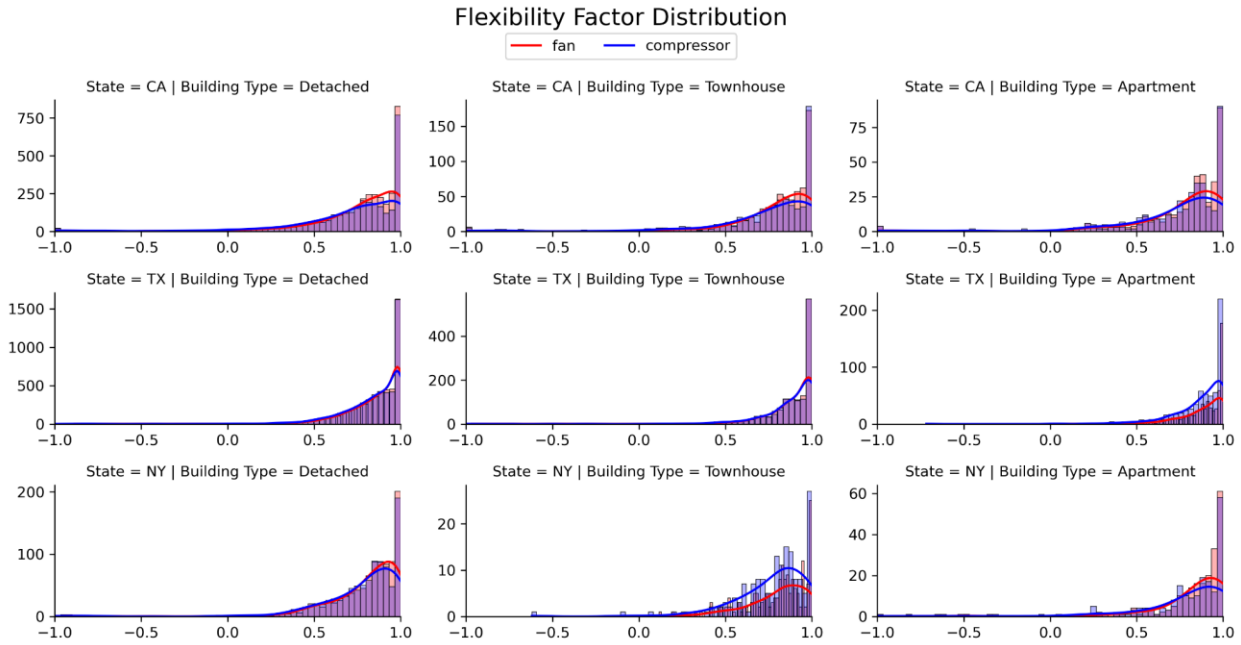


Figure 5. Distributions of flexibility factors by state and building type (x-axis: FF, y-axis, number of buildings)

In addition to the high-level comparisons, we benchmarked individual homes' flexibility factors. Figure 6 shows the flexibility factor benchmarking of two randomly selected houses against the whole sample, where each home has at least 15 DR days. The dashed lines represent the average daily flexibility factor, while the shaded areas show their ranges. It can be seen that

although home A has a higher maximum flexibility factor, it has a wider range and lower average, whereas home B has a consistently high flexibility factor. The benchmarking results could help aggregators and grid operators to better target customers and plan for future DR programs.

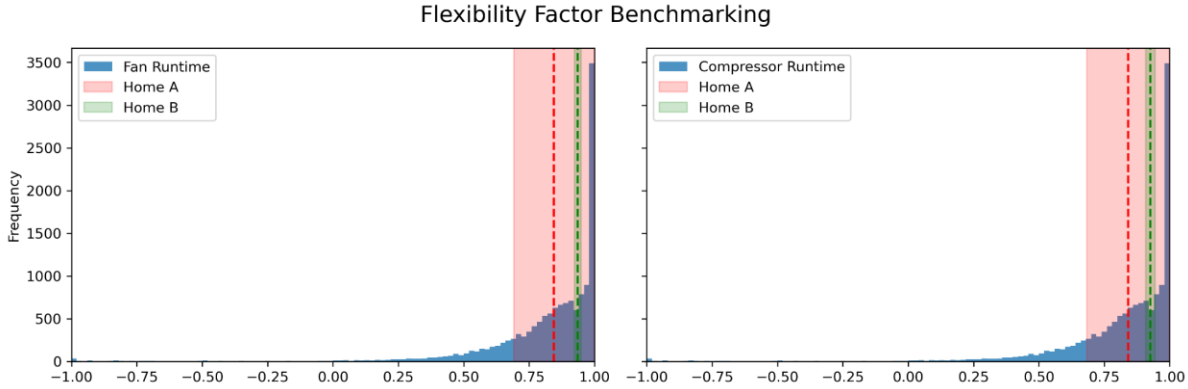


Figure 6. Building-level flexibility factor benchmarking: the blue bars shows the FF for all homes, the red and green bands and dash lines represent the ranges and averages of the FF in two example homes

3.2. Prospective study with PNNL Lab Home data

The second case study demonstrates a prospective study where a data-driven model and a virtual model-predictive controller were developed to improve the energy flexibility of a residential building. Figure 7 shows the implementation of the proposed workflow in the case study.

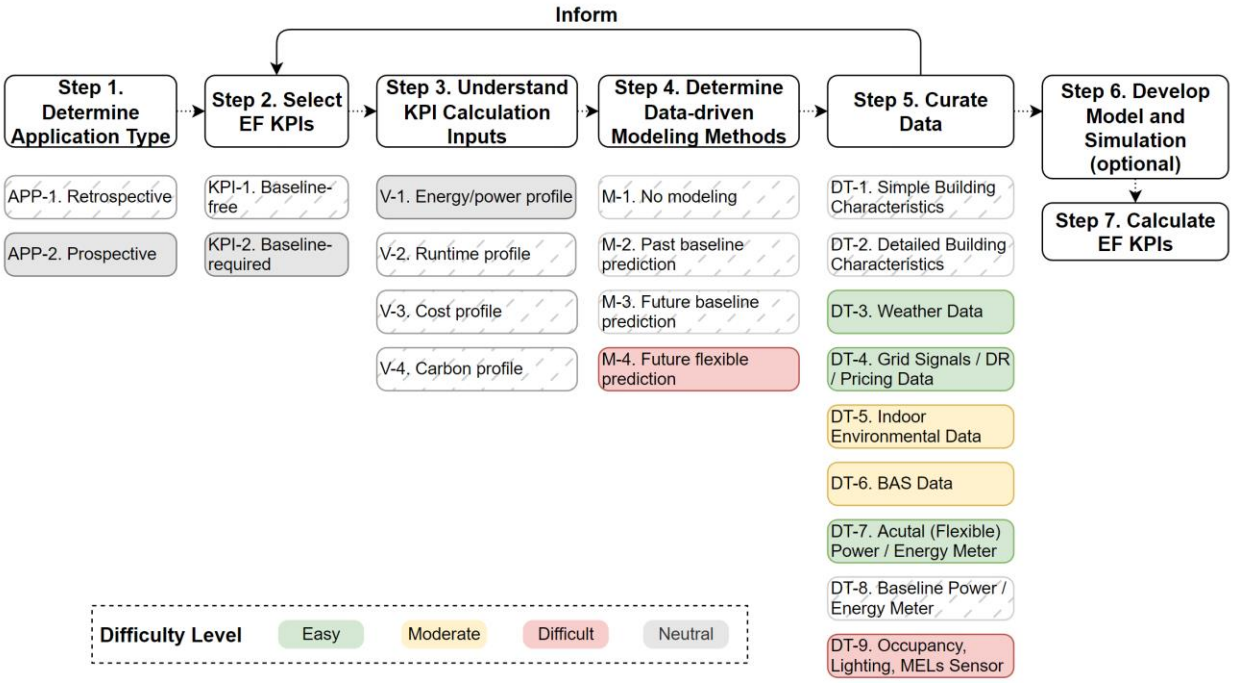


Figure 7. Data-driven EF quantification workflow for case study 2

3.2.1. Determine Application Type

Building control is essential to achieve good energy performance for existing buildings. Understanding how different control strategies would behave under certain future circumstances (e.g., weather conditions, pricing signals, temperature setpoints) is important for building operators. This case study aims to predict an experimental building's energy flexibility under an MPC strategy, comparing with two baseline scenarios. Therefore, it can be considered as a prospective study.

3.2.2. Select KPIs

To demonstrate the quantification of load shedding, three baseline-required KPIs are selected; their formula and terms are listed in Table 3 below. The peak power shedding (PPS) and peak energy shedding (PES) KPIs are both used to quantify how well a building is able to manage the energy consumption during peak periods. However, the PPS only concerns the timestamp when maximum demand occurs, while the PES covers the whole peak period. The former KPI is often used to inform power generation and supply sizing and rotating power outage planning, while the latter can help quantify how much energy and associated emissions are curtailed during peak hours. The building energy flexibility index (BEFI) is essentially the peak energy shedding normalized by the peak period duration, which indicates the average power demand shedding during the peak period.

Table 3. Three energy flexibility KPIs used in case study 2

KPI	**Formula**	**Terms**
Peak Power Shedding (PPS)	$\Delta P = P_{ref,peak} - P_{flex,peak}$	$P_{ref,peak}$ is the peak power demand of the baseline scenario; $P_{flex,peak}$ is the peak power demand of the flexible scenario
Peak Energy Shedding (PES)	$\Delta E = \int_{t_{start}}^{t_{start}+\Delta t} P_{ref}(t)dt - \int_{t_{start}}^{t_{start}+\Delta t} P_{flex}(t)dt$	$P_{ref}(t)$ is the power demand at time t of the baseline scenario; $P_{ref}(t)$ is the power demand at time t of the baseline scenario; t_{start} is the start timestamp of the peak period; Δt is the peak period duration
Building Energy Flexibility Index (BEFI)	$BEFI(t,\Delta t) = \frac{\int_{t_{start}}^{t_{start}+\Delta t} P_{ref}(t)dt - \int_{t_{start}}^{t_{start}+\Delta t} P_{flex}(t)dt}{\Delta t}$	

3.2.3. Understand KPI Inputs

From Table 3, it can be seen that the three selected KPIs all require a baseline and a flexible electric load profile as the inputs. In addition, they need the start time and the duration of the high load periods.

3.2.4. Determine Data-driven Modeling Methods

As illustrated in Figure 7, this case study requires prediction of future flexible operation scenario under the MPC strategy (i.e., counterfactual load profile if the building adopted the MPC strategy). While there are a wide variety of data-driven models for building dynamics, resistance-

capacitance (RC) models are a family of models that use networks analogous to the electric resistors and capacitors to simulate the thermal behavior of buildings. An RC model is physics-informed because its parameters can be identified from real measurements. In the meantime, its simple form allows fast simulation, which suits online optimal control such as MPC. In theory, higher dimension models with more Rs and Cs can better depict the target building's thermal dynamics, while they may also require more data and computational resources. In this case study, we compared a 1R1C model and a 2R2C model and chose the 1R1C structure to model the target residential building because it can predict the building's thermal trend at a high accuracy with fair simplicity[1]. The heat balance of the building and the 1R1C model details can be found in Appendix A1.

3.2.5. Curate Data

The data used in the study was collected from a test facility named Lab Homes in the Pacific Northwest National Laboratory (PNNL) main campus in Richland, Washington [43]. Lab Homes consists of two identical residential buildings; one serves as the baseline (Home A) and the other serves as the experimental (Home B). The two test homes are both equipped with zone-level indoor environmental sensors (e.g., temperature, humidity, occupant) and electric submetering (e.g., lighting, plug-loads, appliances, HVAC), and share a weather station that measures outdoor environmental conditions (e.g., temperature, humidity, solar irradiance), all measured at one-minute intervals. The internal heat gains from equipment and occupants are emulated with electric thermal resistors. Detailed descriptions about the building and system characteristics and the dataset are publicly available via the U.S. Department of Energy funded Benchmark Buildings Datasets project [43]. During the 2021 winter, a control experiment was conducted to test the pre-heating strategy. Its five phases are shown in Table 3. The temperature settings during the experiments can be found in Figure 8.

Table 3. Pre-heating experiment of the PNNL Lab Homes

Phases	Dates	Specifications
Calibration	2021-12-6 to 2021-12-8	Home A and Home B had the same operation schedule and temperature setpoints. Sensor and meter data from both homes were compared to make sure they are about the same.
Setpoint excitation	2021-12-9 to 2021-12-13	Home A kept a constant temperature setpoint while Home B adjusted its temperature setpoint frequently. The operation of other systems except for HVAC were kept the same.
Pre-heating	2021-12-14 to 2021-12-20	Home A kept a constant temperature setpoint while Home B performed a series of pre-heating measures for load shifting. The operation of other systems except for HVAC were kept the same.
Free-floating	2021-12-21 to 2021-12-24	Home A kept a constant temperature setpoint, while Home B was in free-floating mode. The operation of other systems except for HVAC were kept the same.
Warm-up	2021-12-25 to	Home A and Home B had the same operation schedule and

[1] A comparison between 1R1C and 2R2C models can be found in Appendix A1

	2021-12-27	temperature setpoints to wrap up the experiment.

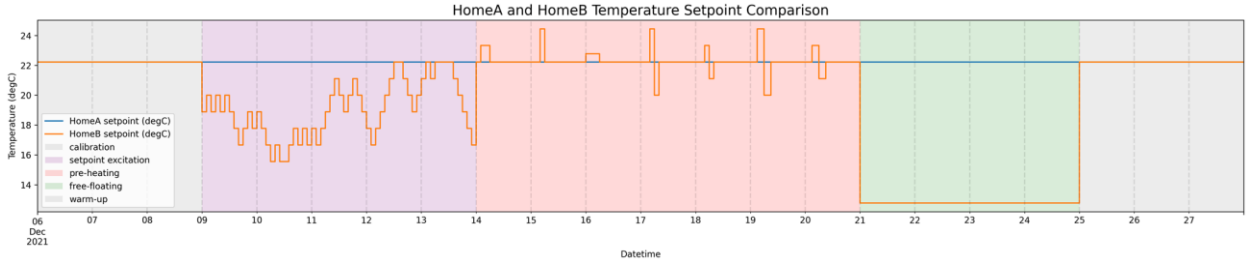


Figure 8. Operation schedule and temperature setpoint comparison of case study 2

In addition to the building characteristics and sensor and meter data, the time-of-use rate of the electricity consumption can be accessed online [44]. Specifically, the on-peak (6 am to 8 am) price is 12 cents/kilowatt-hour (kWh) and off-peak price is 7 cents/kWh during the heating season from October to May.

3.2.6. Develop Model and Simulation

With the model structure stated in Section 3.2.4 and the data curated in Section 3.2.5, we can identify the model parameters and establish the simulation workflow. The details about the 1R1C model and its testing accuracy can be found in the Appendix. This sub section focuses on the simulation setup where the virtual MPC is configured to simulate the prospective flexible operation (i.e., what would happen if the MPC strategy were implemented) and generate input data for KPI calculations.

Previous studies have proven that MPC can notably reduce building energy use and greenhouse gas emissions [45]. For HVAC controls, MPC can be applied at both high levels, such as thermostat setpoint regulation [46], and lower levels, like air handlers [47] and heat pumps [48]. In this case study, the MPC was set up to control the power input of the electric heat pump. Figure 9 shows the schematic of the MPC setup. Here, Ref represents the known variables, $u_k$are the control actions, y are the system outputs, and x are the internal states. The data-driven model identified previously also serves as the virtual system to simulate the building dynamics. The MPC simulation was implemented using Python, where the do-mpc [49] and CaSADi [50] libraries were used to formulate the MPC and perform algorithmic differentiation, respectively.

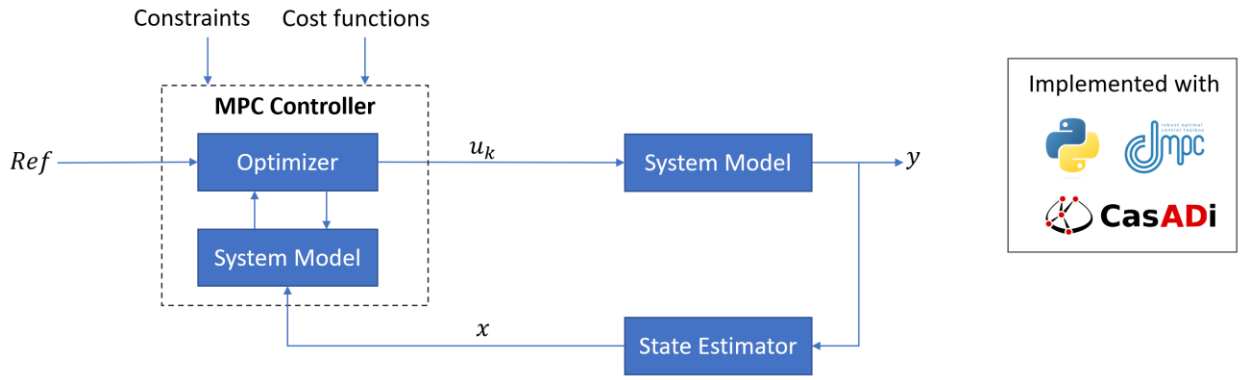


Figure 9. Schematics of the MPC configuration

The objective of the MPC is to reduce power demand during peak price periods while maintaining thermal comfort. The optimal heat pump power input sequence starting at timestamp k until $k+n$ can be obtained by solving Equation 2:

$$min\left(\sum_{k}^{k+n} price^{i} \times {P_{heat\ pump}}^{i} + \begin{cases} w_{comfort} \times \sum_{k}^{k+n}\left({T_{lower}}^{i} - {T_{indoor}}^{i}\right)^{2}, & {T_{indoor}}^{i} < {T_{lower}}^{i} \\ w_{comfort} \times \sum_{k}^{k+n}\left({T_{indoor}}^{i} - {T_{upper}}^{i}\right)^{2}, & {T_{indoor}}^{i} > {T_{upper}}^{i} \end{cases}\right) \quad \text{Equation 2}$$

s.t. $15℃ \leq T_{indoor} \leq 30℃$

$0W \leq P_{heat\ pump} \leq 5200W$

where: $price^{i}$ is electricity price at i-th step.

${P_{heat\ pump}}^{i}$ is the heat pump power input at i-th step.

$w_{comfort}$ is the weighting factor of thermal comfort penalty.

${T_{lower}}^{i}$ is the lower comfort temperatures bound at i-th step.

${T_{upper}}^{i}$ is the upper comfort temperatures bound at i-th step.

In the above equation, the lower and upper comfort temperature bounds, which serve as soft constraints, are set to be 21.2°C and 24.2°C, respectively. However, they can be time-variant in situations, depending on the operation schedule. The hard constraints for indoor temperature are set to be between 15°C and 30°C to prevent overcooling and overheating. The thermal comfort penalty factor $w_{comfort}$ is set to be 20,000 because it provides a good balance between energy and comfort. However, note that some studies have more thoroughly investigated thermal comfort weights into the cost function [51]. Future MPC experiments for energy flexibility can consider those integrated cost functions. The heat pump's electric power demand is assumed to be between 0 and 5,200 watts (W), which is inferred from the measurements. The prediction horizon n is 360 minutes, which is long enough to cover the entire high-price period.

The 1R1C heat-balance equation can be discretized into the form of Equation 3 to simulate the indoor air temperature trend given current inputs.

$${T_{indoor}}^{i+1} = A^{T}x^{i} + Bu^{i} \quad \text{Equation 3}$$

Where:

$$A = \begin{bmatrix} 1 - dt/(R*C) \\ dt/(R*C) \\ A_{sol}*dt/C \\ dt/C \end{bmatrix}$$

$$x^i = \begin{bmatrix} {T_{indoor}}^i \\ {T_{outdoor}}^i \\ {I_{sol}}^i \\ {Q_{internal}}^i \end{bmatrix}$$

$$B = dt/C$$

$$u^i = {P_{heat\ pump}}^i * COP_{heat\ pump}$$

In the system matrix A, parameters R, C, and A_{sol} are identified from the data-driven modeling phase, and dt is the timestep (i.e., one minute). The state vector x consists of indoor and outdoor air temperature, horizontal solar irradiance, and internal heat gains at the i-th step, which are known from measurements. The control input is the heat gain from the electric heat pump at the i-th step.

3.2.7. Calculate KPIs

With the configuration described in Section 3.2.6, a simulation was conducted for the duration from 2021-12-17 to 2021-12-20 to generate the counterfactual scenario, where the outputs include indoor temperature trends and heat pump energy consumption. Figure 10 shows the comparison between the three different operation strategies. The top subplot shows the indoor temperature comparison, where the comfort temperature ranges and peak demand periods are also identified by the shaded areas. It can be seen that with a constant temperature setpoint, Home A's indoor temperature had only slight fluctuations around 22°C. In comparison, Home B's indoor temperature followed its rule-based setpoint schedule, which was raised before peak periods and lowered during peak periods. Similar patterns could be identified from the MPC case, where the indoor temperature was raised before the peak periods arrived. In all cases, the indoor temperature was kept within the comfort range for most of the time. Distinct patterns could be found in the middle subplot, which shows the measured power demand for Home A and Home B, and simulated power demand for the MPC. In all cases, there was a dip in the afternoon of 2021-12-17, which could be explained by the increased solar heat gain indicated by the high solar irradiance during that period shown in the bottom subplot. For other times, the heat pump in Home A was always running in the 1,900 W to 3,200 W range, which was because of its constant temperature setpoint. Both Home B and the MPC exhibited pre-heating behaviors, where the heat pump increased power demands before the peak period and reduced demand during the peak period. However, the MPC showed better load shifting performance as it reduced heat pump operation to almost zero during all peak periods, while Home B still had its heat pump running in parts of the peak periods. Another interesting observation can be found in the transition from 2021-12-20 to 2021-12-21. When Home B and the MPC were both set to free-floating mode, the MPC stopped its heat pump operation about one hour before the new schedule kicked in, while Home B only reduced its heat pump operation until the new schedule started. The comparison shows that the MPC was able to optimize the heat pump operation to consider future schedules.

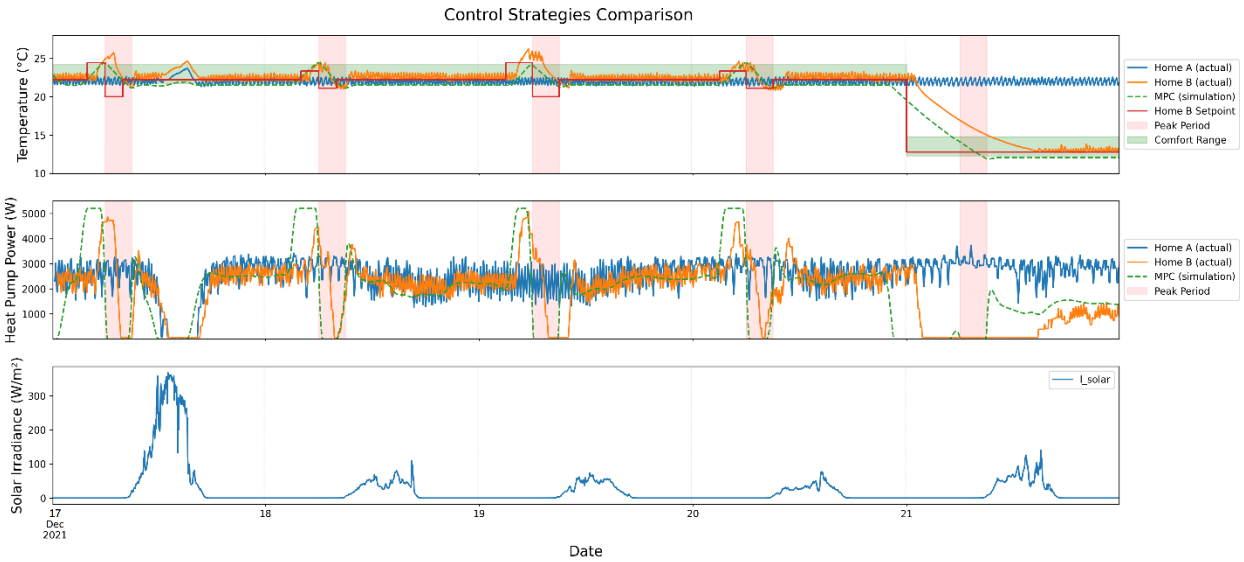


Figure 10. MPC simulation results and the comparison with real measurements. Top: setpoint and indoor temperature in lines, comfort range in green horizontal bands, peak hours in red vertical bands; Middle: heat pump power demand in lines, peak periods in red vertical bands; Bottom: solar irradiance

The daily KPI results of the experiment period are shown in Table 4 below. The KPIs are calculated based on the electrical load of the heat pump system. Depending on the selection of baseline and flexible scenarios, there are three pairs of comparisons: (1) Home B versus Home A, (2) MPC versus Home A, and (3) MPC versus Home B. For PPS and PES, we calculated both the absolute and relative shedding. It is interesting to note that if only looking at PPS, Home B actually performed worse than Home A, which is because the pre-heating of Home B was not strictly limited to non-peak periods (Figure 7). However, both PES and BEFI show that Home B performed better than Home A, because Home B's demand was reduced for most of the peak periods. As for the comparison between the MPC and Home A, all three KPIs indicated that the MPC performed better than Home A, because the MPC was able to almost completely shift the heat pump operation to before and after peak periods. The comparison between the MPC and Home B shows that the MPC can further optimize load shedding performance beyond the rule-based control in Home B, which echoes the visual findings from Figure 7.

Table 4. Energy flexibility KPI results in case study 2

KPI	Date	Home B (Flexible) vs. Home A (Baseline)	MPC (Flexible) vs. Home A (Baseline)	MPC (Flexible) vs. Home B (Baseline)
Peak Power Shedding (PPS)	12/17/2021	(-0.12 kW, -6.02%)	(1.2 kW, 58.56%)	(1.32 kW, 60.91%)
	12/18/2021	(-0.1 kW, -4.95%)	(0.82 kW, 40.62%)	(0.92 kW, 43.43%)
	12/19/2021	(-0.18 kW, -8.46%)	(1.67 kW, 80.33%)	(1.85 kW, 81.87%)
	12/20/2021	(-0.08 kW, -3.86%)	(0.96 kW, 46.96%)	(1.04 kW, 48.93%)
Peak Energy Shedding	12/17/2021	(1.41 kWh, 39.27%)	(3.39 kWh, 94.42%)	(1.98 kWh, 90.81%)
	12/18/2021	(2.01 kWh, 50.6%)	(3.46 kWh, 86.9%)	(1.44 kWh, 73.48%)
	12/19/2021	(1.49 kWh, 51.23%)	(2.85 kWh, 98.21%)	(1.36 kWh, 96.33%)

(PES)	12/20/2021	(1.76 kWh, 47.61%)	(3.33 kWh, 90.18%)	(1.57 kWh, 81.25%)
Building Energy Flexibility Index (BEFI)	12/17/2021	1.07 kW	2.56 kW	1.5 kW
	12/18/2021	1.52 kW	2.61 kW	1.09 kW
	12/19/2021	1.12 kW	2.15 kW	1.03 kW
	12/20/2021	1.33 kW	2.52 kW	1.19 kW

4. Discussion

4.1. Technical contributions

The technical contributions of this paper are centered on data-driven building energy flexibility quantification, which include a framework and two case studies. Overall, the framework is a valuable tool that researchers and practitioners in the building energy field can use to quantify and evaluate energy flexibility using data-driven methods. It streamlines and standardizes the process, which can help to improve the accuracy and reliability of the results and facilitate the comparison of different studies. The case studies showcase how the framework can be applied to different types of data to support energy flexibility quantification applications.

- **Framework:** The framework provides a systematic approach to quantifying building energy flexibility using data-driven methods. The framework is designed to be generic for different use cases (e.g., benchmarking, performance tracking, and building control) and stakeholders (e.g., researchers, building owners, and grid operators) through the process of determining the type of application, checking data availability, applying data-driven modeling techniques, and selecting and calculating energy flexibility KPIs. The framework clearly defines retrospective and prospective studies for building energy flexibility and provides guidance on the data types and selection of data-driven modeling techniques that are suitable for different types of applications and KPIs.
- **Case studies:** The first case study utilized a real, large-scale smart thermostat dataset to enable benchmarking and baseline-free energy flexibility quantification. The benchmarking methodology could be used to help building owners and operators understand how their buildings perform compared to peer groups in terms of energy flexibility, and help aggregators and grid operators to better target customers and plan for future DR programs. The second case study used detailed submetering data from PNNL's twin test homes to develop a data-driven model and carry out an MPC simulation for predicting counterfactual energy consumption, which was then used to calculate EF KPIs. It provides a novel demonstration of how energy flexibility quantification can be incorporated into data-driven optimal building controls in real buildings.

4.2. Limitations

There are several limitations involved with the two case studies presented. In the first case study, the HVAC system runtime was used to calculate energy flexibility because there is no available measured energy consumption data. Although this approximation is commonly used by smart thermostat manufacturers to estimate energy savings [53,54], its accuracy might be influenced

by the HVAC system and building types. For example, single-stage and dual-stage compressors can have different correlations between energy and runtime. Therefore, more dedicated experiments are recommended to quantify the correlations for different buildings. In the second case study, there are some simplifications of the data-driven model that might not work for other times of the year. For example, the entire building's thermal properties were encoded by the thermal resistance and thermal capacitance, the lumped solar heat gain coefficient was assumed to cover the effects of window properties and solar incidence angles, and the COP of the heat pump was assumed to be constant throughout the period. Although those assumptions worked well for this case study, which only spans for a winter week, they might not hold for other periods of the year. For real applications, more complex model types (e.g., 2R2C, 3R2C) may be needed, and the model parameters can be updated online. Another limitation is that the PNNL Lab Home data were collected in experimental settings, where the electric end uses and occupancy were emulated with predefined static schedules. However, the end uses and occupancy can be highly dynamic in reality, depending upon occupant behavior, which can affect the data-driven model prediction and MPC performance. In the future, it is worth investigating how the real-world uncertainties influence the MPC performance and building energy flexibility.

It should also be noted that the framework, as a generic template, does not necessarily satisfy the needs from all stakeholders in its current form. For example, the framework might be more readily usable for building owners and operators who are mainly concerned with specific buildings because the step-by-step guidance on data curation and modeling is more relevant. For utility companies, grid operators, and aggregators who are more concerned with how much load can be altered at building cluster levels, some data types (e.g., building characteristics and BAS data) are no longer relevant. There might be extra needs to collect the data for energy flexibility quantification.

4.3. Future opportunities

The framework presented in this paper provides a comprehensive and standardized approach to quantifying energy flexibility, enabling stakeholders to better understand and optimize building energy performance. While the framework has demonstrated its effectiveness through the two case studies presented in this paper, there are opportunities for future research and applications to improve energy flexibility in real operational buildings.

- **Automated energy flexibility quantification for real buildings:** Looking towards the future, there are opportunities to leverage semantic ontologies and software tools to further automate energy flexibility quantification in real buildings. Semantic ontologies can assist mapping sensors and meters to building energy systems and link them to the variables required for energy flexibility KPI calculations. This approach can help to streamline the process of data collection and preparation, which can be time consuming and resource intensive. By automating this process, building energy flexibility could be quantified more quickly and at a lower cost. In addition, software packages could be developed to standardize and automate the calculation of energy flexibility KPIs. These tools also could help with visualization and interpretation of the KPIs, allowing stakeholders to better understand the energy flexibility of their buildings and the impact of

different operational strategies on them. Such software tools can be useful for building owners, operators, aggregators, and grid operators, as they seek to identify opportunities to optimize energy use and better participate in demand response programs. Overall, the development and integration of semantic ontologies and software tools can lead to more widespread adoption of data-driven modeling techniques, which could, in turn, help to accelerate the transition to more sustainable and resilient building energy systems.

- **Develop prototype data-driven pipelines:** Developing prototype pipelines for data-driven building energy flexibility quantification can be another valuable opportunity. These pipelines can include prototypical datasets, models, and KPIs associated with various use cases. By releasing these prototypes, different stakeholders can easily select suitable data-driven modeling techniques and KPIs for their energy flexibility quantification. This will reduce the time and resources required to develop customized models for different use cases, making the process more efficient and accessible.

5. Conclusion

As the society and building industry continues to strive towards achieving electrification and decarbonization, the concept of building energy flexibility is becoming increasingly important. In this paper, we presented a data-driven framework for quantifying energy flexibility in buildings, which can be adopted by stakeholders for applications with different levels of data availability in real-world buildings. The proposed framework has been demonstrated through two case studies using real measurements. The first case study used a large-scale smart thermostat dataset from real residential buildings and conducted a retrospective energy flexibility quantification at the building cluster level. The second case study dived deep into data-driven modeling and MPC simulation at the single building level. The two case studies have shown promising results in terms of the framework's scalability. The proposed framework and two case studies shed light on standardized data-driven building energy flexibility quantification for real world scenarios.

Acknowledgements

This research was supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, Office of Building Technologies of the United States Department of Energy, under Contract No. DE-AC02–05CH11231.

Appendix

A1. RC Model Configuration and Identification Results

This section shows the RC model configuration, parameter identification results, and performance evaluation. The heat balance of the building can be formulated as Equation 4(a) ~ 4(d).

$$C\frac{dT}{dt} = \frac{T_{indoor} - T_a}{R} + Q_{solar} + Q_{internal} + Q_{HVAC} \quad \text{Equation 4(a)}$$

$$Q_{solar} = A_{sol} I_{sol} \quad \text{Equation 4(b)}$$

$$Q_{internal} = Q_{occupants} + Q_{lights} + Q_{appliances} \quad \text{Equation 4(c)}$$

$$Q_{HVAC} = Q_{furnace} + P_{heat\ pump} * COP_{heat\ pump} \quad \text{Equation 4(d)}$$

Where C is the thermal capacitance, R is the thermal resistance, T is the lumped indoor air temperature, T_a is the outdoor air temperature, and t is time. Among the right-hand side terms, the solar heat gain Q_{solar} is the product of the lumped solar heat gain coefficient A_{sol} (considering window area, solar heat gain coefficient, and incidence angle) and the horizontal solar irradiance I_{sol}. The internal heat gains $Q_{internal}$ include those from occupants, $Q_{occupants}$; artificial lighting, Q_{lights}; and appliances $Q_{appliances}$. The supplied energy from the HVAC system Q_{HVAC} consists of the heat gain from the furnace, $Q_{furnace}$, or the cooling energy from the heat pump, which can be calculated with the power input $P_{heat\ pump}$ and a known average efficiency of coefficient of performance $COP_{heat\ pump}$. It should be noted that real heat pumps usually have on/off or multi-stage operations. However, to simplify the problem setting, we assumed the power input of the heat pump could be continuously adjusted. The unknown parameters are R, C, and A, which can be identified with a least-squares estimation.

Table A1 shows the model parameters identified using the measurements of Home B between 2021-12-6 to 2021-12-13, where the thermostat is frequently adjusted during the calibration, setpoint excitation, and pre-heating phases.

Table A1. RC model parameters identified from the measurements (2021-12-6 to 2021-12-14)

Model	Parameter	Meaning	Unit	Value
1R1C model	R	The overall thermal resistance of the building	K/W	7.96e-3
	C	The overall thermal capacitance of the building	$W \cdot S/K$	7.97e6
	A_{sol}	The overall solar heat gain coefficient of the building	m^2	7.12
2R2C model	R_i	The overall thermal resistance between envelop and air	K/W	7.68e-4
	R_e	The overall thermal resistance of the envelope	K/W	7.15e-3
	C_i	The thermal capacitance of internal objects	$W \cdot S/K$	1.66e6
	C_e	The thermal capacitance of the envelope	$W \cdot S/K$	7.5e6
	A_{sol_i}	The solar heat gain coefficient of the internal objects	m^2	8.56
	A_{sol_e}	The solar heat gain coefficient of the envelope	m^2	1.02

To evaluate the model performance, we ran simulations with both the 1R1C and 2R2C models following the actual temperature setpoint schedule in Home B between 2021-12-14 to 2021-12-20. Figure A1 shows the Home B ground truth temperature and predictions using the RC models.

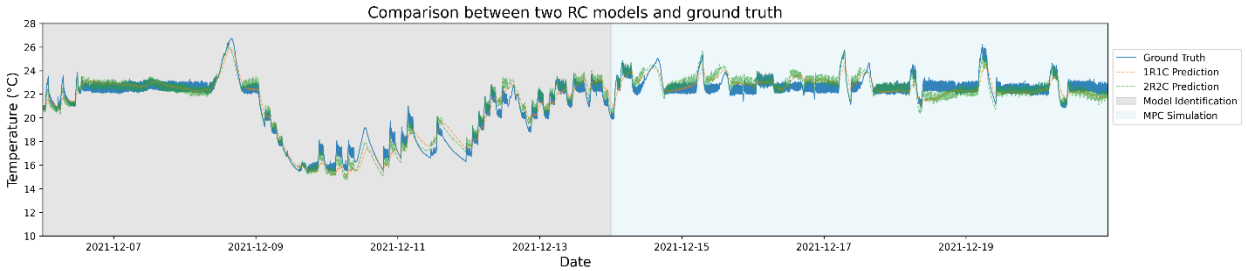


Figure A1. Evaluation of the RC models

Table A2 shows the performance metrics comparison between the two RC models during the simulation period. It can be seen that both models could predict the temperature trend with good accuracy. For this case study, we chose the 1R1C model for its simplicity.

Table A2. Accuracy metrics comparison of the RC models

Model Type	R^2	Mean Absolute Error (°C)	Mean Squared Error (°C)
1R1C	0.93	0.59	0.6
2R2C	0.94	0.54	0.54

A2. Impact of Comfort Weighting Factor on MPC

This section shows the influence of different comfort weighting factors on indoor temperature (Figure A2), out-of-range degree hours, and heat pump energy inputs (Table A3). Note that the selected $w_{comfort} = 20000$ is based on heuristics, and the value can be adjusted per user needs in practice.

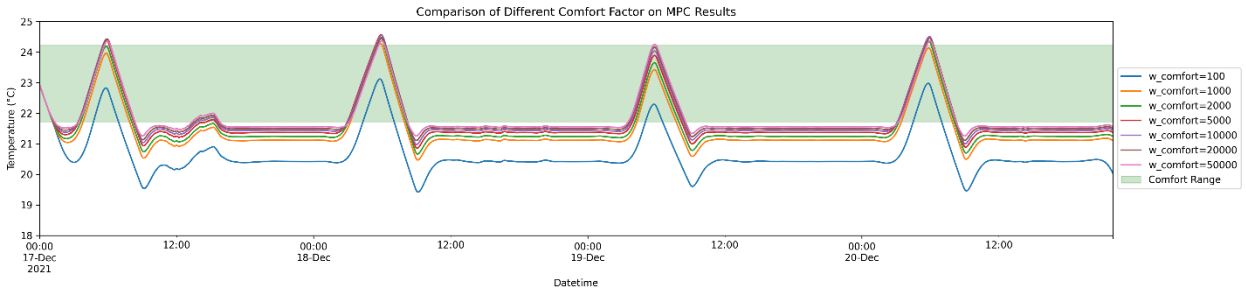


Figure A2. Comparison of different comfort weighting factors' impacts on indoor temperature

Table A3. Different comfort weighting factors' impacts on out-of-range degree-hours and heat energy

Comfort Weighting Factor	Degree-hours	Heat Pump Heating Energy (kWh)
100	106.7	42.3
1000	46.3	44.9
2000	36.3	45.3
5000	26.6	45.8
10000	21.2	46.0
20000	15.8	46.1
50000	14.3	46.8

References

[1] Renewable capacity statistics 2021, International Renewable Energy Agency (IRENA), Abu Dhabi, 2021.

[2] T. Eckman, L.C. Schwartz, G. Leventis, Determining Utility System Value of Demand Flexibility From Grid-interactive Efficient Buildings, (2020). https://escholarship.org/uc/item/0s30v1hj (accessed May 12, 2021).

[3] A. Satchwell, M.A. Piette, A. Khandekar, J. Granderson, N.M. Frick, R. Hledik, A. Faruqui, L. Lam, S. Ross, J. Cohen, K. Wang, D. Urigwe, D. Delurey, M. Neukomm, D. Nemtzow, A National Roadmap for Grid-Interactive Efficient Buildings, Ernest Orlando Lawrence

Berkeley National Laboratory, 2021. https://escholarship.org/uc/item/78k303s5 (accessed June 11, 2021).

[4] S.Ø. Jensen, A. Marszal-Pomianowska, R. Lollini, W. Pasut, A. Knotzer, P. Engelmann, A. Stafford, G. Reynders, IEA EBC Annex 67 Energy Flexible Buildings, Energy and Buildings. 155 (2017) 25–34. https://doi.org/10.1016/j.enbuild.2017.08.044.

[5] T. Weiß, A.M. Fulterer, A. Knotzer, Energy flexibility of domestic thermal loads – a building typology approach of the residential building stock in Austria, Advances in Building Energy Research. 13 (2019) 122–137. https://doi.org/10.1080/17512549.2017.1420606.

[6] J. Niu, Z. Tian, Y. Lu, H. Zhao, Flexible dispatch of a building energy system using building thermal storage and battery energy storage, Applied Energy. 243 (2019) 274–287. https://doi.org/10.1016/j.apenergy.2019.03.187.

[7] A. Gallardo, U. Berardi, Evaluation of the energy flexibility potential of radiant ceiling panels with thermal energy storage, Energy. 254 (2022) 124447. https://doi.org/10.1016/j.energy.2022.124447.

[8] E. Mlecnik, J. Parker, Z. Ma, C. Corchero, A. Knotzer, R. Pernetti, Policy challenges for the development of energy flexibility services, Energy Policy. 137 (2020) 111147. https://doi.org/10.1016/j.enpol.2019.111147.

[9] C. Xu, S. Li, X. Zhang, Energy flexibility for heating and cooling in traditional Chinese dwellings based on adaptive thermal comfort: A case study in Nanjing, Building and Environment. 179 (2020) 106952. https://doi.org/10.1016/j.buildenv.2020.106952.

[10] G. Airò Farulla, G. Tumminia, F. Sergi, D. Aloisio, M. Cellura, V. Antonucci, M. Ferraro, A Review of Key Performance Indicators for Building Flexibility Quantification to Support the Clean Energy Transition, Energies. 14 (2021) 5676. https://doi.org/10.3390/en14185676.

[11] G. Reynders, R. Amaral Lopes, A. Marszal-Pomianowska, D. Aelenei, J. Martins, D. Saelens, Energy flexible buildings: An evaluation of definitions and quantification methodologies applied to thermal storage, Energy and Buildings. 166 (2018) 372–390. https://doi.org/10.1016/j.enbuild.2018.02.040.

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

[13] S. Stinner, K. Huchtemann, D. Mueller, Quantifying the operational flexibility of building energy systems with thermal energy storages, APPLIED ENERGY. 181 (2016) 140–154. https://doi.org/10.1016/j.apenergy.2016.08.055.

[14] A. Arteconi, A. Mugnini, F. Polonara, Energy flexible buildings: A methodology for rating the flexibility performance of buildings with electric heating and cooling systems, Applied Energy. 251 (2019) 113387. https://doi.org/10.1016/j.apenergy.2019.113387.

[15] H. Tang, S. Wang, Energy flexibility quantification of grid-responsive buildings: Energy flexibility index and assessment of their effectiveness for applications, Energy. 221 (2021) 119756. https://doi.org/10.1016/j.energy.2021.119756.

[16] R. Ruusu, S. Cao, B. Manrique Delgado, A. Hasan, Direct quantification of multiple-source energy flexibility in a residential building using a new model predictive high-level controller, Energy Conversion and Management. 180 (2019) 1109–1128. https://doi.org/10.1016/j.enconman.2018.11.026.

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

[18] A. Wang, R. Li, S. You, Development of a data driven approach to explore the energy flexibility potential of building clusters, Applied Energy. 232 (2018) 89–100. https://doi.org/10.1016/j.apenergy.2018.09.187.

[19] N. Qi, L. Cheng, H. Xu, K. Wu, X. Li, Y. Wang, R. Liu, Smart meter data-driven evaluation of operational demand response potential of residential air conditioning loads, Applied Energy. 279 (2020) 115708. https://doi.org/10.1016/j.apenergy.2020.115708.

[20] G. Pinto, D. Deltetto, A. Capozzoli, Data-driven district energy management with surrogate models and deep reinforcement learning, Applied Energy. 304 (2021) 117642. https://doi.org/10.1016/j.apenergy.2021.117642.

[21] A. Kathirgamanathan, M. De Rosa, E. Mangina, D.P. Finn, Data-driven predictive control for unlocking building energy flexibility: A review, Renewable and Sustainable Energy Reviews. 135 (2021) 110120. https://doi.org/10.1016/j.rser.2020.110120.

[22] H. Johra, A. Marszal-Pomianowska, J.R. Ellingsgaard, M. Liu, Building energy flexibility: a sensitivity analysis and key performance indicator comparison, J. Phys.: Conf. Ser. 1343 (2019) 012064. https://doi.org/10.1088/1742-6596/1343/1/012064.

[23] I. Vigna, R. Lollini, R. Pernetti, Assessing the energy flexibility of building clusters under different forcing factors, Journal of Building Engineering. 44 (2021) 102888. https://doi.org/10.1016/j.jobe.2021.102888.

[24] S. Zhou, S. Cao, Energy flexibility and viability enhancement for an ocean-energy-supported zero-emission office building with respect to both existing and advanced utility business models with dynamic responsive incentives, Energy Reports. 8 (2022) 10244–10271. https://doi.org/10.1016/j.egyr.2022.08.005.

[25] E.A.M. Klaassen, C.B.A. Kobus, J. Frunt, J.G. Slootweg, Responsiveness of residential electricity demand to dynamic tariffs: Experiences from a large field test in the Netherlands, Applied Energy. 183 (2016) 1065–1074. https://doi.org/10.1016/j.apenergy.2016.09.051.

[26] P. Taddeo, A. Colet, R.E. Carrillo, L. Casals Canals, B. Schubnel, Y. Stauffer, I. Bellanco, C. Corchero Garcia, J. Salom, Management and Activation of Energy Flexibility at Building and Market Level: A Residential Case Study, Energies. 13 (2020) 1188. https://doi.org/10.3390/en13051188.

[27] D. Blum, J. Arroyo, S. Huang, J. Drgoňa, F. Jorissen, H.T. Walnum, Y. Chen, K. Benne, D. Vrabie, M. Wetter, L. Helsen, Building optimization testing framework (BOPTEST) for simulation-based benchmarking of control strategies in buildings, Journal of Building

Performance Simulation. 14 (2021) 586–610. https://doi.org/10.1080/19401493.2021.1986574.

[28] J. Vivian, U. Chiodarelli, G. Emmi, A. Zarrella, A sensitivity analysis on the heating and cooling energy flexibility of residential buildings, Sustainable Cities and Society. 52 (2020) 101815. https://doi.org/10.1016/j.scs.2019.101815.

[29] S.S. Amiri, M. Mottahedi, S. Asadi, Using multiple regression analysis to develop energy consumption indicators for commercial buildings in the U.S., Energy and Buildings. 109 (2015) 209–216. https://doi.org/10.1016/j.enbuild.2015.09.073.

[30] N. Fumo, M.A. Rafe Biswas, Regression analysis for prediction of residential energy consumption, Renewable and Sustainable Energy Reviews. 47 (2015) 332–343. https://doi.org/10.1016/j.rser.2015.03.035.

[31] B. Bueno, L. Norford, G. Pigeon, R. Britter, A resistance-capacitance network model for the analysis of the interactions between the energy performance of buildings and the urban climate, Building and Environment. 54 (2012) 116–125. https://doi.org/10.1016/j.buildenv.2012.01.023.

[32] Z. Wang, T. Hong, M.A. Piette, Building thermal load prediction through shallow machine learning and deep learning, Applied Energy. 263 (2020) 114683. https://doi.org/10.1016/j.apenergy.2020.114683.

[33] D. Kim, J.E. Braun, A general approach for generating reduced-order models for large multi-zone buildings, Journal of Building Performance Simulation. 8 (2015) 435–448. https://doi.org/10.1080/19401493.2014.977952.

[34] G. Gokhale, B. Claessens, C. Develder, Physics informed neural networks for control oriented thermal modeling of buildings, Applied Energy. 314 (2022) 118852. https://doi.org/10.1016/j.apenergy.2022.118852.

[35] K. Amasyali, N.M. El-Gohary, A review of data-driven building energy consumption prediction studies, Renewable and Sustainable Energy Reviews. 81 (2018) 1192–1205. https://doi.org/10.1016/j.rser.2017.04.095.

[36] M. Bourdeau, X. qiang Zhai, E. Nefzaoui, X. Guo, P. Chatellier, Modeling and forecasting building energy consumption: A review of data-driven techniques, Sustainable Cities and Society. 48 (2019) 101533. https://doi.org/10.1016/j.scs.2019.101533.

[37] H. Li, T. Hong, A semantic ontology for representing and quantifying energy flexibility of buildings, Advances in Applied Energy. 8 (2022) 100113. https://doi.org/10.1016/j.adapen.2022.100113.

[38] H. Johra, energy-flexibility-kpis: Python package for the computation of energy flexibility KPIs, (n.d.). https://github.com/HichamJohra/energy_flexibility_kpis (accessed March 13, 2023).

[39] L. Zhang, J. Wen, A systematic feature selection procedure for short-term data-driven building energy forecasting model development, Energy and Buildings. 183 (2019) 428–442. https://doi.org/10.1016/j.enbuild.2018.11.010.

[40] Donate your Data Smart Wi-Fi Thermostats by ecobee, (2016). https://www.ecobee.com/donate-your-data/ (accessed November 7, 2022).

[41] ENERGY STAR Program Requirements for Connected Thermostats Version 1.0, (2017).

[42] J.L. Dréau, Demand-Side Management of the Heating Need in Residential buildings, in: 2016. https://hal.science/hal-01988234 (accessed March 13, 2023).

[43] PNNL: Lab Homes, (n.d.). https://labhomes.pnnl.gov/ (accessed March 13, 2023).

[44] Time of Use, (n.d.). https://www.pacificpower.net/savings-energy-choices/time-of-use.html (accessed March 13, 2023).

[45] All you need to know about model predictive control for buildings - ScienceDirect, (n.d.). https://www.sciencedirect.com/science/article/pii/S1367578820300584 (accessed April 18, 2022).

[46] J. Drgoňa, M. Klaučo, M. Kvasnica, MPC-based reference governors for thermostatically controlled residential buildings, in: 2015 54th IEEE Conference on Decision and Control (CDC), 2015: pp. 1334–1339. https://doi.org/10.1109/CDC.2015.7402396.

[47] W. Liang, R. Quinte, X. Jia, J.-Q. Sun, MPC control for improving energy efficiency of a building air handler for multi-zone VAVs, Building and Environment. 92 (2015) 256–268. https://doi.org/10.1016/j.buildenv.2015.04.033.

[48] S. Rastegarpour, S. Gros, L. Ferrarini, MPC approaches for modulating air-to-water heat pumps in radiant-floor buildings, Control Engineering Practice. 95 (2020) 104209. https://doi.org/10.1016/j.conengprac.2019.104209.

[49] S. Lucia, A. Tătulea-Codrean, C. Schoppmeyer, S. Engell, Rapid development of modular and sustainable nonlinear model predictive control solutions, Control Engineering Practice. 60 (2017) 51–62. https://doi.org/10.1016/j.conengprac.2016.12.009.

[50] J.A.E. Andersson, J. Gillis, G. Horn, J.B. Rawlings, M. Diehl, CasADi: a software framework for nonlinear optimization and optimal control, Math. Prog. Comp. 11 (2019) 1–36. https://doi.org/10.1007/s12532-018-0139-4.

[51] H. Zhang, A. Tzempelikos, X. Liu, S. Lee, F. Cappelletti, A. Gasparella, The impact of personal preference-based thermal control on energy use and thermal comfort: Field implementation, Energy and Buildings. 284 (2023) 112848. https://doi.org/10.1016/j.enbuild.2023.112848.

[52] Lab Homes, (n.d.). https://bbd.labworks.org/ds/bbd/lh (accessed March 13, 2023).

[53] Savings from your ecobee, (n.d.). https://www.ecobee.com/en-us/savings/ (accessed March 13, 2023).

[54] Energy Savings from the Nest Learning Thermostat: Energy Bill Analysis Results, (2015). https://storage.googleapis.com/nest-public-downloads/press/documents/energy-savings-white-paper.pdf.