

UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

DESIGN AND ANALYSIS OF SMART HOME ENERGY MANAGEMENT SYSTEM
FOR ENERGY-EFFICIENT AND DEMAND RESPONSE OPERATIONS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

Yilin Jiang
Norman, Oklahoma
2023

DESIGN AND ANALYSIS OF SMART HOME ENERGY MANAGEMENT SYSTEM
FOR ENERGY-EFFICIENT AND DEMAND RESPONSE OPERATIONS

A DISSERTATION APPROVED FOR THE
SCHOOL OF AEROSPACE AND MECHANICAL ENGINEERING

BY THE COMMITTEE CONSISTING OF

Dr. Li Song, Chair

Dr. Choon Yik Tang

Dr. Jie Cai

Dr. Pejman Kazempoor

Dr. Meijun Zhu

DEDICATION

To my father and grandfather in law, in loving memory.

ACKNOWLEDGMENTS

The success of this study would not have been possible without the guidance and collaboration of my supervisor, Dr. Li Song (chair of my dissertation committee). Her insights and advice have shaped every phase of this project, from model development to results analysis. I would like to also express my gratitude to Dr. Choon Yik Tang for his substantial contributions and support throughout the project.

I would like to thank lab mates Junke Wang, Andrew Ejenakevwe, Rodeney Hurts, Luyao Xie, and Marwan Hashem in the Building Energy Efficiency Laboratory (BEEL) at OU. Without their kindness help and support, I would not gain so much knowledge and happiness during my PhD journey. Special thanks to Dr. Michael Brambley and his team working on connected homes, for helping with water heater construction and many technical suggestions in our early stage.

Great appreciation goes to the homeowners who participate in our precooling tests.

I would like to thank my dissertation committee, who are my qualify exam committee too, Dr. Cai Jie, Dr. Pejman Kazempoor, and Dr. Meijun Zhu.

Part of this work was supported through the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Building Technologies Office. I am very grateful to the support.

I would like to thank my family. Without their support, I would have been impossible to finish the study.

ABSTRACT

In the movie “Iron Man”, Tony Stark, with his highly connected and smart home system, shows the audience an appealing vision of future work and domestic life. Many audiences desire such a living environment where they can not only interact with their homes but also let the homes manage their operation automatically. As technology progressively steps into such a future, realizing a responsive and autonomous smart home is not just a fantasy. To establish grid-interactive homes that help save costs for users and improve grid reliability, this study introduces an energy management framework for smart home environments. This framework provides optimal operation of multiple appliances, taking into account dynamic responses to external factors such as outside weather conditions, homeowner’s preferences, and particularly, grid conditions like time-varying pricing in demand response programs.

As one of the largest energy consumers in the home, the operation of the HVAC system holds great potential for cost savings and energy flexibility—the latter being the ability to adjust its consumption based on grid signals such as time-of-use (TOU) pricing. Achieving cost savings and energy flexibility requires intelligent strategies, one of which is precooling—a control strategy where an air conditioner (AC) cools space when the electricity price is low to avoid expensive operation when the electricity price is high. In previous studies, Model Predictive Control (MPC)-based precooling strategies are typically analyzed through simulations, and field studies in residential buildings are quite limited. In this study, we developed an MPC agent and carried out extensive field tests on nine homes over a period of

four months in Oklahoma and Miami. Filed test results show that the MPC agent can reduce energy cost by 28.72%–51.31% on hot summer days and by up to 60.32% on mild summer days, in addition to achieving significant energy flexibility. Moreover, the agent's performance is found to be most impacted by weather conditions, AC performance, user comfort preferences, and floor areas of the homes.

In addition, to further comprehend diverse factors that may impact the results of MPC-based precooling, an EneyPlus virtual testbed and a corresponding control framework for co-simulation are developed. The purpose of developing such a virtual testbed is to create a simulation environment that enables experiments without the limitation and variability of field tests. The virtual testbed is modified by using the Python script to mimic the on/off cycle in the majority of U.S. residential building HVAC systems. By conducting the sensitivity analysis and ablation study, the MPC-based precooling co-simulation results are evaluated. It was observed in our case study that cost savings achieved through MPC-based precooling were primarily influenced by the use of forecast weather. The accuracy of the models and the prediction horizon of the MPC models also plays a substantial but lesser extent role.

With the optimal operation framework shifting from the HVAC system to multiple appliances, the proposed energy management framework has a broader scope, encompassing not only the HVAC system but also water heaters, non-thermal appliances, and the power flow between photovoltaics panel (PV), batteries, and the grid. Apart from the cost-savings and energy flexibility that can be achieved, the proposed framework also provides a more realistic

simulation scenario by considering the user's appliance time usage preference, water usage, and thermal comfort preferences. Finally, the framework also embedded multi-objective optimization to support the homeowner's decision-making between cost saving and thermal comfort.

Overall, this study aims to realize the optimal operation of various load-flexible resources under demand response programs in residential buildings. This study investigates the fundamental research for the investigation of methodologies to enhance and understand the interactions between buildings, homeowners, and the grid. Due to the flexibility of the model, this study can be adapted to other residential buildings and even in larger communities.

Table of Contents

DEDICATION	iv
ACKNOWLEDGMENTS	v
ABSTRACT	vi
1 Introduction	1
1.1 Motivation.....	1
1.2 DRPs and Gap Analysis in Residential Buildings	4
1.2.1 <i>DRPs offered by utilities</i>	4
1.2.2 <i>Building energy management strategies and the focus of this study</i>	6
1.2.3 <i>Available load-flexible resources in residential buildings</i>	7
1.2.4 <i>Deployment barriers in demand response participation</i>	10
1.2.4.1 <i>Technology barriers: hardware</i>	11
1.2.4.2 <i>Technology barriers: software</i>	14
1.2.4.3 <i>Economic and social barriers</i>	15
1.3 Research Problem	17
1.4 Dissertation Outline	21
2 In-depth Literature Review	23
2.1 Optimal Control of the HVAC System	23
2.1.1 <i>HVAC precooling strategy and control methods</i>	23
2.1.2 <i>Current MPC-based precooling controller issues</i>	27
2.1.3 <i>Possible solutions for MPC-based precooling controller</i>	30
2.2 Smart Home Energy Management (Home-EMS) System	31
3 Load-flexible Resources Modeling in Residential Buildings	36
3.1 Introduction.....	36
3.2 Model Type Selection for Load-flexible resources in a Home	36
3.3 HVAC Modeling and Parameter Identification Approach	39
3.3.1 <i>Reduced-order RC-based gray-box model of home thermal dynamics</i> ...	39
3.3.2 <i>2R2C parameter identification approach</i>	41
3.3.3 <i>HVAC system power estimation</i>	45
3.4 Water Heater Modeling.....	46
3.4.1 <i>Reduced-order RC-based gray-box model of water heater thermal dynamics</i>	46
3.4.2 <i>Parameter identification for water heater model</i>	47
3.5 Short-term PV Power Generation Prediction.....	48
3.5.1 <i>Data preprocessing</i>	49
3.5.2 <i>Feature selection</i>	50
3.5.3 <i>Hyperparameter tuning</i>	51
3.6 Model Validation.....	52
3.6.1 <i>The Lab House and data sources</i>	53
3.6.1.1 <i>HVAC system data acquisition</i>	54
3.6.1.2 <i>Retrofit a conventional electric water heater into a smart one</i>	57

3.6.1.3	PV forecast data sources	60
3.6.2	<i>HVAC parameter identification results</i>	60
3.6.3	<i>Water heater calibration and parameter identification results</i>	61
3.6.4	<i>Short-term PV power generation results</i>	62
3.7	Summary	65
4	Development, Implementation, and Impact Analysis of MPC-based Precooling for Homes	67
4.1	Introduction.....	67
4.2	Real-time Control Framework of the MPC-based Precooling.....	68
4.2.1	<i>Mixed-integer linear programming problem formulation</i>	68
4.2.2	<i>MPC agent</i>	70
4.2.3	<i>Performance indicators for MPC agent</i>	71
4.3	MPC-based Precooling in Real Homes.....	72
4.3.1	<i>Cloud-based platform for data management and control</i>	72
4.3.2	<i>Implementation setup</i>	74
4.3.2.1	Introduction of eleven (11) participating test homes	74
4.3.2.2	Temperature and power measurements.....	78
4.3.3	<i>MPC precooling test filed test results</i>	79
4.3.3.1	MPC performance assessment at one test home	79
4.3.3.2	MPC performance assessment at multiple test homes	81
4.4	MPC-based Precooling at Virtual Testbed	90
4.4.1	<i>Virtual Testbed Development</i>	91
4.4.1.1	Residential building modelling	91
4.4.1.2	EMS for on/off control with deadband	92
4.4.1.3	Internal thermal mass modification	94
4.4.1.4	EnergyPlus model validation	95
4.4.2	<i>Evaluate MPC-based precooling performance at the virtual testbed</i>	96
4.4.2.1	MPC-based precooling control workflow.....	97
4.4.2.2	MPC hyperparameter choice	98
4.4.2.3	Test results and impact analysis for MPC hyperparameters	100
4.5	Summary	105
5	Connected Home Energy Management for Grid-Interactive Building	108
5.1	Introduction.....	108
5.2	Smart Home Scheduling Formulation	110
5.2.1	<i>Baseline model</i>	111
5.2.1.1	Non-thermal appliances	112
5.2.1.2	HVAC system.....	114
5.2.1.3	Water heater system and its coupling effect.....	115
5.2.2	<i>Grid-tie PV system</i>	117
5.2.3	<i>Hybrid PV system with AC battery</i>	119

5.2.4	<i>Occupancy behavior and user preference model</i>	121
5.2.4.1	<i>User time preference on appliance usage</i>	121
5.2.4.2	<i>Water usage profile</i>	123
5.2.5	<i>User-preferred thermal comfort range</i>	124
5.3	<i>Case Study</i>	127
5.4	<i>Result and Discussion</i>	132
5.4.1	<i>Three optimal operations</i>	132
5.4.2	<i>Energy flexibility provided by different system operations</i>	137
5.4.3	<i>Impact of user time preference and weather condition</i>	140
5.4.4	<i>Cost and comfort trade-off</i>	143
5.5	<i>Summary</i>	146
6	Conclusion, Limitations and Future Studies	149
6.1	<i>Conclusion of this study</i>	149
6.2	<i>Limitations</i>	152
6.2.1	<i>Internal heat gain and occupancy</i>	152
6.2.2	<i>Growth of the problem size</i>	152
6.2.3	<i>Thermal comfort</i>	153
6.3	<i>Future Studies</i>	154
6.3.1	<i>Meta-model to quickly evaluate energy flexibility potential of a home</i>	154
6.3.2	<i>Data-driven occupancy behavior model</i>	155
6.3.3	<i>Stochastic optimizations consider uncertainty</i>	155
	Appendix	157
	Appendix A: Matrix Profile-based Weather Clustering Algorithm	157
	References	159

List of Tables

Table 1.1 Five ways building can provide value to the grid [11,12].....	6
Table 1.2 Research questions and hypothesis.	18
Table 2.1 Advantage and disadvantage of major control strategies on implementing precooling	26
Table 2.2 Summary of SHEM Scheduling Models.....	33
Table 3.1 Model characteristics [99] and their application in this study	37
Table 3.2 Possible values of hyperparameters used in the grid search	51
Table 3.3 The performance results on testing dataset	65
Table 4.1 Three Types of Dataset.....	73
Table 4.2 Test home information.	75
Table 4.3 Cost and energy consumption summary for homes on hot summer days.....	85
Table 4.4 Cost and energy consumption summary for homes on mild summer days.....	86
Table 4.5 Cost and energy consumption summary for homes on cold summer days.....	86
Table 4.6 EnergyPlus model variables for on/off control	93
Table 4.7 MPC hyperparameters choice for impact analysis.....	99
Table 5.1 Other appliances specifications.....	128
Table 5.2 PV panel specifications for test location.....	129
Table 5.3 Battery system specifications.....	130
Table 5.4 User time preference interval used in case study	131

List of Figures

Figure 1.1 The official “Duck Chart” during spring, shows steep ramping needs and over-supply risk, published by CASIO [3].....	2
Figure 1.2 Demand side management and DRPs in retail and wholesale markets.....	5
Figure 1.3 TOU rate examples in Oklahoma and Miami.....	6
Figure 1.4 Load-flexible resources, technologies, and their deployment limitations in residential buildings	18
Figure 2.1 Two precooling strategies under DRPs	24
Figure 2.2 Prediction horizon, control horizon and timestep for MPC	25
Figure 3.1 Schematic of the RC network of home thermal dynamics	40
Figure 3.2 Physics constraints explained for system identification process.....	44
Figure 3.3 Water heater heat and water flow	47
Figure 3.4 The schematic diagram of PV power generation prediction	49
Figure 3.5 Lab House information: (a) outside view of the front of the Lab House (b) an outdoor unit in the back yard of the Lab House (c) an indoor unit in the attic of the Lab House	54
Figure 3.6 Lab House sensors: (a) the floor plan of the Lab House and sensor locations (b) a smart thermostats and thermal couples to measure indoor air temperature (c) a node sensor to measure interior wall surface temperature (d) a power meter inside the electrical panel.	57

Figure 3.7 Retrofitted smart water heater system.	58
Figure 3.8 (a) 2R2C model validation result and (b) AC outdoor unit consumption regression result (power lower than 3.2kW is excluded).....	61
Figure 3.9 Water heater system identification result.....	62
Figure 3.10 Model performance improvement (a) SHAP value for every train data points in feature selection for September (b) RMSE results on validation sets on September in grid search.....	64
Figure 3.11 PV power generation, G _{hi} , and relative humidity in original data on September 2, 2019	64
Figure 3.12 PV power generation point and interval prediction compared with actual generation from August 3, 2019 to August 5,2019	65
Figure 4.1 MPC agent.....	71
Figure 4.2 Proposed cloud-based platform for data management and control.	73
Figure 4.3 Test homes with diverse characteristics. (a) Building floor area distribution. (b) Count plot of AC tonnage. (c) Building ages (inner) and AC unit service time (outer).	76
Figure 4.4 Setpoint distribution in several test homes.....	77
Figure 4.5 Equipment installed at test homes. (a) Smart thermostat. (b) Node sensor. (c) Smart power meter.	79
Figure 4.6 Seven consecutive days of MPC-based precooling test results. (a) Smart	

thermostat and node sensor measurements. (b) AC signals and TOU rate. (c) Actual weather conditions.	81
Figure 4.7 Weather clustering results in (a) Norman and (b) Miami from July to October	83
Figure 4.8 Cost saving summary for each home on hot summer days for normal operation days and MPC operation days.....	84
Figure 4.9 Cost saving summary for each home on mild summer days for normal operation days and MPC operation days.....	85
Figure 4.10 Flexibility factors during the entire test period except for extreme days.....	87
Figure 4.11 MPC agent performance assessment: (a) FF versus floor area in Norman test homes (b) distribution of FF for different temperature bands in the Lab House (c) average FF for different refrigerant charging conditions in the Lab House	89
Figure 4.12 Energy model built for the Lab House in Norman, Oklahoma	92
Figure 4.13 Proposed Erl for on/off control with deadband	94
Figure 4.14 Indoor air temperature comparison between actual smart thermostat measurement and EnergyPlus simulation	95
Figure 4.15 Daily energy consumption comparison with outdoor air temperature	96
Figure 4.16 MPC-based precooling control through Python plugin and EnergyPlus virtual testbed.	98
Figure 4.17 An example of a 2R2C model, a 2R2CreduceAC model, and a	

2R2Creducter model simulation results compared with the EnergyPlus outputs. (a)	
Indoor air temperature results. (b) Interior wall surface temperature results. (c) AC	
impact and reduced AC impact to the home thermal dynamics.....	100
Figure 4.18 MPC-based precooling test result on August 6 at virtual testbed.....	101
Figure 4.19 Sensitivity analysis on optimization horizon in terms of cost saving with	
forecasted weather and with perfect weather.	102
Figure 4.20 Ablation study using cumulative AC operation cost: compare ideal case (bold	
green) to baseline case (dashed black) and five different ablations.....	104
Figure 4.21 Performance comparison using cumulative AC operation cost: compare	
baseline case (bold black) to idea case (dashed green) and five different cases	
between them.	104
Figure 5.1 Three different types of PV Systemsm: Off-grid (left), Grid-tie (middle), and	
Hybrid System (right).	109
Figure 5.2 An example of ATUS dataset.....	122
Figure 5.3 Example of the probability density function of (a) Kitchen clean up and (b)	
Laundry.	123
Figure 5.4 Hot water usage data collected from a single-family water heater [139].....	124
Figure 5.5 Comfort reward examples.	126
Figure 5.6 Hot water usage profiles (a) hot water usage profile from July to August (b)	
most representative hot water usage 24-hour profile.	131

Figure 5.7 Baseline operation of five appliances on August 7.	134
Figure 5.8 Grid-tie system operation of five appliances on August 7.	135
Figure 5.9 Hybrid system operation of five appliances on August 7.....	136
Figure 5.10 Power flow comparison (a) total power buy from grid for three different operations (b) battery (dis-)charging profiles.	137
Figure 5.11 Power flow for (a) Grid-tie system (b) Hybrid System with a medium battery (c) Hybrid System with a small battery (d) Hybrid System with a large battery...	138
Figure 5.12 Performance indices for different system operations. (a) PV self-consumption (b) PV self-sufficiency (c) Flexibility factor.....	140
Figure 5.13 Cost-saving variation in terms of different weather conditions, user preference, and operation scenarios.....	141
Figure 5.14 User preference impact on appliance operation. (a) Grid-tie system appliance operation given strict user time preference. (b) Grid-tie system appliance operation given flexible user time preference.....	143
Figure 5.15 Multi-objective optimization results where HVAC system behavior shown only. (a) choose w_1 equals 0.1 (b) choose w_1 equals 0.9.	144
Figure 5.16 Scheduling result with the multi-objective framework with equal weights on cost and comfort (w_1 equals 0.5)	145
Figure 5.17 Cost and comfort trade-off.	146

Figure A. 1 Three stages of the modified matrix profile-based weather clustering algorithm
..... 158

1 Introduction

This study aims at developing and implementing an intelligent system to coordinate multiple electric loads in residential buildings. In this chapter, the motivation of the study is introduced first, followed by state of the art and gap analysis. Research questions, the scope of the research, and the main contributions of this dissertation are summarized at the end.

1.1 Motivation

In 2022, about 22% of electricity produced at utility-scale generation facilities was from renewable energy [1]. Although variable renewable energy (VRE) resource emits little to no greenhouse gas and is sustainable, it brings reliability risk to the grid due to their inherent variability. A well-known phenomena, the California Duck, reveals that one of the main manifestations of the grid reliability problem is that when solar energy has a high penetration in the grid, there is a high probability of an demand and supply imbalance at different times [2,3]. The “Duck Chart”, which describes the change in net load (the difference between the forecasted demand and the expected generation), was first published in 2016 in CASIO’s report [3], see Figure 1.1. The “belly” appearance in the mid-afternoon is due to the excess capacity of renewable energy (oversupply), and the “duck neck” appearance links to the ramping demand in the late afternoon and the decrease in the renewable energy when sunsets (undersupply). Since a reliable grid requires a continuous supply/demand balance to maintain the frequency of the power system (e.g., 60Hz in the US), facility operators may choose to manually curtail to maintain the base load of generators when oversupply occurs; or they could

dispatch conventional generators resources to meet ramping demand to handle undersupply issue. However, these strategies cause economic or environmental loss and thus are not long-term solutions.

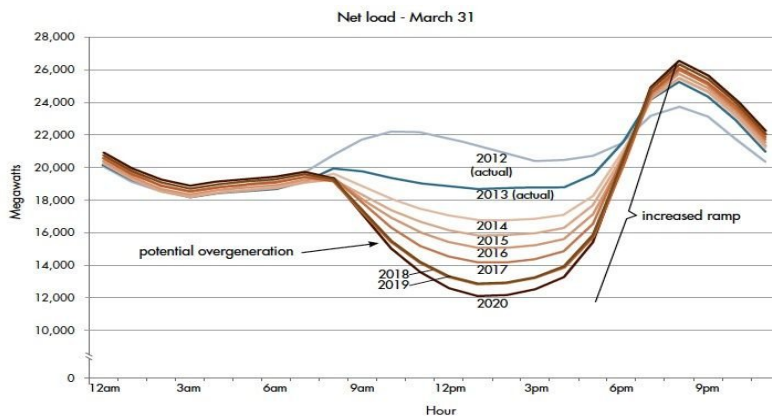


Figure 1.1 The official “Duck Chart” during spring, shows steep ramping needs and over-supply risk, published by CASIO [3].

In [4], by comparing different solutions to handle grid reliability issues, the authors indicated that only flexibility technologies, i.e., the technology that contributes to the system flexibility by generating and consuming active power, can address the supply-demand balance issue. However, some flexibility technologies, including developing grid-friendly PV plants or modifying conventional generators, require significant investments. In recent years, a simpler and less expensive approach, called Demand Side Management (DSM), has attracted attention.

DSM is a collection of flexibility strategies and techniques designed to manage electricity consumption at the end-user side to facilitate reliable and sustainable grid [5]. Among all end users, residential buildings hold great potential to address grid reliability issue. As one of the key drivers of electricity demand, residential buildings consume 21% of total U.S. energy in

2021 when energy losses during generation, transmission, and distribution are included [6]. Space cooling contributes about 16% of total household electricity consumption and will generally increase over the next 30 years due to climate and population changes [7], therefore HVAC system operation becomes one of the focuses of this study. Meanwhile, as the penetration of solar panels and active energy storage increases in residential buildings, PV-battery systems become another effective source for DSM [5]. In 2015, since roughly half of the energy usage was contributed by electrical appliances, on top of space heating, air conditioning, water heating, and refrigerators, EIA added 22 categories such as dishwashers, clothes washers, and clothes dryers into household energy usage [8]. These electrical devices also bring opportunities for DSM since they are more flexible compared with thermostatically appliance.

Due to its significant potential on addressing grid reliability, many utilities offered demand response programs (DRPs) in retail level, in hoping to facilitate the residential buildings to adjust their power demand based on grid signals. Unfortunately, the participation of residential users in DRP remains low [9]. The primary reason is the lack of an intelligent control and coordination system, designed for the residential customers, to optimize the operation of appliances under the DRPs. This intelligent control, which ideally should be low-cost, effective, and align with user preference, could improve operation efficiency, increase home appliances' response speed to the grid signal, and save times for users. However, designing such intelligent control for home appliances often involves navigating a complex system: the thermal

interactions between indoors and outdoors, occupancy behavior patterns, and the operation characteristic of the appliances all needed to be predicted or simulated. In the following section, current deployment barrier are reviewed for development of such intelligent control.

1.2 DRPs and Gap Analysis in Residential Buildings

In this chapter, the concepts and examples of demand response programs (DRPs) in residential buildings are introduced first. To deploy the DRPs, the load-flexible resources and their accompanying smart technologies that hold potential for enabling energy management in the residential buildings (from demand side) are then reviewed. Building on these discussions, the deployment barriers of the DRPs in residential buildings are discussed. It is important to recognize that improving the participation or involvement of DRPs from the demand side is not only use the current load-flexible resources and their technologies, but also understand and overcome these deployment barriers which encompass hardware, software, financial and social factors. Afterward, research questions are formulated based on this review. Finally, the outline of this dissertation are demonstrated.

1.2.1 DRPs offered by utilities

New challenges arise after the integrating of renewable energy into the grid, particularly in supply-demand balance issue. The building's high demand is one of the reasons for this balance issue, but it can also be the most efficient solution due to its flexibility. Building load, of which 55% is consumed by residential buildings [8], can be adjusted through proper management.

If a building adjusts its load profile based on the signal that comes from the grid, it is called demand response (DR). As utilities realize the value of Demand Side Management (DSM), more Demand Response Programs (DRPs), aiming to motivate users to conduct DR by providing time-based rates or incentives, start to be focused on the retail level.

Figure 1.2 Demand side management and DRPs in retail and wholesale markets lists major DRPs at the wholesale market level and at the retail level. Among those DRPs, time-based program such as Time-of-Use (TOU), which aims to provide time-varying price to stimulate users to adjust their load [10], is the most established program and has already been used in commercial and residential buildings.

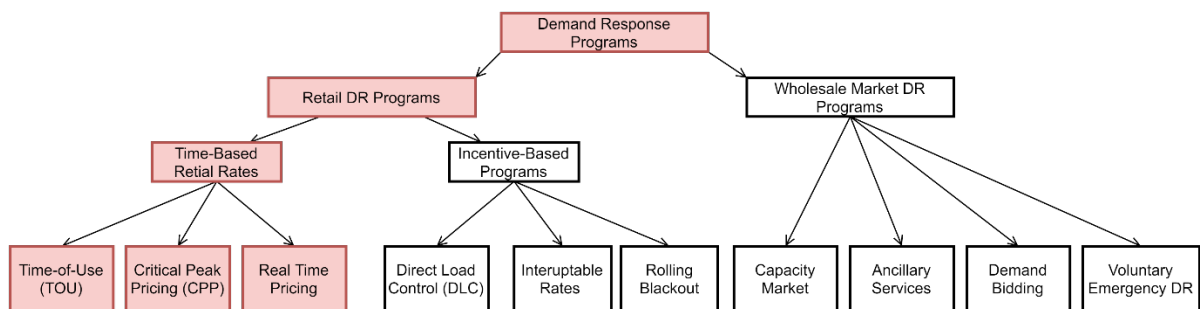


Figure 1.2 Demand side management and DRPs in retail and wholesale markets

Figure 1.3 displays a comparison of Oklahoma’s and Miami’s 2022 TOU rates, which are actual quotes offered by utilities to their customers. In some periods the price is higher (the on-peak hour) and in some other periods the price is lower (the off-peak hour). But the time span of the high price period, and the magnitude of the high prices of two TOU rates are different, depending on different preferences from different utilities.

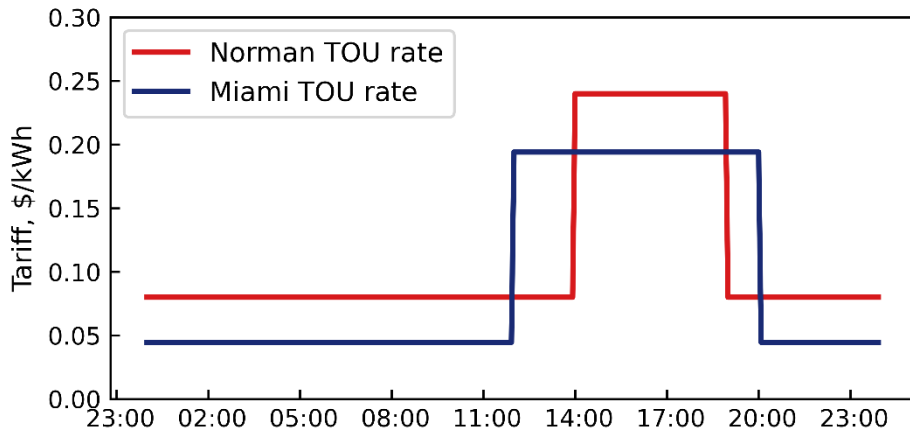


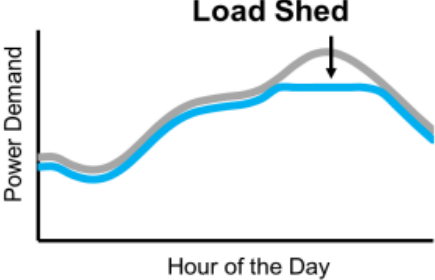
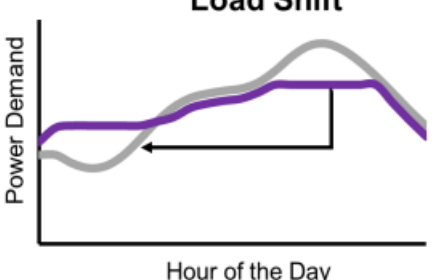
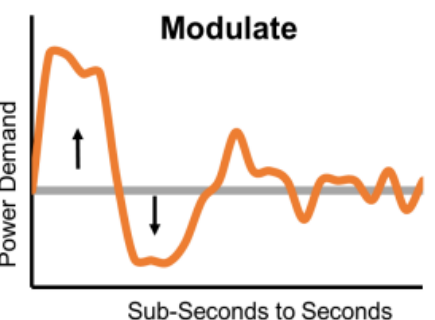
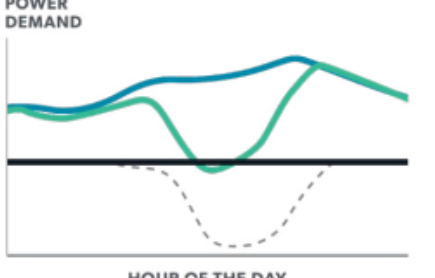
Figure 1.3 TOU rate examples in Oklahoma and Miami

1.2.2 Building energy management strategies and the focus of this study

Since the time-based DRPs, such as the TOU rate, are designed to incentivize users rather than directly control the electricity load, the demand-side management strategies for building load adjustments becomes particularly important. Based on the DOE report in 2019 [11], there are five management strategies that a building can conduct to adjust its load, as shown in Table 1.1. The load shifting and generation in residential buildings are the focus in this study.

Table 1.1 Five ways building can provide value to the grid [11,12]

Name	Characteristics	Example load profile
Efficiency	Aims to reduce long-term energy use regardless of time	

Load shed	Aims to provide immediate energy reduction during peak hours or an emergent event	
Load shift	Buildings change the timing of energy usage; usually means intentional and planned shifting	
Modulate	Building automatically responds to the grid signals during the dispatch period in seconds and sub-second	
Generation	Building generates energy for its on-site energy usage	

1.2.3 Available load-flexible resources in residential buildings

To enable the full potential of residential buildings' capability on load shifting and generation, lots of research and development efforts have been devoted to sensing and control of specific resources that can facilitate load coordination in residential buildings. Several review studies have been done to discuss those resources[4,13,14]. It is commonly accepted

that four major resources can provide the shifting and generation capabilities:

1) *Thermostatically controlled load*: This energy load mainly refers to the traditional energy end-users like space heating, air conditioning, and water heating, which consumes 42%–46% of energy in residential buildings [8,15]. Depending on the operation method, the thermostatic control load can be divided into two parts:

a) *Passive thermal storage*: passive thermal storage refers to the ability of building itself to help coordinate the electricity demand of the HVAC system. A building's envelope and its internal thermal mass can unleash significant potential serving as a heating or cooling energy storage in precooling or preheating, i.e., cooling or heating the space at the appropriate time to shift the load [16]. It should not be overlooked that the HVAC system directly affects the user's thermal comfort. To conduct the precooling or preheating without compromising the user experience, the indoor air temperature dynamics under HVAC operation must be accurately modeled. However, this modeling is one of the most complicated tasks in DSM, since indoor air temperature dynamics involve time-varying outside disturbance (e.g., weather conditions), occupant preferences, and complicated HVAC system (different components and phase change involved) [17–19].

b) *Active thermal storage*: active thermal storage refers to thermal storage tank which can charge and discharge energy through the thermal medium at the appropriate time to shift or shed the load. In residential buildings, activated thermal storage usually

refers to a water heater. Noticed that except for electricity, water is the medium that connects many different appliances such as clothes washers. This coupling effect contributes to the difficulty of its thermal modeling. In addition, for user satisfaction and safety, thermal comfort limitation of water heaters is also needed to consider. For example, it is reported that harmful bacteria such as Legionella may occur if the water temperature at the tank is lower than 120°F [20]. This requires accurate modeling of the heat transfer inside the water tank as well.

- 2) *Non-thermostatically controlled load*: non-thermostatically controlled load refers to end-users that are not affected by the thermal dynamic mechanism. In residential buildings, they can be lighting, electronic, and electric appliances. However, not all the appliances are suitable for adjusting in residential buildings. It is not proper to coordinate some appliances such as TV, computer, or lights, as they not only do not consume high levels of energy, but also adjusting them may cause significant inconvenience to the users. In energy management studies, those appliances are called non-schedulable appliances, while some appliances such as dishwashers, clothes washers, and dryers are called schedulable appliances [13].
- 3) *Distributed energy resources*: distributed energy resources (DER) include the self-generator such as PV panels and wind turbines. DER is another important resource that can provide energy flexibility to the home. The self-generation energy provides load-shedding capability, and meanwhile, it can combine with the other appliances or energy

storage equipment such as battery to further provide load-shifting capabilities [21]. The most used self-generator is rooftop photovoltaic (PV) panels in residential buildings. Noticed that the generation efficiency of the PV is not only limited by outside weather, its distribution networks, and the rooftop availability but also the inverter efficiency (DC to AC) [22].

- 4) *Electricity storage*: Electricity storage usually refers to batteries or electric vehicles. Similar to active thermal storage, it can provide both load shifting and shedding by charge and discharge operation, but no need to involve with the thermal medium. Different from passive or active thermal storage, which might need time to precool or preheat the space or thermal medium, electricity storage provides electricity flexibility immediately. However, it is also limited by battery cell temperature, state of charge, depth of cycle, C-rate (charging rate), etc. In general, its efficiency is almost certain to change over time due to the limitations of its chemical mechanism [23,24].

1.2.4 Deployment barriers in demand response participation

In 2021, the U.S. energy information administration (EIA) reported more than 300 DRPs offered by utilities in the US [25]. Most of these DRPs are targeted toward commercial and residential buildings, with some of them are for industrial sectors, and only a few intended for transportation. However, participation in DRPs remains low, especially in residential buildings. For example, in a report, approximately 9.4 million enrolled in DRPs versus 134 million residential buildings in the survey [9]. Although the residential building's participation has

grown with the implementation of various policies and guidelines, some reports still indicate that the level of response from users who enrolled in programs is not as high as expected [26,27]. Several studies have discussed the barriers to DRPs deployment [28–32]. In these studies, the most frequently mentioned barriers to demand response deployment in residential buildings are technology, high cost, operation complexity, and user consumption routine.

1.2.4.1 Technology barriers: hardware

The technology barriers refer to insufficient hardware and software in sensing, computing, and communication in demand side management. In [30], the authors pointed out that the technology barriers are the fundamental barrier due to their importance. The emergency of the Internet of Things (IoT) may serve as a promising hardware solution to overcome the technology barrier. Its rapid development and increasing penetration rate in the past few years have demonstrated its potential as well. As mentioned in Section 1.2.3, there are multiple load-flexible resources and limitations that need to be considered in demand side management, the current hardware development in residential buildings will briefly be introduced in the rest of this section.

In the application of using passive thermal storage in residential buildings, the role of the thermostat is crucial. It not only serves as a temperature sensor but also provides control to the HVAC system based on user-defined setpoints, thereby contributing significantly to the energy management system. The programmable thermostat which allows the user to manually schedule a setpoint at a different time of the day was developed and applied as early as 1906

[33]. Although the programmable thermostat achieves up to 30% energy savings on the HVAC system operation, many users misuse, override, and even abandon the programmable function due to the confusion caused by their lacks of user-friendliness interface and flexibility in schedule feature [34]. In [34], the authors emphasized the need to develop connected thermostats (CTs), also called communicating thermostats. The CTs integrate advanced information and communication technologies to better serve the homeowners in the HVAC system control and build a data-rich solution for future energy management opportunities. Smart thermostats [35,36] emerged in recent years and are a prime example of CT. Another technological innovation in smart thermostats is the use of node sensors [35,36]. The primary intention of developing the node sensors is to enhance the user's thermal comfort: manufacturers want the exact temperature setpoint to be maintained at a certain room when it is detected occupied. Therefore, such node sensors usually include both occupancy sensors (to detect if a specific room is occupied) and temperature sensors (to sense the temperature inside specific room). However, different in the commercial building, there are no terminal boxes to control the amount of cooled or heated air delivered in each zone (room) in residential buildings. Using the node sensor to maintain the temperature in one room may lead to thermal comfort loss in other rooms, or an overall energy waste. Besides, due to the concerns regarding the technology and information security, the occupancy sensor inside the thermostat has no head count feature yet, making it difficult to further improve thermal comfort.

Although not spreading as rapidly as smart thermostats, one of the most common active

thermal storage, smart water heaters—particularly those water heaters that can interact with the grid—are starting to appear. For example, water heaters equipped with CT-2045 protocol, which refers to a physical socket add-on for appliances, are starting to emerge and are being heavily promoted in Washington, Oregon, and California [37]. With CT-2045, the user can directly change the setpoint of the water heater. Moreover, water usage can directly influence a home’s energy consumption, particularly when hot water is used. However, current “smart” capability primarily focus on detecting water leakage [38], with only a few focus on implementing functionality to learn water usage patterns.

Although it is relatively straightforward for users to adjust non-thermal appliances (the appliances that are not thermostatically controlled), such as manually postponing the use time of a dishwasher, there are two primary limitations to this approach. First, manual operations cannot be promptly response to grid signals. Second, manual operation may only allow for simple start/stop operation, rather than optimal control. As such, an optimal intelligent control through energy management system still holds great potential on demand management. On the other hand, the smart features in current smart appliances mainly reply on the monitor rather than control. Given the unique functionalities and cycles of each non-thermal appliances, optimal energy management may require more sophisticated technologies, which leave more room on the development of advanced communication and control protocols.

Distributed energy resources, such as photovoltaic panels (PV), are being increasingly installed in residential buildings. Due to its environmental impact, many states provide policy

support, such as net metering [39]. A recent national lab report shows the PV system's cost in the residential building has dropped from \$7.53/W_{DC} in 2010 to \$2.71/W_{DC} in 2020 [40]. The major price drop in the PV system comes from the hardware cost, but the authors also mentioned that the improvement of module efficiency reduces the number of modules required in a single house, which helps to reduce the installation and labor costs.

Several manufacturers have developed the batteries system that is connected to PV for residential buildings [41–43]. Compared with the lead-acid battery, the sodium sulphur battery, and the vanadium redox battery, it is reported that lithium-ion (Li-ion) batteries are the most promising solution for smart home appliances [22], and by taking trade-off between energy and power density, charging and discharging rates, security, lifecycle, and the cost, the lithium-iron phosphates (LFP/C) are the most popular choice [44,45]. However, electricity storage is still an emerging area, and a battery management system is still needed. Despite the proven benefits of battery energy storage, most batteries are still served as energy backups in residential buildings rather than using them to support daily activities.

1.2.4.2 Technology barriers: software

As DR introduces increased operational complexity in coordination and management of all the relevant devices in residential buildings—a task usually beyond homeowner's—the use of algorithms that are embedded with expert-knowledge has become necessary. A noticeable technical barrier highlighted in [31] is the difficulties for small residential consumers to check price or load modification signals and have a timely and optimal response. In general, one

solution to overcome the challenge operation could be a development of a building management system (BMS) is a centralized software platform that monitors, regulates, and controls electrical and mechanical equipment to reduce cost, save energy, or improve comfort in buildings [46]. Different from the commercial building, which has a relatively mature BMS, residential buildings rarely have an energy management framework in place. From the software perspective, some open-source home BMS such as “Home Assistant” and “openHAB” are promising solutions, which support third-party applications, cloud services, and data visualization. However, those BMSs do not support advanced optimal control yet.

In summary, many developments effort have already started in hardware development, such as smart meters, CTA-2045 protocols, low-cost distributed energy sources, and reliable battery systems, for establishing interactions between the user and the grid. Although some limitations still exist, the progress on hardware offers more opportunities for improving the demand side adjustment. Although many open-source BMSs have been developed for home energy management, they have not been involved in real-time advanced control.

1.2.4.3 Economic and social barriers

Other deployment barriers for DRPs can be classified as economic and social factors. The economic factor (i.e., financial barrier), associated with the cost or benefits that the DRP can bring to homeowners or other stakeholders, is the most important factor since it is the key motivation for users to adopt the DRP. There are various forms of financial barriers: some users are informed of the time-varying price, but not directly informed of the cost reduction [29,31];

some users concern about hidden fees after enrolling in DRPs [28,32]. To make cost savings clearer to those users, and eliminate their worries about hidden fees, in [32], the authors proposed a saving calculator for users in DRP. However, in [29], the authors mentioned that such calculation could be challenging, since the baseline, i.e., the consumption by a customer in the absence of a demand response event, is based on an estimation. In addition to the difficulty of estimating the baseline cost, it is also difficulty for the other stakeholders, who may have an impact on program implementation or further development, to accurately calculate the specific benefits they can gain from the DRP [30]. Therefore, the lack of clear understanding of the cost saving by homeowners and the benefits by these stakeholders may have negative impact on the short-term feasibility and long-term sustainability of DRPs.

Social factor is mainly related to the human responses and interactions with the DRPs, and is of great significance since the operations of appliances (demand side operations) rely solely on the user's intention [28,29,32]. Studies [29,31,32,47] indicate that the complexity on DRPs, whether due to the complicated of devices or operation, can be overwhelming or inconvenient for users, thereby providing a strong barriers for users to participate in DRPs. Moreover, in [28,48], the authors suggests that the bounded rationality [49], where users cannot always make rational decision, is another social factor may decrease the effectiveness of DRPs. This issue is particularly evident in user's unwilling to change their consumption routine, such as charging electric vehicles at night time or using a dishwasher after dinner [50]. Therefore, ensuring respects to user's time preference and simplify the complexity of DRPs are the

important considerations for the successful deployment of demand side management.

1.3 Research Problem

Residential buildings have a large potential on providing grid-service by load shifting and load generation flexibility under Demand Response Programs (DRPs) through proper energy management. Although significant progress has been made in hardware development, resulting in smart devices (including IoT devices) are available and affordable in residential buildings, enables more control opportunities. However, the advanced control for residential building home energy management is still in its early stage. Figure 1.4 summarizes the load-flexible resources in a residential building, highlighting that different appliances have unique operational characteristics. These distinctions, along with the current technological limitations, must be appropriately considered in the demand side management. In addition, beyond the technical barriers, economic and social factors also play an important role in improving the demand response efficiency. In summary, *the knowledge gap is a lack of a systematic framework to establish a grid-interactive home automated energy management framework that can enable advanced optimal control and coordination among multiple flexible resources, provide more insights to different stakeholders, and be applicable for different residential buildings, under DRPs.* More importantly, the HVAC system, which dominates energy usage in residential buildings due to space cooling and heating, is particularly complex. *Developing a framework to conduct the user-driven optimal precooling or preheating and overcome the deployment barrier of the HVAC system is one of the main challenges of this home energy*

management framework.

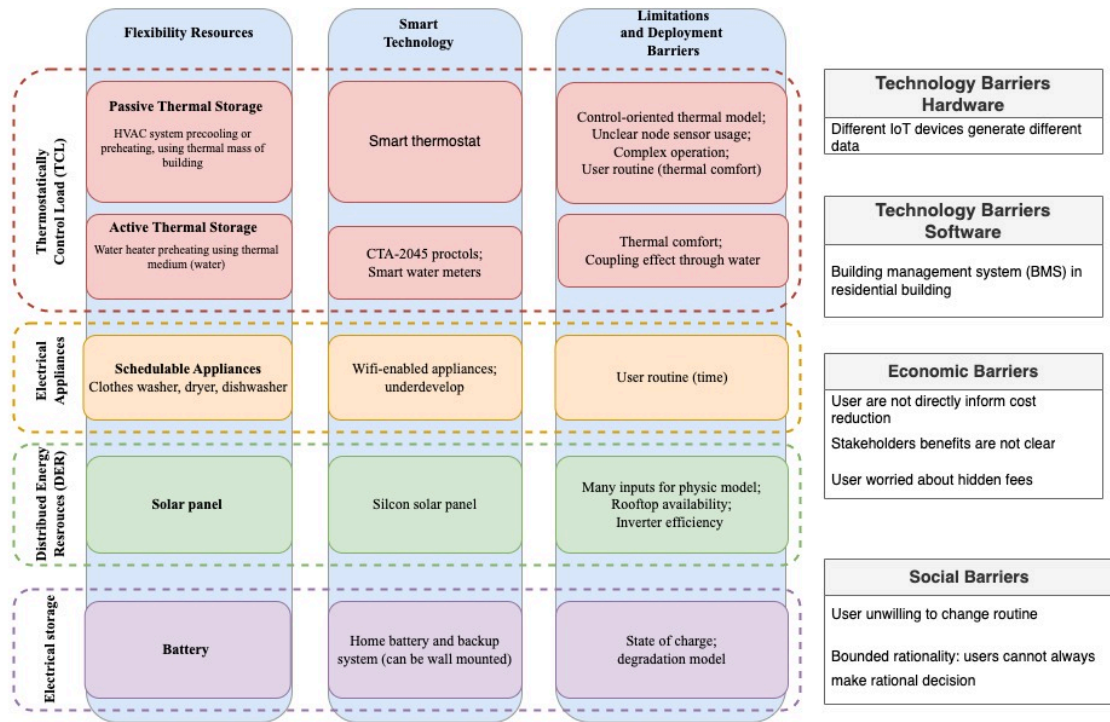


Figure 1.4 Load-flexible resources, technologies, and their deployment limitations in residential buildings

To fill these research gap, the following research questions are raised followed by the hypothesis on how to address each question.

Table 1.2 Research questions and hypothesis.

Primary Research Question:	
How can residential buildings be transformed into energy-efficient, cost-effective, and grid-interactive systems under a DRP by overcoming hardware, software, and social/economical barriers?	
Sub-research Questions:	Hypothesis:

<p>Q1: How can a home energy management system be facilitated in residential buildings in response to grid signals without the need for user's inputs and knowledge?</p>	<p>H1: The implementation of learning-based model in a home energy management framework can facilitate optimal operation of appliances in one and multiple residential buildings without the need for use's input.</p>
<p>Q2: Can the feasible solutions for conducting optimal operation such as optimal precooling in a residential building be implemented through smart thermostat?</p>	<p>H2: The feasible solutions can be achieved by establishing learning-based models for appliances, implementing practical control based on an optimization framework, and developing remote control capabilities for the energy management planform.</p>
<p>Q3: What are the key factors that impact the cost saving and demand flexibility of optimal precooling control under DRPs?</p>	<p>H3: Numerous factors can influence the cost-saving and demand flexibility, therefore, to identify these key impact factors, both simulation study which allowed for controlled variables, and field test study which provide may provide insights into aspects difficult to</p>

	model in simulation are required.
Q4: How can a home energy management framework be designed to incorporate user thermal comfort preference and appliance use time preference when optimally coordinating multiple appliances in residential buildings under DRPs?	H4: To incorporate user preference in the home energy management framework, user thermal comfort and appliance usage time can be designed through understanding and learning user behavior patterns and applying multi-objective optimization framework.
Q5: What are the proper methods to quantify the energy flexibility potential of appliances in residential buildings that provide insights to different stakeholders such as homeowners and energy providers under DRPs?	H5: The energy flexibility potential in residential buildings can be effectively quantified using different key performance indices tailored for the unique interests and needs of different stakeholders, and a single index might not adequately capture the benefits to all stakeholders.

Based on current research and technology development, the following knowledge gap has found:

- 1) Lack of an effective software platform that can enable advance control and operate with minimal user input.

- 2) Lack of low-cost hardware that does not require significant retrofits to conduct the advance control.
- 3) Lack of field demonstration of MPC-based optimal precooling in multiple residential buildings under DRPs.
- 4) No impact analysis for MPC-based optimal precooling based on filed test results.
- 5) Lack of systematic algorithm to incorporate both user thermal comfort and appliance use time comfort when optimal coordinate multiple appliances under DRPs.
- 6) Lack of quantification metric to evaluate the advance control performance that can provide valuable information to various stakeholders involved in demand response programs.

Build upon previous discussion, the research hypothesis for each research questions are in the following. *The primary research hypothesis in this study is that such a home energy management framework can be achieved by using the proposed framework including model development, control implementation, and post-analysis procedure in this dissertation to produce actionable instruction for home appliances control and to provide contextualized insights to different stakeholders.*

1.4 Dissertation Outline

It is evident from the above discussion that there is a need for a home energy management framework that is adaptable to residential buildings, respects users' thermal comfort and time-

use preferences, provides differentiated performance indices for various stakeholders, and carry out in-depth impact analysis. This integrated framework forms the central theme of this study.

The dissertation is divided into six chapters.

- Chapter 1 provide an introduction and gap analysis of this study.
- Chapter 2 describe the literature review.
- Chapter 3 describes the development of different models for the HVAC system, water heater system and the PV power generation.
- Chapter 4 is presenting the field study and simulation study of the MPC-based precooling in residential buildings and the impact analysis results.
- Chapter 5 describes the development of a home energy management multi-objective framework incorporating user preferences.
- Chapter 6 includes conclusions and future work.

2 In-depth Literature Review

2.1 Optimal Control of the HVAC System

The HVAC system operation is one of the most important topic in demand side energy manage, this is because the spacing cooling and heating dominated the energy consumption in household [8,15], the rising market of smart thermostat [9], and the complexity of the HVAC system itself. A report which studies the response of representative 1300 California households to the price signal indicated that the household with spacing heating and air conditioners, responded most [26].

2.1.1 *HVAC precooling strategy and control methods*

As mentioned in Section 1.2.3, adjusting HVAC operation based on price signal, in the most of case refers to the precooling or preheating strategy, which take advantage of building's own thermal mass, (also called passive thermal storage), has also been shown an effective strategy for building load shifting [51–55]. In a precooling strategy, an AC overcools space during off-peak hours to avoid expensive operation during on-peak hours because the building structure serves as passive thermal storage to prevent indoor air increase rapidly. Figure 2.1 demonstrates two precooling strategy examples under DRPs. It can be seen from the figure, although the two strategies choose lower setpoints before the on-peak hours, the timing to precool and the specific setpoint choice are different.

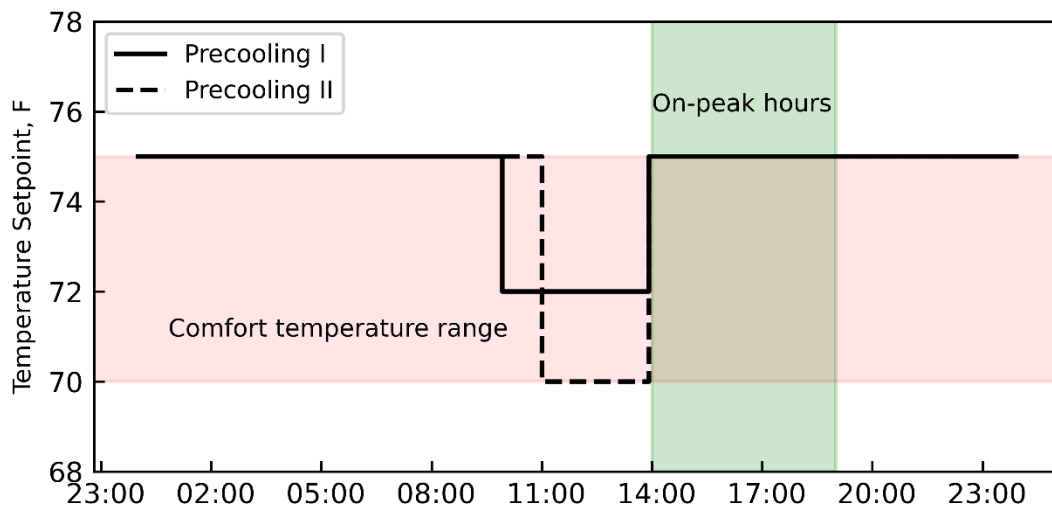


Figure 2.1 Two precooling strategies under DRPs

Previous literature has introduced many control methods for implementing precooling strategies, including rule-based precooling control [51,52,54], optimization pre-cooling control without feedback [53,55,56], reinforcement learning control [57,58], and model predictive control (MPC)-based optimal precooling control [59–66]. In rule-based control, the control signals are based on the pre-defined rule. Other than typical examples of precooling signals shown in Figure 2.1, the pre-defined rule can be flexible in precooling strategies in terms of setpoints and precooling time [54]. The optimization precooling control without feedback means the control signal of the HVAC system is obtained by solving the predictive model and the optimization problem in one shot. In general, the solution to the optimization problem decided the optimal time to start the cooling or heating process of the HVAC. Notice that, most of the optimization solutions for the precooling studies in residential buildings are the on/off signals for the HVAC system, but in some literature, such as [53], the solution is the setpoint

of the AC which is achieved by using the linear approximation of the relationship between the thermal mass characteristics and setpoint. The reinforcement learning is a type of machine learning control strategy where the agent are iteratively train to make sequential decisions in given environment to maximize the cumulative rewards [67]. In the precooling application, the reinforcement learning agent interact with current environments such as indoor air temperature, humidity, occupancy, system status, take sequential actions such as on/off signals or setpoint choices over time, and maximize the cumulative rewards based on cost saving or other objectives [57,58]. The MPC control, whose concepts is shown in Figure 2.2, is based on the optimization control but in iterative way. The MPC predicts future outputs of a system over a *finite receding prediction horizon* through a system model, determines control signals through optimization over a *control horizon* based on the predictions, time-variant inputs, and system feedback, and repeats the predication and optimization for every iteration [68,69].

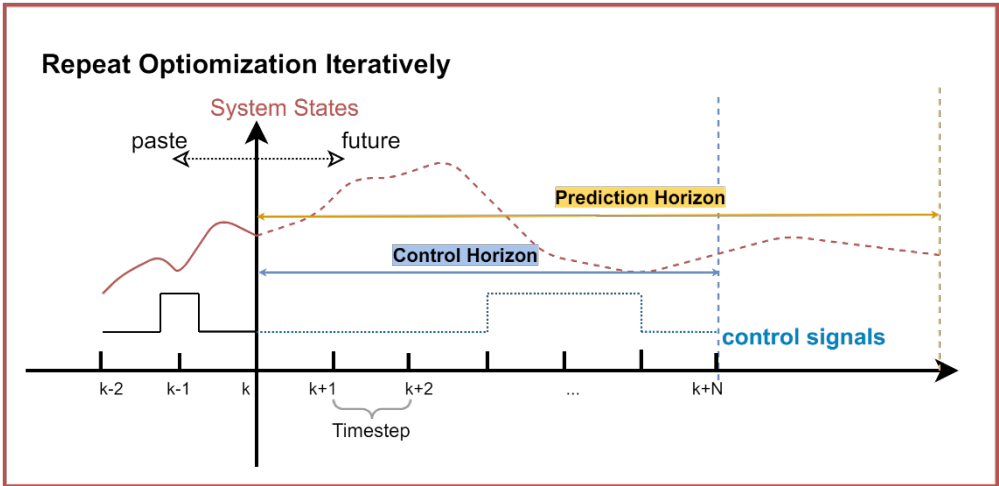


Figure 2.2 Prediction horizon, control horizon and timestep for MPC

The rule-based control and optimization control without feedback cannot incorporate

complicated, uncertain dynamic, and time-varying disturbances in the HVAC system into decision-making. Furthermore, reinforcement learning requires a large number of training scenarios to make optimal decisions [70]. MPC has been proposed to solve many HVAC system control problems already [66,71,72]. Compared with PID controller [73], rule-based controller [74], and reinforcement learning [70], MPC controller shows the more robust and stable operation and meanwhile achieves better energy and comfort performance in the HVAC system applications. In summary, MPC provides a more efficient control solution for optimal precooling [75,76]. However, as the name implies, the predictive model is the core for the MPC control. In building operation, the model used in MPC must be able to capture the thermal dynamics of the building and therefore is often referred to as a thermal model. Therefore, an accurate and efficient thermal model, along with a computationally economical optimization framework, is essential to run MPC for control purposes.

Table 2.1 Advantage and disadvantage of major control strategies on implementing precooling

Precooling Control	Advantages	Disadvantages
Rule-based	Simple; Easy to implement;	Hard to determine the specific timing and setpoint for precooling; Cannot incorporate with outside disturbance

Optimization no feedback	Optimal or near optimal solution	Cannot incorporate with outside disturbance. Rely on prediction model;
Reinforcement learning	Learning by interaction rather than model or data; Can explore the complicated solutions;	Long time interaction needed to explore the environment; Retraining needed if environment change
Model Predictive Control	Robust and stable optimal solution; Feedback control;	Rely on prediction model;

2.1.2 Current MPC-based precooling controller issues

Although an MPC controller can be a promising solution to implement precooling, its implementation and assessment are still difficult. First of all, quantification the effectiveness of the MPC controller is naturally challenging. Due to the fluctuation and non-recurring external disturbance, including the weather conditions, quantifying the pre- and post-implementation of MPC controller is hard. As mentioned in Section 1.2.1, when investigating the DRPs benefits, the cost reduction or other benefits are hard to evaluate, due to the hardness of quantify the baseline consumption. It is impossible to repeat the same external disturbance,

especially for the outside weather disturbance, hence it is very difficult to evaluate what would have occurred [77]. Additionally, although retail-level DRPs is entirely depending on demand-side operation, many stakeholders such as homeowners, grid operator, and utility companies, can all benefit from it, which raise the higher standards for MPC precooling result quantification [78]. However, majority of studies only evaluate the effectiveness of MPC by using the value of objectives .

The performance of MPC controller can be affected by many different factors, making it even harder to conduct the impact analysis. Many impacts analysis studies of MPC controller are based on simulation studies. In [79], the impacts of weather conditions, home thermal conditions, prediction horizon, price structure, and rated cooling capacity on the MPC controller were investigated using MATLAB simulation by assuming the model had perfect temperature predictions. In [62], MATLAB simulations were performed to investigate the impacts of different TOU structures and tradeoff coefficients between comfort and cost. In [70], the authors conducted MPC study through co-simulations in EnergyPlus based on 500 buildings and discussed the impact of climate, floor size, and building age on energy saving and thermal comfort. Since MPC controllers require the feedback of the system to make future operation decisions, the impact analyses in the simulation studies used the one provided by virtual testbeds, such as EnergyPlus [70], or TRNSYS [72], which do not represent real building dynamics precisely.

Even though there are a few studies found implemented the MPC precooling controller in

field tests [59,60,80], they have limitations to some extent. A laboratory building mimicking a residential building was tested to evaluate the load-shifting potential with a geothermal heat pump and water tank [62], where the MPC controller signal was obtained from simulation and then implemented in the laboratory building. Although the MPC controller did not use actual feedback from the system, it enabled load shifting and saved 25%-35% in electricity cost. In [63], two different economic MPCs, i.e. operation costs were included in the MPC optimization objective function, were implemented for heating tests in a residential building, where the heat pump and the PV panel in the tests were installed virtually. Compared with the simulated rule-based controller, the MPC controllers were proven to obtain 3-7% cost savings and increase energy flexibility. Although MPC controllers are effective for cost savings and load shifting, the previous studies reported additional sensors or controllers were needed, adding additional cost and installation effort. Only one study was found in multiple residential buildings for impact analysis, where an MPC controller was implemented for improving heating efficiency [81]. The authors adopted a neural network approach to construct the thermal model for space heating in ten different residential buildings in Switzerland and Germany. In this study, a NiQ system, which includes a controller and a set of sensors, was custom-made and installed in each home to performance the MPC operation. Given the limitation presented by experiments, only the effect of outdoor air temperature impact on energy reduction was investigated by periodically alternating between the MPC controller and reference controller under similar conditions and filtering out abnormal days. About 1.8kWh extra daily energy was concluded

for each degree of daily average outdoor air temperature reduction in one single-family house in Brugg. The impact analysis of MPC-based precooling controllers in residential buildings is limited due to the difficulties of implementing it considering high computational demand of the MPC and lack of advanced control system in residential buildings.

2.1.3 Possible solutions for MPC-based precooling controller

Widespread use of the Internet of Things (IoT), including smart thermostats, provides an opportunity for the implementation of advanced control without prohibitive expenses. In [82], the authors reviewed the application of IoT devices in smart homes and pointed out IoT is an intelligent solution for small residential. Therefore, to investigate the impact of the aforementioned factors on an MPC-based precooling controller, an implementable MPC-based precooling agent through the smart thermostat is developed in this study. The development of such MPC control agent requires an efficient thermal model and optimization approach. A grey-box model approach with an efficient parameter identification approach is proposed. The grey-box model is formulated using system knowledge and therefore has a greater chance to generate reliable predictions even if the operational condition for prediction is beyond the condition of the training data or the training dataset is limited. In addition, the grey-box model is usually a model based on differential equations with a simpler model structure, which allows the mixed-integer linear programming (MILP) technique [83] to solve the optimization without the use of computationally expensive exhaustive optimal search methods.

The complicated quantification of benefits of MPC controller due to the complicated

environment including the rapid outside weather changes and the variation of the stakeholder can be solved by using weather clustering algorithm and multiple key performance indicators (KPIs). The weather condition is a complicated factor due to its inherent variability and higher dimension. Using a single point, such as daily average outdoor air temperature, as a factor affecting AC energy consumption is too simple while using all weather conditions can result in over complex analysis. Weather clustering, which merges similar weather conditions into one variable, is required in the result analysis. Additionally, optimal precooling is sensitive to the time of the day, making many available clustering algorithms focusing on values rather than similar **shapes** of subsequences of time series data not appropriate. Therefore, two matrix-profile based algorithms are modified and combined for multi-dimensional (including temperature, humidity, and solar irradiation) weather pattern clustering. To further investigate which homes or operation conditions provide more energy flexibility with the proposed MPC agent, in addition to the cost reduction, a flexibility factor that quantified the load shifting ability is employed in impact analysis across multiple homes.

2.2 Smart Home Energy Management (Home-EMS) System

In a smart home energy management (Home-EMS) system, thermal appliances often account for most of the building's electricity consumption; additionally, they have a direct impact on human comfort. It is critical to incorporate them into the Home-EMS. In [84–86], the researchers implemented the optimal HVAC operations derived by several control algorithms. Their findings demonstrate that a 9%-20% cost savings under the time of use (TOU)

rate can be realized while maintaining a reasonable room air temperature bound when compared to baseline cases. However, these studies mainly focused on HVAC systems and ignored consideration of the smart home environment. Integrating thermal and non-thermal appliances is critical for Home-EMS because one appliance's electricity consumption limits the energy consumption of another, and thermal appliances cannot be turned off for a long time without compromising thermal comfort. Hence, their interaction should be considered together. [87–89] integrated the HVAC system with several appliances and considered the dynamic model of HVAC system. However, the studies cited above did not consider the process of system identification before to using the HVAC model. Unlike washing machines and dryers, the energy consumption of the HVAC system is determined by the house's thermal dynamic, which is influenced by a variety of factors, including the house's thermal characteristics and the outside weather. The system should be identified in advance so that the optimal control algorithm can be automated across different buildings subjected to varying weather conditions. All the studies cited previously do not incorporate photovoltaic and battery storage into the Home-EMS.

[11]–[15] included solar panels and a variety of household appliances into their scheduling model. In [90], the author evaluated a photovoltaic (PV)-battery system integrated with an electric vehicle (EV) and an electric heater, and addressed the smart house optimization problem using a genetic algorithm (GA) called the shuffling frog jumping method (SFLA). In [91], the researchers classified the load on home appliances into three categories: flexible

interruptible load, flexible uninterruptible load, and rigid load, and used the GA to solve the optimization problem and minimize the utility bill. In [92], the optimization problem has two objectives: minimize the peak load, and minimize the utility bill, the problem was solved as MIP by using YALMIP. In [93], the researchers considered a home energy generation including a solar panel, a wind turbine, and a battery system. The smart home scheduling problem are solved by modified NSGA II in this study. The researchers in [95] pointed out that the study in [93] did not take into account the uncertainty associated with the load and the coordination of numerous residences. The study considered the multiple houses that share the PV generation. While [11]–[15] could be practical to the SHEM but none of them considered the dynamic of the HVAC system. Table 2.2 is the summary of the features of the scheduling models for aforementioned studies. The notation "X" indicates that the researcher considered correspondence appliances when conducting their study. We observed from the "Method" columns that, although there is debate about the efficiency of GA in solving these types of optimization problems, some genetic algorithms are still popular, particularly for the GA that involving multiple objectives. Multiple words with different arrows in the "Objective" column indicate that their model incorporates a multi-objective framework. Although the author considered both homeowner's thermal comfort and the utility bill, in [86], the author was simply combining two objectives together and thus treating it as a single objective problem.

Table 2.2 Summary of SHEM Scheduling Models

References	P	Battery	HVAC	Water Heater	Non-thermal Appliances	HVAC Dynamic	Method	Objective
[96]					X		MIP	Cost
[85]			X				PSO	↓Cost, ↑Comfort
[86]			X				GA	↓ (Cost – Comfort)
[97]					X		NSGA II	↓Cost, ↑Comfort
[93]	X	X	X	X	X		Modified NSGA II	↓Cost, ↑Comfort
[98]				X	X		Compromise Programming	↓Cost, ↑Comfort, ↓Inconvenience
[88]			X	X	X	X	MIP	↓Cost
[87]			X		X	X(ARX)	MIP	↓Peak Load
[89]			X		X	X	MIP	↓Cost
[90]	X	X		X	X(EV)		GA	↓Cost
[95]	X	X		X	X		(Stochastics) MIP	↓Cost
[91]	X	X	X	X	X		GA	↓Cost
[92]	X	X	X	X	X		MIP	↓Peak Load, ↓Cost
Legend: ↓: Minimize, ↑: Maximize, MIP: Mixed Integer Programming, GA: Genetic Algorithm, PSO: Particle Swarm Optimization, ARX:								

Although several studies have focused on the Home-EMS, many of them either focus on individual components such as PV or HVAC systems. In contrast, this study integrates various components including PV, battery, the HVAC system, the water heater, and the flexible load such as clothes washer. Furthermore, while the trade-offs between cost and comfort have been discussed in previous works, these studies often do not sufficiently take into account realistic

user behavior. Particularly, user may wish to reduce their utility bill through the energy management system but prefer to do so without changing their existing usage patterns. This includes the user might have preferences on appliance use time, water use, and the preferred setpoint. The proposed energy management framework takes into account these user preferences and offers a more realistic simulation scenario.

3 Load-flexible Resources Modeling in Residential Buildings

3.1 Introduction

In predictive control and scheduling algorithms, the model, which is used to predict the system behavior, directly determines the effectiveness and robustness of final decisions, and therefore improving its accuracy is a focal point in this research. In general, there are four aspects contributes to establish an accurate model, including selecting an appropriate model type, attaining a thorough understanding of the energy and information transmission in the system, using an effective training method, and collecting high-quality training data. However, it is less economical or infeasible to achieve perfection in all four aspects. This chapter describes complete modeling procedures of load-flexible resources including the HVAC system, water heater system, and PV. In the following sections, we first conduct the four aspects mentioned above, while considering some barriers in current states, and finally validating proposed models in real-world operation data obtained from a Lab House and open datasets.

3.2 Model Type Selection for Load-flexible resources in a Home

A model, which is a mathematical representation to describe the transfer of mass, energy, or information between systems, in general includes tree main branches: a white-box model, a black-box model, and a gray-box model [99]. The white-box model is typically has known model structure and known parameters, the black-box model, which aims to build a general function governing the input-output behavior of the system, usually has no clear model

structure and unknown parameters, the gray-box model, which lies between white-box and black-box model, usually has known model structure and unknown parameters. The gray-box model, depending on the assumption of model structures, can be further divided into the light gray-box and dark-gray box model [99]. In [99], dark-gray box model refers to the neuron/fuzzy model, where the expert knowledge is required in model training process, making the model development is limited by human thinking. Therefore, in this study, only light gray-box model will be discussed. Table 3.1 shows the model characteristic and their applications in this study.

Table 3.1 Model characteristics [99] and their application in this study

Model	Characteristics	Examples	Application in this study
White-box model	Physical laws known; Parameters known	Linear/non-linear differential equations	Electrical appliances; Battery
Light-gray box	Physical laws known; Parameters unknown	Differential equations with known parameters; polynomial regression	Passive and active thermal storage
Black-box	Input/output governing function; Assumption of a model structure	Deep learning models; Ensemble models	PV panel generation

Light gray-box models are used in this study to describe the thermal dynamics of homes which related to the HVAC system and the thermal dynamics inside a tank which related to the water heater. There are three motivations for developing light gray-box model to describe the dynamics of thermostatically resources. First, a gray-box model for HVAC applications and water heater applications typically requires less effort to learn than a white-box model [100]. Second, because it is partially physics-based, a gray-box model also generally yields more accurate prediction than a black-box model, especially in operating regimes where they are not trained [65,101]. This feature can be very useful in developing a model for control purposes since the control operation usually is different from the operations in the training dataset. Finally, the light gray-box model describes the physical process rather than general input-output relationship, making the model adjustment easier when additional processes need to be included (e.g., coupling effect between the water heater and clothes washer occur).

While the light gray-box model has many advantages, it is not suitable for modeling every load-flexible resources in homes. The parameter identification process makes the light gray-box are not suitable for modeling the electrical appliances or battery. These on/off or (dis-)charging behaviors do not rely on complicated thermal dynamics, but rather on their own specifications, which can be easily found from information provided by the users or manufactures. Therefore, in this study, the white-box model is used to model electrical appliances and batteries. Besides, the light gray-box model is often simplified or generalized model of physical reality, and its capability is limited in a system that is inherently variable,

such as a distributed PV panel system. In contrast, the black-box model is often more suitable in handling these complicated input output relationships. In the following section, the HVAC system, the water heater system, and the PV panel generation are introduced due to their modeling development complexity. On the other hand, the characteristics of battery and electrical appliance are direct inputs to the optimization framework in Chapter 5 due to their white-box model are relatively simple and straightforward.

3.3 HVAC Modeling and Parameter Identification Approach

3.3.1 Reduced-order RC-based gray-box model of home thermal dynamics

The home thermal dynamics involves several heat and mass transfer processes, including heat transfer through opaque building envelope, solar radiation through windows (or other transparent elements in homes), infiltration through cracks of windows or doors, and internal heat gains caused by people, lights, and equipment. A RC network model, which mainly models the thermal energy flow driven by temperature difference (thermal resistance), and the ability of a building to store energy (thermal capacitance), is a widely used simplified model to describe these heat and mass transfer processes in building [102]. In this paper, a second-order resistance-capacitance (RC)-based gray-box model is used to describe the thermal dynamics of a home, as shown in Figure 3.1.

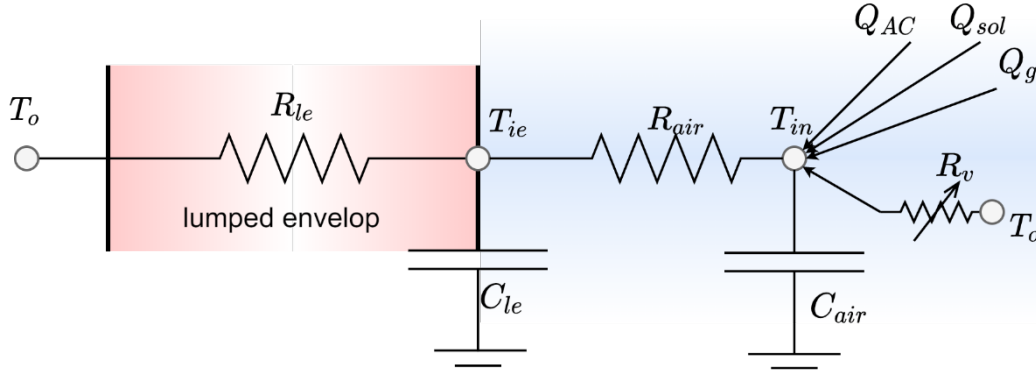


Figure 3.1 Schematic of the RC network of home thermal dynamics

Based on Figure 3.1, the second-order RC-based model is given by:

$$C_{le} \frac{dT_{ie}}{dt} = \frac{(T_o - T_{ie})}{R_{le}} + \frac{(T_{in} - T_{ie})}{R_{air}} \quad (1)$$

$$C_{air} \frac{dT_{in}}{dt} = \frac{(T_{ie} - T_{in})}{R_{air}} + Q_{sol} + \frac{(T_o - T_{in})}{R_v} + Q_{AC} + Q_g, \quad (2)$$

where t denotes time; T_{ie} , T_{in} , and T_o are the interior wall surface temperature, indoor air temperature, and outdoor air temperature, respectively; R_{le} and C_{le} are the thermal resistance and thermal capacitance of the lumped envelope, which includes but not limited to external wall, roof, ceiling, floor, interior wall, window and outside air of the building; R_{air} and C_{air} are the thermal resistance and thermal capacitance of the air; Q_{sol} represents the heat transfer rate introduced by solar irradiation acting on the indoor space; R_v represents the variable convective thermal resistance associated with the infiltration of building, whose value depends on cracks of the building, wind speed, and outside air temperature; Q_{AC}

represents the cooling output of the AC; Q_g is the internal heat gain contributed by people, lights, and equipment

The second-order RC-based model is also called 2R2C model. RC networks of various orders have been widely used to model building thermal dynamics (e.g., 1R1C in [58], 3R2C plus 2R2C in [103], and 6R2C model [104]). For residential buildings, it is believed that 2R2C model has the best trade-off between model accuracy and implementation cost: a first-order model, although simpler, would not be able to effectively separate the indoor air temperature dynamics (that usually respond faster to external changes) from the interior wall surface temperature dynamics (that usually respond slower). In contrast, a higher-order model, although could be potentially more accurate, would require more sensors to measure its states.

3.3.2 2R2C parameter identification approach

To clearly express the relationship between home thermal dynamics, weather conditions and AC on/off signal, and make the parameter identification easier, (1) and (2) are converted into:

$$\frac{dT_{ie}}{dt} = \frac{1}{\tau_1}(T_o - T_{ie}) + \frac{1}{\tau_2}(T_{in} - T_{ie}) \quad (3)$$

$$\begin{aligned} \frac{dT_{in}}{dt} = & \frac{1}{\tau_3}(T_{ie} - T_{in}) + \frac{1}{\tau_3}(a_1S + a_2S^2) + \frac{1}{\tau_3}(T_o - T_{in})(b_1W + b_2W^2) \\ & + \frac{1}{\tau_3}(c_1T_o + c_2T_o^2)u, \end{aligned} \quad (4)$$

where S is the solar irradiation, W is the wind speed, and u is the AC on/off signal.

Equation (3) and (4) contain a total of nine unknown parameters: τ_1 is the time constant of

the building envelope; τ_2 is the coefficient associated with the thermal resistance of air and thermal capacitance of the building envelope; τ_3 is the time constant of the indoor air; $a_1, a_2, b_1, b_2, c_1, c_2$ are coefficients representing the impact of solar irradiation, wind speed, and AC output, respectively. In (4), we only use the outdoor air temperature as the dependent variable of the output of the AC system (containing one outdoor unit and one indoor unit), i.e., $\frac{1}{\tau_3}(c_1 T_o + c_2 T_o^2)u$. Although the indoor air temperature, which can be considered as the evaporator inlet temperature, may also affect the AC output, we neglect the impact of the wet bulb temperature of indoor air due to its smaller variation compared with the outdoor air temperature and the willing to keep Equation (4) linear.

To identify the unknown parameters in (3) and (4), we first discretize these two equations and then consider the optimization problem with the following objective function:

$$\min_P = \sum_{k=1}^{N_{\text{train}}} (T_{in}^k - T_{in,mes}^k)^2 + (T_{ie}^k - T_{ie,mes}^k)^2 \quad (5)$$

where $P = \{\tau_1, \tau_2, \tau_3, a_1, a_2, a_3, b_1, b_2, c_1, c_2\}$, T_{in}^k and T_{ie}^k represent the indoor air temperature and the interior wall surface temperature at time step k , and $T_{in,mes}^k$ and $T_{ie,mes}^k$ are their measurements, N_{train} represents the total number of time steps in the training dataset.

The constraints of the optimization problem are:

$$T_{ie}^k = \frac{\Delta t}{\tau_1} (T_o^{k-1} - T_{ie}^{k-1}) + \frac{\Delta t}{\tau_2} (T_{in}^{k-1} - T_{ie}^{k-1}) + T_{ie}^{k-1} \text{ for } k = 2, \dots, N_{\text{train}} \quad (6)$$

$$T_{in}^k = \frac{\Delta t}{\tau_3} (T_{ie}^{k-1} - T_{ie}^{k-1}) + \frac{\Delta t}{\tau_3} (a_1 S^{k-1} + a_2 S^{k-1^2}) + \frac{\Delta t}{\tau_3} (T_o^{k-1} - T_{in}^{k-1}) (b_1 W^{k-1} + b_2 W^{k-1^2}) \quad (7)$$

$$+ \frac{\Delta t}{\tau_3} (c_1 T_o^{k-1} + a_2 T_o^{k-1^2}) u + T_{ie}^{k-1} \text{ for } k = 2, \dots, N_{\text{train}}$$

$$a_1 \geq 0 \quad (8)$$

$$a_2 \geq -0.5a_1 \quad (9)$$

$$c_2 \geq 0 \quad (10)$$

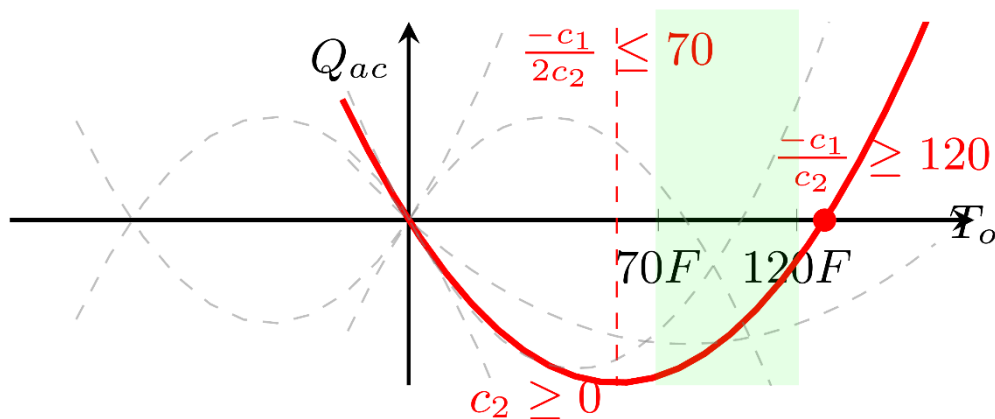
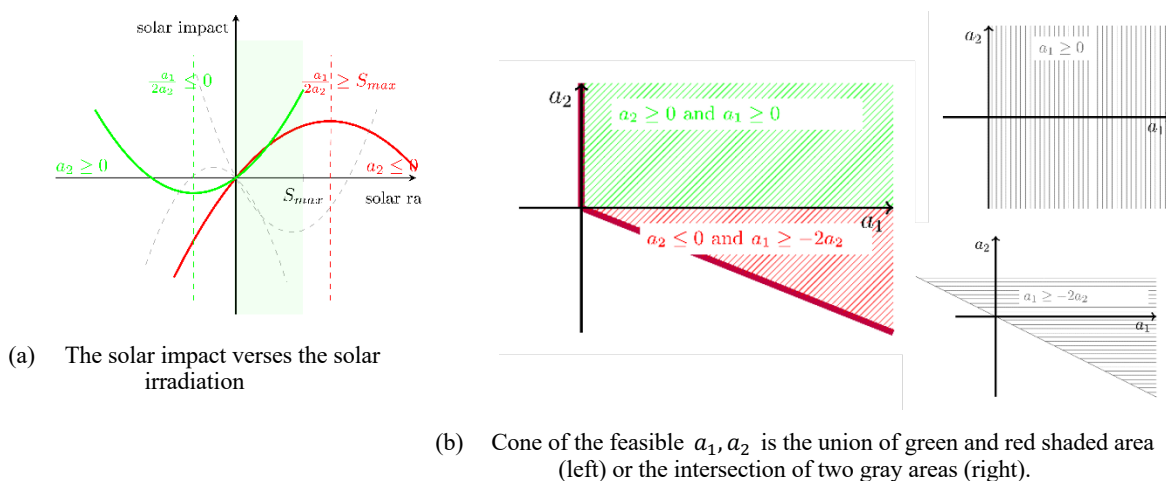
$$c_1 \leq -120c_2 \quad (11)$$

$$c_1 \geq -140c_2. \quad (12)$$

Here, (6) and (7) are obtained from (3) and (4) by discretization using the forward Euler method. (8)-(12) are added to ensure the identified parameters are realistic, which we now explain in detail.

The Constraints (8) and (9) are motivated by the fact that the solar impact $\frac{1}{\tau_3} (a_1 S + a_2 S^2)$ to temperature dynamics should always be a positive, increasing function in solar irradiation S , from 0 to S_{max} (maximum solar irradiation). Among all parabolas defined by $\frac{1}{\tau_3} (a_1 S + a_2 S^2)$, which are demonstrated in Figure 3.2 (a), only the red or green solid curve satisfies this fact. We then have inequalities $a_2 \leq 0$ and $\frac{-a_1}{2a_2} \leq 0$ (green curve), or $a_2 \geq 0$ and $\frac{-a_1}{2a_2} \geq 1$ (red curve since solar irradiation is normalized). This leads to constraints $(a_2 \geq 0 \wedge a_1 \geq 0) \vee (a_2 \leq 0 \wedge a_1 \geq -2a_2)$, where \wedge meaning "and", \vee meaning "or."

However, a logical "or" constraint may require introducing new variables in the optimization problem formulation. Fortunately, a reformulation is possible in this case. As shown in Figure 3.2(b), the union of green and red areas (associated with the green and red curves) is also the intersection of two gray-shaded areas on the right ($a_1 \geq 0$ and $a_1 \geq -2a_2$), which leads to Constraint (8) and (9).



(c) Possible shape of the quadratic equation for AC output.
Figure 3.2 Physics constraints explained for system identification process.

The Constraints (10)-(12) are motivated by the fact that the AC output impact $\frac{1}{\tau_3}(c_1 T_o + c_2 T_o^2)u$ to the temperature dynamics should always be a negative, increasing

function in outdoor air temperature T_o from 70°F to 120°F (light green shaded area), the typical summer outdoor air temperature condition. This is because once the outdoor air, which can be considered as the air entering the condenser, gets higher, the condensation temperature gets higher, reducing the AC output. As shown in Figure 3.2(c), the red curve satisfies this fact, which leads to Constraints (10)-(12).

3.3.3 *HVAC system power estimation*

In Section 3.3.2, we have discussed the modeling of the thermal dynamics of homes, i.e., the model describes the heat transfer between the outside environment, building envelope, and indoor space. Although modeling such dynamics is a crucial part of understanding the behavior of the HVAC system, we also need to consider other parts of the HVAC system—its power estimation. In this section, we will discuss the modeling of the HVAC system power consumption, which is an important procedure for achieving minimum energy or cost in optimal control.

In general, a common AC system in a residential building consists of an outdoor unit and an indoor unit, where the power of the indoor (P_{indoor}) unit is relatively constant and the power of the outdoor unit (P_{outdoor}) may be affected by two factors: the wet bulb temperature of the air entering the evaporator, and the temperature of the air entering the condenser. If the outdoor unit is placed in the outdoor and the indoor unit is placed inside the building, the indoor air can be considered as the air entering the evaporator and the outdoor air can be considered as the air entering the condenser. Therefore, P_{outdoor} is a function in $T_{\text{in,wet}}$ (indoor air wet bulb

temperature) and T_o (outdoor air temperature). However, same as the reason for only use the outdoor air as the dependent variable of the AC output to the temperature dynamics, we neglect the impact of $T_{in,wet}$. In this study, we conduct third-order polynomial regression to estimate the outdoor unit power and constant value for indoor air unit power:

$$P_{HVAC} = A_1 T_o + A_2 T_o^2 + A_3 T_o^3 + A_4 + P_{indoor} \quad (13)$$

3.4 Water Heater Modeling

3.4.1 Reduced-order RC-based gray-box model of water heater thermal dynamics

Besides electricity, water is another important medium that connects energy consumption in a residential building. For example, a clothes washer withdraws hot water from a hot water tank; a hot water tank releases heat to the indoor space, air temperature of which is determined by the HVAC system. Therefore, a model that can provide accurate prediction and capable to be easily modified to adjust to different physical scenarios (e.g., coupling effects), such as the RC network model used in this study, is a popular choice for water heater modeling [105–107].

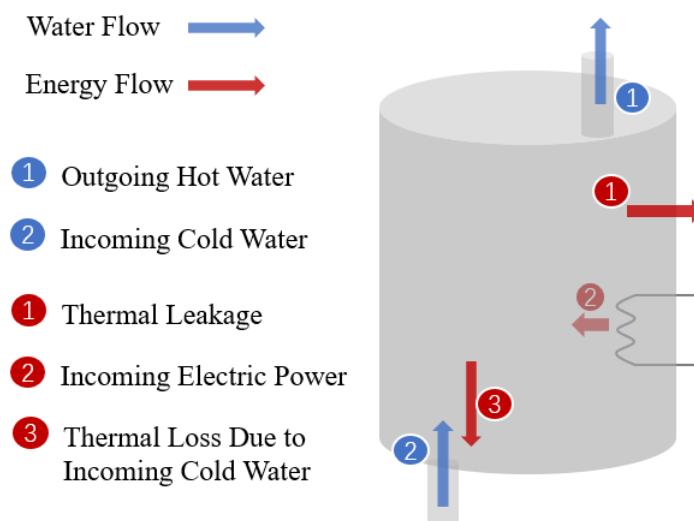


Figure 3.3 Water heater heat and water flow

Figure 3.3 demonstrates energy and water transfer from the water tank to the environment. Additionally, it is assumed the water heater is exposed to the indoor air temperature, and the changing indoor air temperature T_{in} influences the thermal leakage of the hot water tank. Let's use t denotes time, the first-order RC-based gray-box water heater model is as following:

$$C_{wh} \frac{dT_{wh}}{dt} = \frac{1}{R_{wh}} (T_{room} - T_{wh}) + \rho C_p W (T_{inlet} - T_{wh}) + Q_{wh}, \quad (14)$$

Where T_{wh} is the temperature of the water inside the tank; R_{wh} is the thermal resistance of the water heater (equals to R/A where R refers to a common used measure for thermal resistance of water heater, called R-value, and A refers to the water tank surface area); C_{wh} is the thermal capacitance of the water heater, which equals to $\rho C_p V$, and ρ, C_p, V refers to the density of water, the specific heat of water, the tank volume, respectively; In this study, we assume that once the hot water is withdrawn, cold water with temperature T_{inlet} will enter the tank, and here we assume that T_{room} represents the indoor air temperature, which also equals T_{in} for the HVAC system; Q_{wh} is the power consumption of the water heater.

3.4.2 Parameter identification for water heater model

Since the water heater system has a simpler heat transfer process (i.e., it is first-order and has less variables) than the HVAC system, a Least-square method is used in the training process.

Before introduced the formulation of Least-sqaure method, let's k denotes time step, and the first-order RC-based gray-box water heater model is given by:

$$\frac{T_{wh}^k - T_{wh}^{k-1}}{\Delta t} = \frac{1}{\alpha} (T_{in}^{k-1} - T_{wh}^{k-1}) + \frac{1}{\beta} \rho C_p W (T_{inlet}^{k-1} - T_{wh}^{k-1}) + \frac{1}{\beta} Q_{wh}^{k-1} \text{ for } k = 2, \dots, N, \quad (15)$$

where α refers to an unknown parameter equals to $C_{wh}R_{wh}$, β refers to an unknown parameter equals to C_{wh} , and N refers to the number of training data points. Equation (15) can be rewritten into a form of $X\beta = Y$:

$$\begin{bmatrix} T_{in}^1 - T_{wh}^1 & \rho C_p W^1 + Q_{wh}^1 \\ T_{in}^2 - T_{wh}^2 & \rho C_p W^2 + Q_{wh}^2 \\ \vdots & \vdots \\ T_{in}^{N-1} - T_{wh}^{N-1} & \rho C_p W^{N-1} + Q_{wh}^{N-1} \end{bmatrix} \begin{bmatrix} \frac{1}{\alpha} \\ \frac{1}{\beta} \end{bmatrix} = \begin{bmatrix} \frac{T_{wh}^2 - T_{wh}^1}{\Delta t} \\ \frac{T_{wh}^3 - T_{wh}^2}{\Delta t} \\ \vdots \\ \frac{T_{wh}^N - T_{wh}^{N-1}}{\Delta t} \end{bmatrix} \quad (16)$$

where the Least square solution is $(X^T X)^{-1} X^T Y$.

3.5 Short-term PV Power Generation Prediction

In this section, the methodology of the ensemble framework to forecast short-term PV power generation is discussed. One of the most popular implementations of Gradient Boosting Decision Tree, LighGBM, is adopted to forecast PV power generation. Compared to other machine learning or deep neural network models, LightGBM uses less memory and usually has a higher training efficiency [108]. Figure 3.4 shows the schematic outline of the prediction framework. First, the residential PV generation and the metrological dataset are collected from different data sources (data sources will be introduced in Section 3.6). Then, original data are preprocessed, and features are calculated. In this process, a training and a test dataset are carefully split. Next, the SHAP value then used to conduct a feature selection. Further, a grid-search method and a k-fold validation are employed to optimize the hyperparameters. Finally, the tuned model will be evaluated using the test dataset.

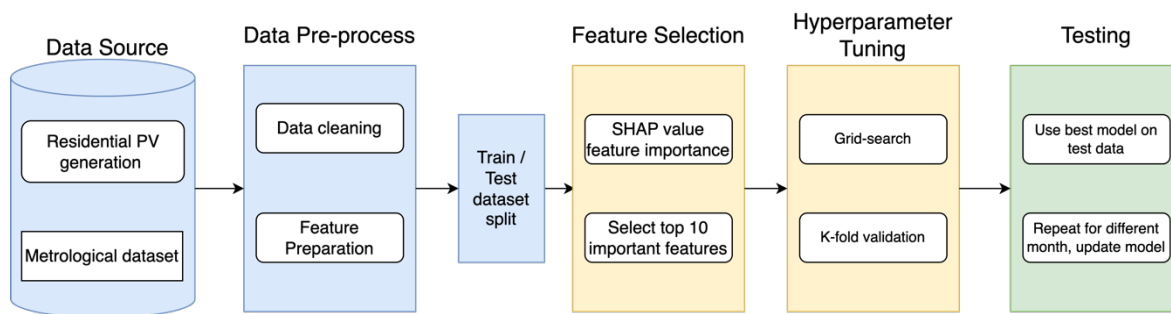


Figure 3.4 The schematic diagram of PV power generation prediction

3.5.1 Data preprocessing

In this study, six-month datasets of a residential rooftop PV power generation and the historical metrological dataset are obtained from the data sources. To make the proposed automation system work as soon as possible after installation, a random continuous three days of each month are selected as the test dataset, which mimics any time of the year that the automation system is installed, whereas the previous two-month data of this randomly selected days are used for training and validating the developed model. Weather forecast is the only input of this model, and the output of this model is the PV power generation in Watt.

Since the PV power generation dataset is collected from a real system, it is necessary to preprocess some abnormal data that may be caused by system or sensor faults before training the model. First, the days with no observable solar but still has PV power generation are removed (the rare case that the PV panel still have small currents generated through thermal effects or other mechanism are neglected). Afterward, the days with almost constant PV power generation throughout the day are filtered out, utilizing the fact that very small standard

deviations in PV power generation within a whole day of such days may occur (i.e., “no change” in generation throughout the day under normal solar radiation).

3.5.2 *Feature selection*

The original meteorological dataset (input dataset) includes a total of ten different parameters such as air temperature, global horizontal irradiance (Ghi) and wind speed. In addition, time information includes hours, months, years and seasons of the forecasted days are calculated based on the given timestamps and are added into input dataset as well. Among these meteorological parameters, wind direction in units of degrees may not be a good model input. For example, the model may treat 0° and 360° differently but they represents the same meaning in reality. To make the model easier to understand this input dataset, the wind vectors W_x and W_y are calculated by using the following equations:

$$W_x(t) = W(t) \cos(\theta(t)) \quad (17)$$

$$W_y(t) = W(t) \sin(\theta(t)) \quad (18)$$

where $\theta(t)$ and $W(t)$ represents the wind degree and wind speed at time t .

The train dataset mentioned in Section 3.5.1 is then divided into the first 80% as training set and the last 20% as validation dataset for feature selection purpose. Since the training dataset is in less than a two-month period, it is necessary to eliminate irrelevant features to prevent overfitting issue. A state-of-the-art feature explanation approach called Shapley Additive exPlanations (SHAP), which is based on cooperative game theory and is first proposed by Lundberg and Lee [108], is utilized in the feature selection stage. The SHAP

explains the contribution of each feature to the prediction *in each data point*. Since SHAP conducts the calculation based on the model, it is less sensitive to input features, even if they are correlated with each other (e.g., air temperature and hour of the day in this case). The details information and of SHAP and its application on energy estimation can be found in [108–111], respectively. Finally, the features corresponding with top ten mean absolute SHAP values are selected.

3.5.3 *Hyperparameter tuning*

To investigate the best lightGBM model, the grid search method is used to find out best hyperparameter combination and K-fold cross validation is used to validate the model performance. The hyperparameters tested in the grid search method are `n_estimator` represents the number of boosting iterations (i.e., number of trees) that will be performed; `max_depth` represents the maximum depth of the decision trees in the training; `num_leaves` represents the maximum number of leaves in one tree; and `learning_rate` represents the step size taken during the algorithm updates the weights of the model. Table 3.2 shows the possible values of each hyperparameter, and 72 ($2 \times 3 \times 4 \times 3 = 72$) combinations of hyperparameter are tested in the grid search.

Table 3.2 Possible values of hyperparameters used in the grid search

<code>n_estimators</code>	[100,200]	<code>max_depth</code>	[10,20,40]
<code>num_leaves</code>	[50,100,150,200]	<code>learning_rate</code>	[0.05,0.1,0.2]

Due to the size of the training dataset is limited, the K-fold cross validation is used to

estimate each hyperparameter combination. By further dividing the whole train set into K-partition and selecting each partition as the new testing dataset and the rest as the new training dataset, the K-fold cross validation makes full use of the whole training dataset (two-month period) and meanwhile prevent the possible information leakage occur when using a fixed validation set [112]. Through iterations of all hyperparameter combinations in grid search, the average performance score for each combination (i.e., the average of the performance scores each K-fold test) will be recorded and compared, and the hyperparameter combination with the highest (or lowest) performance score will be selected as the best hyperparameter. In this study, the performance of the lightGBM is evaluated based on the root mean square error (RMSE):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}, \quad (19)$$

Where n is the number of data points, and y_i and \hat{y}_i are the actual and predicted PV power generation.

3.6 Model Validation

In this section, we present the model validation results. Firstly, a data acquisition system is introduced for a lab house located in Norman, Oklahoma, and then the HVAC system parameter identification results are included. Next, the procedure of retrofitting an electric water heater into a smart water heater is presented and the parameter identification results of the retrofitted water heater are also demonstrated. Finally, due to practical constraints, we validated our PV model using a real operation dataset collected from an open dataset.

3.6.1 *The Lab House and data sources*

The Lab House is in Norman, OK, as shown in Figure 3.5 (a). It is a single-family, one-story home with a floor area of 1,658 ft², built in 1940. The home is equipped with 3.5 tons (42,000 Btu/h) of cooling capacity, 3 tons (36,000 Btu/h) of indoor unit, and 1,400 cfm of air flow rate. As shown in subplots (b) and (c) in Figure 3.5, the outdoor unit locates at the backyard of the Lab House and the indoor unit locates at the attic.



Figure 3.5 Lab House information: (a) outside view of the front of the Lab House (b) an outdoor unit in the back yard of the Lab House (c) an indoor unit in the attic of the Lab House

3.6.1.1 *HVAC system data acquisition*

The floor plan of the Lab House and the sensor locations is demonstrated in Figure 3.6

(a). The entire three-bedroom and one-living-room house is equipped with sensors uniformly

distribuent throughout the space. These data, including air temperature and velocity entering and leaving the key subcomponents of the HVAC system, indoor air and wall surface air temperature, and the power use in the entire house, are measured using different sensors, as shown in Figure 3.6 (a). These data are logged in Raspberry Pis [113], which are connected to a data server (i.e., a Linux desktop) in the Lab House.

The installation of these sensors aims to investigate sensor accuracy, sensor locations, and development of other applications. For the HVAC application, since it aims to explore the potential for practical implementation, we only utilize the data recorded by IoT (internet of things) devices that can be easily installed and connected to the network. Such IoT devices include a smart thermostat installed in the living room to measure the indoor air temperature, four built-in node sensors that paired to the smart thermostat to measure the four interior wall surface temperatures of the house, and a smart power meter inside the electrical panels to measure the power usage for both indoor and outdoor unit, as shown in Figure 3.6 (b), (c) and (d), respectively.

High-quality weather data is essential for parameter identification of the HVAC system. In this study, five-minute historical weather data in Norman, Oklahoma are obtained from Mesonet [114,115] data services. Taking into account that not every region has a free local weather data service like Mesonet, we used the historical weather data provided by Solcast [116] and OpenWeather [117]. These two data sources have their advantages and disadvantages on data accuracy and services, since their comparison goes beyond the scope

of this study, data source will only be referenced when they are used without accuracy comparison. Although all three data services are stable, preprocessing missing and abnormal data is still needed. In this study, second-order polynomial interpolation is used for handling the missing and abnormal data, which is a trade-off between computational effort and interpolation accuracy in real-time operation.

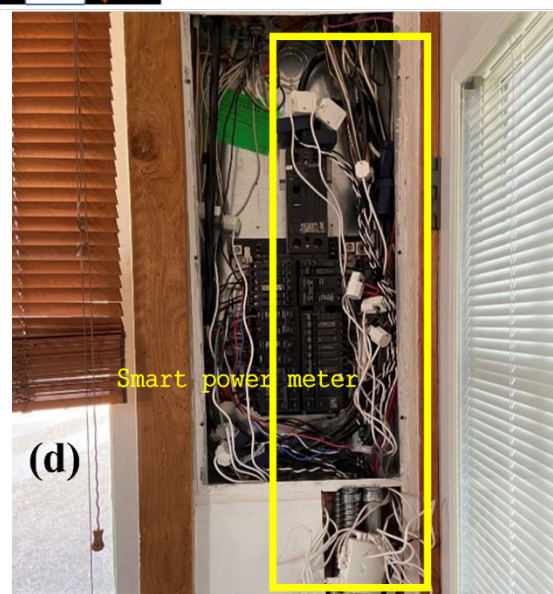
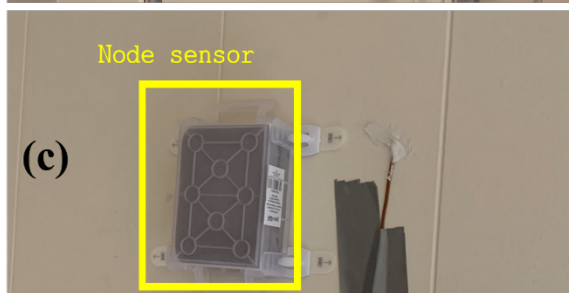
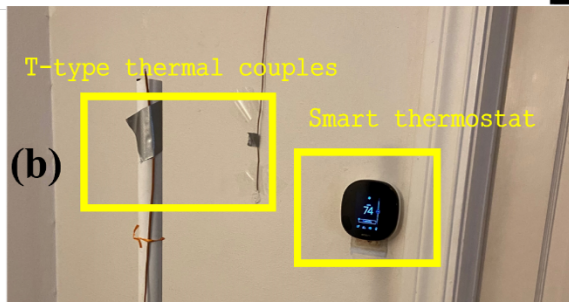
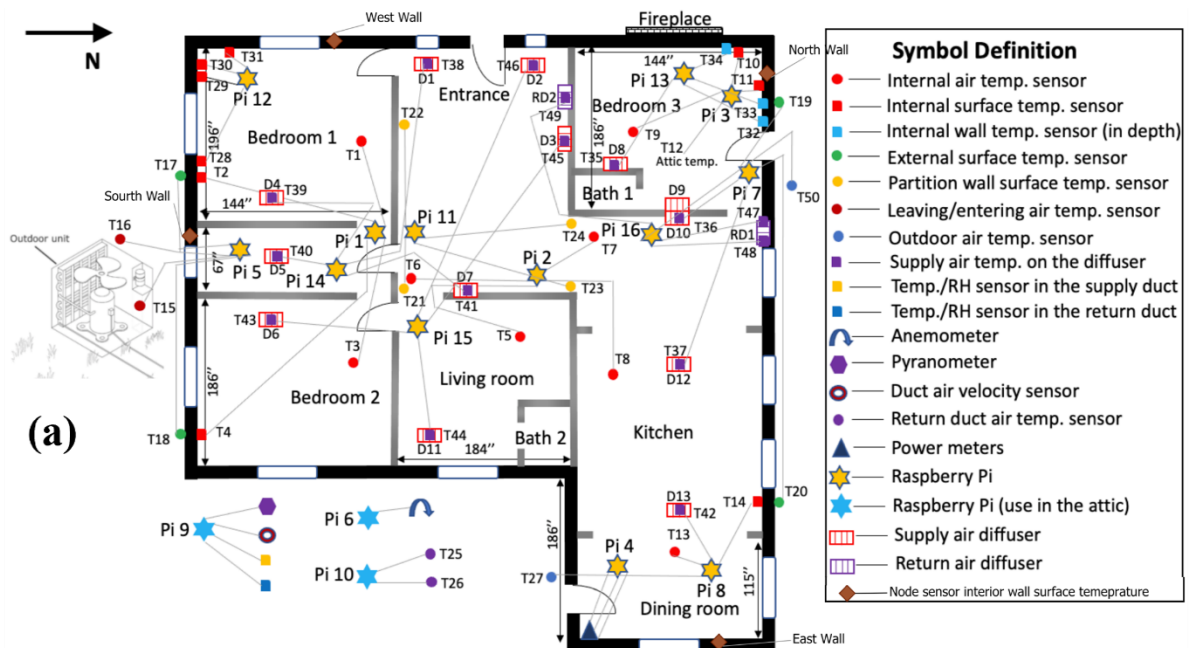


Figure 3.6 Lab House sensors: (a) the floor plan of the Lab House and sensor locations (b) a smart thermostats and thermal couples to measure indoor air temperature (c) a node sensor to measure interior wall surface temperature (d) a power meter inside the electrical panel.

3.6.1.2 *Retrofit a conventional electric water heater into a smart one*

A conventional 50-gallon, 4500-watt electric water heater [118] is also installed in attic of the Lab House to replace an old natural gas water heater. To integrate the water heater usage in the home energy management system discussed in Chapter 5, a proof-of-concept IoT system is developed to capture the hot water flow and temperature, as well as enable the remote water event scheduling, transforming a conventional electric water heater into a smart one. The retrofitted water heater is shown in Figure 3.7.

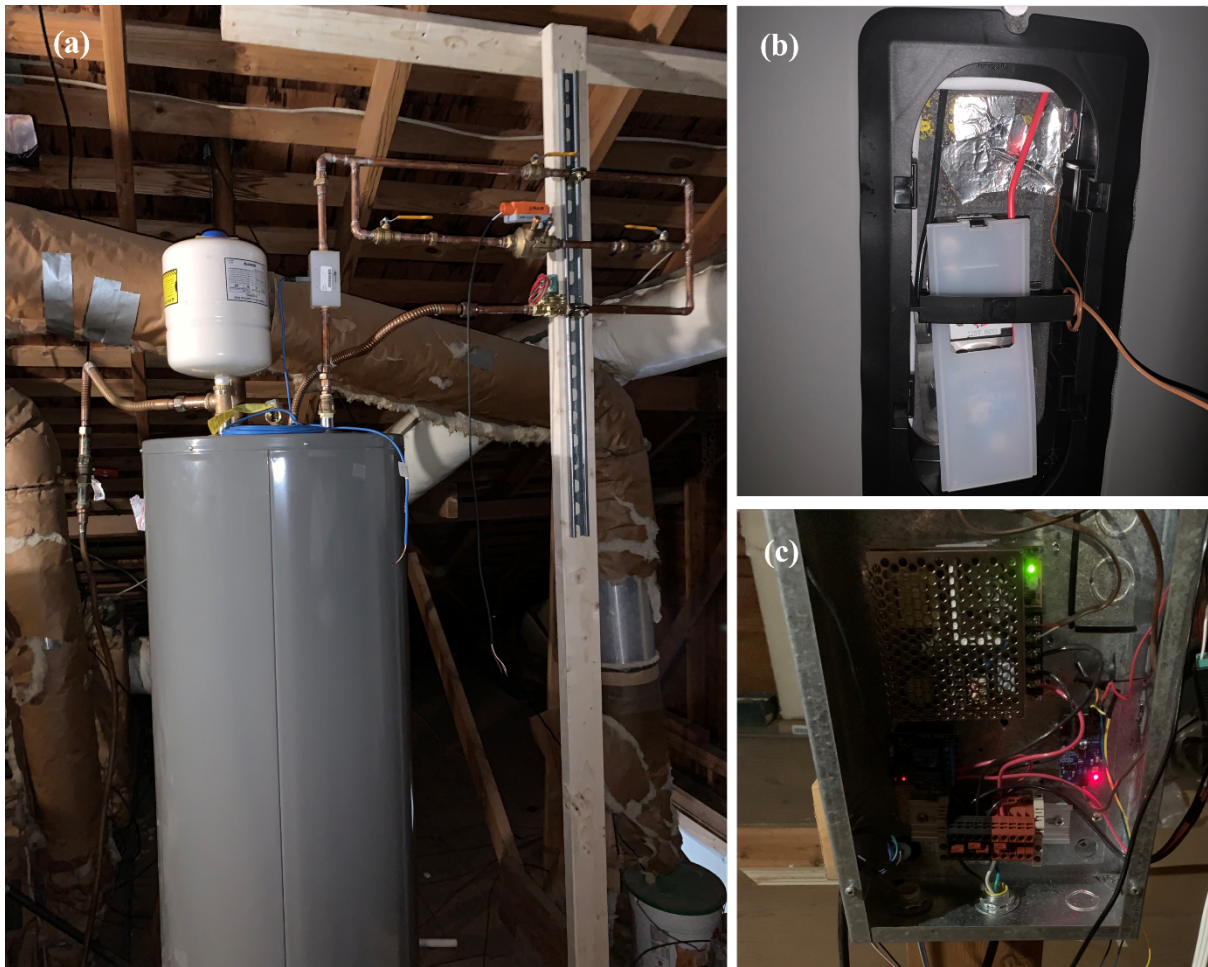


Figure 3.7 Retrofit smart water heater system.

A flow meter and temperature sensors are installed and connected with a Raspberry Pi [113], which serves as a data logger and connects to the data server (the Linux desktop) in the Lab House. Four T-type thermal couples are utilized to measure the upper element water temperature, bottom element water temperature, inlet and outlet temperature. They are located at the upper elements well, the lower element well, the hot water discharge, and the cold-water inlet of the water heater system, respectively. To measure the hot water flow rate withdrawal from the hot water tank, a low flow polypropylene liquid flow meter [119] is installed at the

hot water discharge.

To enable remote water event scheduling, a water scheduling system, including a Raspberry Pi controller, a two-position normally closed Solenoid Control Valve (SCV), a pressure independent Variable Flow Control Valve (VFCV), and a web-based remote scheduling interface, is designed and installed. In the water scheduling system, a Raspberry Pi serves as a controller which offers analog and digital outputs that can trigger solid state electric relay, which powers the SCV that controls the on/off of the hot water withdrawal. To achieve a precise flow rate for each water withdrawal, a VFCV is also installed which controls the water flow rate by adjusting its clip position based on an analogy signal from the Raspberry Pi controller. The web-based remote scheduling interface is a web application that can receive input from the user to conduct local or remote real-time water withdrawal schedules.

Prior to the start of the test, at least three values (i.e., the start time, water draw (in gallons), and VFCV position (0 to 1.0)) are required to be entered into the web-based remote scheduling interface to complete a continuous water withdrawal process, which can be repeated multiple times. These signals will be immediately fed into different Python scripts and calculated into two control sequences, including volume control sequence sent to the SCV and flow rate control sequence sent to the VFCV. Once the time reaches the scheduled start time in volume control sequence, an “on” signal is given to the SCV and starts to record the pulse count; once the number of pulses indicate that the total withdrawal volume has been reached, a “off” signal is given to the SCV. As the flow rate control sequence, the VFCV position (0 to 1) obtained

from the web-based scheduling interface will be directly sent to a digital analog converter (DAC) that convert this VFCV position to a 0-10 VDC analog command 60 seconds prior to the scheduled start time. Calibration to investigate the relationship between VFCV position and the actual water flowrate is needed because it can be affected by many factors other than VFCV characteristics.

3.6.1.3 PV forecast data sources

Due to limitations in the Lab House, a PV rooftop real time operation dataset is collected from National Renewable Energy Laboratory's (NREL) PV Rooftop Database (PVRDB) [120]. The dataset includes PV rooftop real time operation data for both residential and commercial buildings for multiple years across US.

3.6.2 HVAC parameter identification results

To identify the parameters of the 2R2C model, we solved the optimization problem stated in (20)– (21) using measured data. To validate its accuracy, we simulated the model using a different set of measured data. Figure 3.8(a) shows the simulated (black) and measured (red) indoor air temperature, from which it can be seen that the model is quite accurate, achieving a mean absolute error (MAE) of 0.503°F . Note that because of model simplicity, the identified parameters can be sensitive to training datasets. Moreover, although being a gray-box model it has some extrapolation capability, regular model retraining and parameter updating are still needed while the optimal control operates.

To identify the parameters of the total AC power usage model, we solved a polynomial regression problem of the form (23) using measured T_o and $P_{outdoor}$ data. To eliminate the impact of transient AC operation, only data with sufficiently large outdoor unit power (e.g., exceeding 3.2kW) were used. Figure 3.8(b) shows the resulting polynomial (red) alongside the measured data (black). The coefficient of determination (R^2) of 0.845 and root-mean-square error ($RMSE$) of 0.162kW suggest that the polynomial fit is satisfactory.

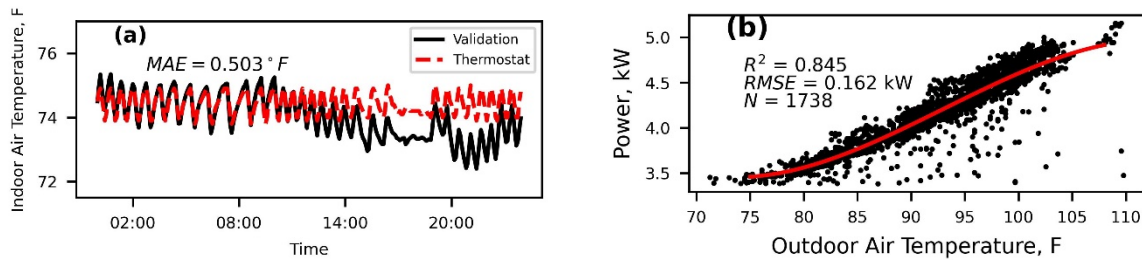


Figure 3.8 (a) 2R2C model validation result and (b) AC outdoor unit consumption regression result (power lower than 3.2kW is excluded).

3.6.3 Water heater calibration and parameter identification results

To obtain the relevant dataset which described the dynamic model in Equation (14), we collected the bottom and top elements' water temperature from the thermal couples. The water schedule event (in Gallon per meter) is calculated from the pulse in a water flow meter. The total power consumption from both bottom and upper elements is obtained from the energy monitor. In the parameter identification, we use the average value of the bottom and top elements as the measurement of T_{wh} . Although a thermal couple was put in to measure the inlet water temperature, the reading was affected by the water supply, hence is not used. Besides, the city water inlet is relatively constant, we assume it keeps $73.4^{\circ}F$ ($23^{\circ}C$) during the test.

The water tank was exposed to the house’s attic, which was also assumed to be a constant temperature.

Figure 3.9 is the simulation result of the RC network model of the water heater. Since the test is conducted in November, we assume the attic temperature is kept at $68^{\circ}F(20^{\circ}C)$. The mean absolute value for this simulation is 2.144, which is slightly high. The major reason is the thermal couple’s reading of the bottom and top elements have a large variation, making the T_{wh} measurement is not accurate. The estimation of the R-value is $20.04 \text{ ft}^2\text{hr/Btu}$.

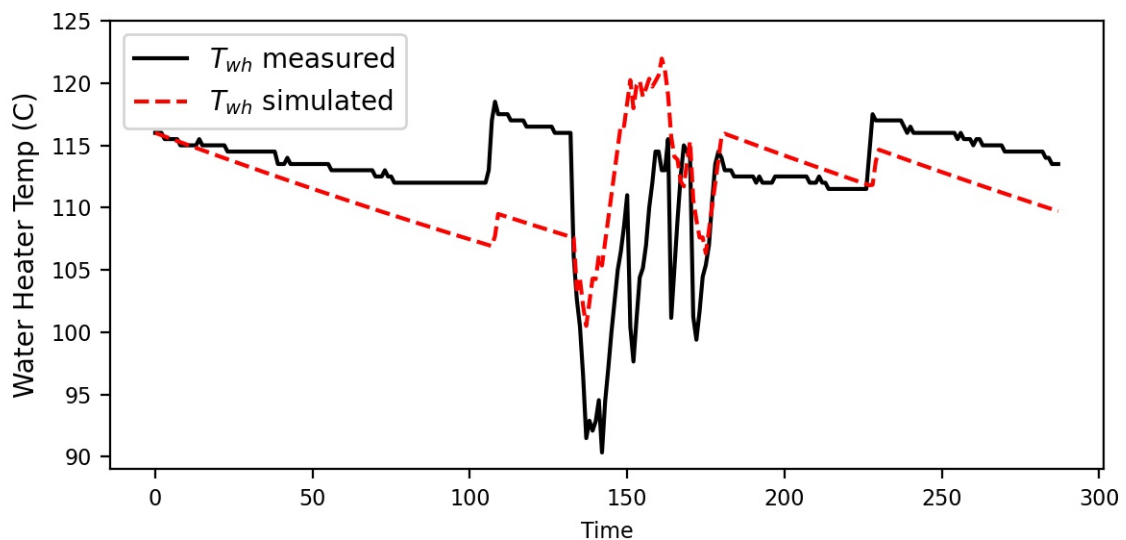


Figure 3.9 Water heater system identification result

3.6.4 Short-term PV power generation results

To test the robustness of the proposed prediction model, random three-day period in each month from July to October are selected as a test dataset and the dataset from the two months prior to these three-day periods are used as a train dataset. In these experiments, different random seeds are utilized to generate these random three-day periods, and to further partition

the train and valid subsets for K-fold cross validation or feature selection.

Figure 3.10(a) demonstrate the SHAP values of every data point (e.g., air temperature = 30°C) for the PV power prediction on September dataset using random seed 123, and the features in the dashed box are selected. As expected, the GHI has a high contribution on the model, whereas some features such as relative humidity has a low contribution. Figure 3.11 illustrates PV power generation, GHI, and relative humidity in original dataset on September 2. As can be seen in Figure 3.11, the trend of GHI is very similar to that PV power generation, while the trend of relative humidity is almost opposite to it, which further indicates remove relative humidity feature might benefits the model. A reasonable explanation for why the trend of relative humidity is opposite to that PV power generation is because of the increase in temperature and sufficient illumination in the afternoon, a large amount of moisture in the air evaporates, resulting in decreasing in relative humidity. Figure 3.10(b) shows the RMSEs of the lightGBM model trained on September dataset with random seed 123 for different values of max_depth and num_leaves when the learning rate and n_estimators are chosen from the best model and fixed. The optimal hyperparameters chosen in this model is:

- Max_depth = 20
- num_leaves = 25
- learning rate = 0.2
- n_estimator = 300

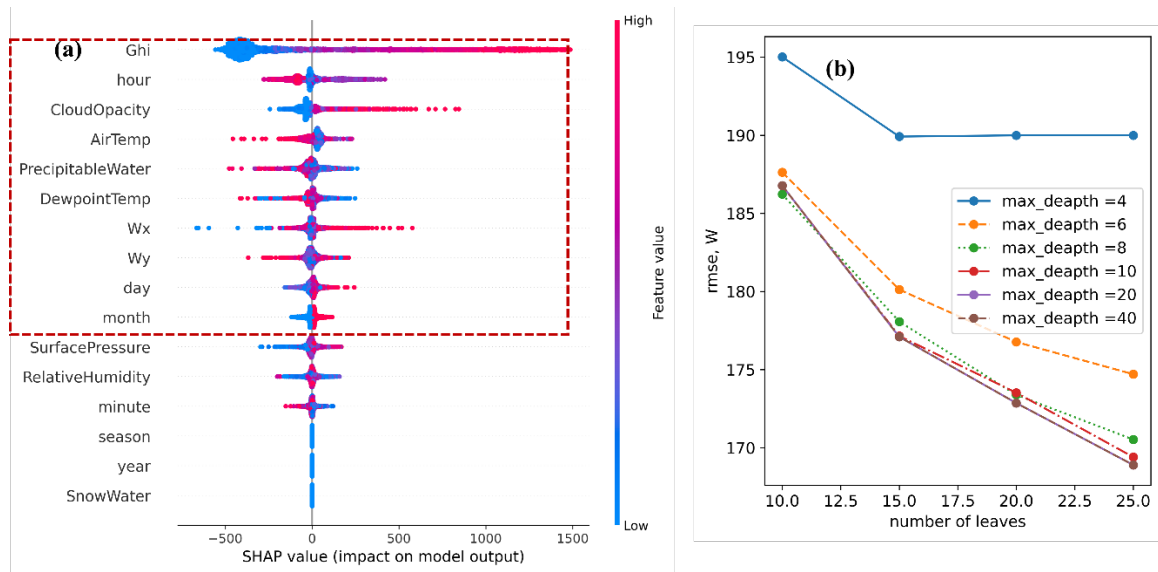


Figure 3.10 Model performance improvement (a) SHAP value for every train data points in feature selection for September (b) RMSE results on validation sets on September in grid search

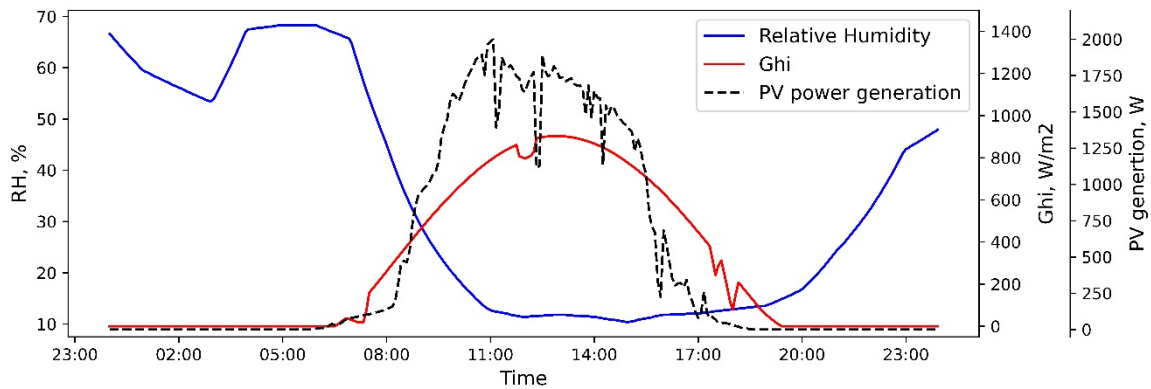


Figure 3.11 PV power generation, GHI, and relative humidity in original data on September 2, 2019

Table 3.3 shows the performance on four months testing datasets of prediction in terms of RMSE in unit of W. Overall, the prediction model has highest performance on October and lowest performance in July. Figure 3.12 shows the prediction results of lightGBM model on August dataset (from August 3 to August 5 in 2019). As can be observed from the figure, the model prediction is accurate when the illumination is sufficient. However, under the condition

of insufficient illumination, for example, a sudden drop in afternoon sunlight and subsequently affecting the amount of PV power generated in August 4, the model prediction results are not as good as that in other time periods.

Table 3.3 The performance results on testing dataset

	RMSE (<i>kW</i>)			
Random seed	July	August	September	October
42	0.323	0.351	0.191	0.286
100	0.397	0.296	0.347	0.213
123	0.358	0.342	0.297	0.220

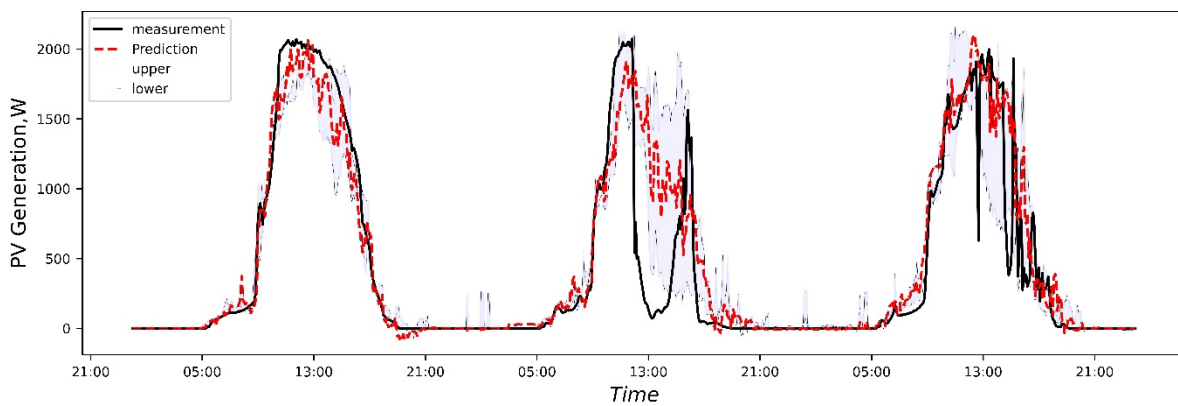


Figure 3.12 PV power generation point and interval prediction compared with actual generation from August 3, 2019 to August 5, 2019

3.7 Summary

In this section, the comparison of white, light-gray, and black box models is conducted first, aiming at proper model type selection for each different load-flexible resources in

residential building. Due to the advantages of the light-gray box model in modeling dynamic system, it is applied to model thermostatically controlled appliances. The black-box model, however, is more suitable for modelling the PV panel generation, which is affected by many different factors.

For different modeling approach, we have introduced different parameter identification framework considering both model accuracy and implantation effort. A novel set of constraints to ensure that the model parameters identified are realistic is applied to parameter identification for the light gray-box model established for describe the home thermal dynamics involving the HVAC system. A feature selection procedure with SHAP value and hyperparameter tuning with grid-search and K-fold cross validation is applied to improve the lightGBM model.

Finally, the modeling process and the model validation results for the HVAC system, the water heater tank, and the PV panel are shown. In this study, several results for different model are achieved. The HVAC system outdoor unit power can be predicted with a root mean square error of 0.162 kW using a regression model. Forecasting of the water temperature of dynamic within the water heater tank, using the first-order model, resulting in a mean absolute value of 2.144. Lastly, the LightGBM model facilitates a forecast of 3 day ahead PV power generation with a 0.191 root mean square error.

4 Development, Implementation, and Impact Analysis of MPC-based Precooling for Homes

4.1 Introduction

After describing the flexibility resources model and validating these models by using real-world operation data, in this chapter, we introduce the optimal operation of the HVAC system under DR programs. As can be seen in Figure 1.4, the passive thermal storage, in general, is the most complicated and sophisticated component among available flexibility resources, making the development of automated software platform challenge. On the other hand, the HVAC system, in most of the case, is the biggest energy “eater” in a home, making the development, implementation, and impact analysis on its optimal operation has high research and practical value.

In this study, an MPC-based precooling real time control framework has been developed, and this framework has been implemented in real homes (field test study) and a virtual testbed (simulation study). Both methods have validated the robustness and effectiveness of the MPC-based precooling. More importantly, field test study, which focuses on impact analysis across multiple homes, is used to analyze MPC-based precooling in different homes and operation conditions; while the simulation study, which focus on impact analysis in a virtual building mimicking the real house, is used to analyze the benefits of MPC-based precooling in difference choice of MPC hyperparameters (including prediction model).

4.2 Real-time Control Framework of the MPC-based Precooling

4.2.1 Mixed-integer linear programming problem formulation

Having described the reduced-order gray-box model and its parameter identification, in this subsection the MPC agent is introduced. At the heart of the MPC agent is a mixed-integer linear programming (MILP) problem for optimal precooling to be solved at each time step k . The objective function of this MILP problem is given by:

$$\min_U \sum_{m=k}^{k+N} price^m P^m u^m + w v^m, \quad (22)$$

where $U = (u^k, \dots, u^{k+N}, v^k, \dots, v^{k+N})$ is a vector of decision variables, N is the length of the optimization horizon, $price^m$ is the TOU rate at time step m , P^m represents the total AC power usage, $u^m \in \{0,1\}$ is the AC on/off signal, $v^m \in \mathbb{R}_{\geq 0}$ is a non-negative slack variable intended for penalizing indoor air temperature deviation from a thermal comfort range, and w is a weight for the slack variables. The total AC power usage P^m is modeled by:

$$\begin{aligned} P^m = P_{indoor} + P_{outdoor}^m = P_{indoor} + A_1 T_o^m + A_2 (T_o^m)^2 + A_3 (T_o^m)^3 \\ + A_4 \text{ for } m = k, \dots, k + N, \end{aligned} \quad (23)$$

where P_{indoor} is the constant indoor unit power, $P_{outdoor}^m$ is the variable outdoor unit power assumed to be a polynomial function of T_o^m , and A_1, A_2, A_3, A_4 are coefficients that can be experimentally determined.

The constraints of the MILP problem are given by:

$$T_{ie}^m = \frac{\Delta t}{\tau_1} (T_o^{m-1} - T_{ie}^{m-1}) + \frac{\Delta t}{\tau_2} (T_{in}^{m-1} - T_{ie}^{m-1}) + T_{ie}^{m-1} \quad (24)$$

for $m = k + 1, \dots, k + N$

$$\begin{aligned}
T_{in}^m = & \frac{\Delta t}{\tau_3} (T_{ie}^{m-1} - T_{in}^{m-1}) + \frac{\Delta t}{\tau_3} (a_1 S^{m-1} + a_2 (S^{m-1})^2) \\
& + \frac{\Delta t}{\tau_3} (T_o^{m-1} - T_{in}^{m-1}) (b_1 W^{m-1} + b_2 (W^{m-1})^2) \\
& + \frac{\Delta t}{\tau_3} (c_1 T_o^{m-1} + c_2 (T_o^{m-1})^2) u^{m-1} + T_{in}^{m-1}
\end{aligned} \tag{25}$$

for $m = k + 1, \dots, k + N$

$$T_{lb}^m - v^m \leq T_{in}^m \leq T_{ub}^m + v^m \quad \text{for } m = k, \dots, k + N \tag{26}$$

$$u^m \in \{0,1\} \quad \text{for } m = k, \dots, k + N \tag{27}$$

$$v^m \in \mathbb{R}_{\geq 0} \quad \text{for } m = k, \dots, k + N, \tag{28}$$

where (24) and (25) are taken from discretizing 2R2C model using Euler forward method, (26) represents soft constraints on thermal comfort, T_{lb}^m and T_{ub}^m are lower and upper bounds on T_{in}^m , (27) ensures that u^m is binary, and (28) forces v^m to be non-negative. Notice that the MILP problem makes use of two initial conditions T_{in}^k and T_{ie}^k , which come from real-time measurements at time step k .

In general, the MILP problem may be infeasible because, for example, the forecasted outdoor air temperature is too high for the AC to keep up, or the homeowner's thermal comfort range is too narrow for the indoor air temperature to stay in. Such infeasibility may cause the MPC agent to crash. To ensure that the MILP problem is always feasible, instead of using hard temperature constraints $T_{lb}^m \leq T_{in}^m \leq T_{ub}^m$ for $m = k, \dots, k + N$, the constraints are

"soften" by inserting into (26) a non-negative slack variable v^m which represents the degree to which the hard temperature constraints are violated. Since such violation is undesirable, a term wv^m is also insert into the objective function (22), which only imposes a penalty when T_{in}^m turns out to be higher than T_{ub}^m or lower than T_{lb}^m . In other words, the temperature constraints are relaxed only when necessary.

4.2.2 *MPC agent*

Building upon the MILP problem, the MPC agent workflow is now introduced, shown in Figure 4.1. As can be seen from the figure, at each time step k , the home thermal model (i.e., 2R2C model) receives the identified values of its parameters, weather forecasts over the optimization horizon, and both the indoor air and interior wall surface temperature measurements. The MPC agent then receives the TOU rate, the identified parameters of the total AC power usage model, and the thermal comfort upper and lower bounds, through which the objective function and constraints of the MILP problem are generated. The MILP problem is subsequently solved to generate an optimal sequence of AC on/off signals and slack variables. Finally, as is typical with MPC algorithms, only the first element of the optimal sequence of AC on/off signals is physically implemented. The MPC agent can implement this decision by adjusting the temperature setpoint at the smart thermostat in such a way that the AC is forced to turn on or off until the next time step.

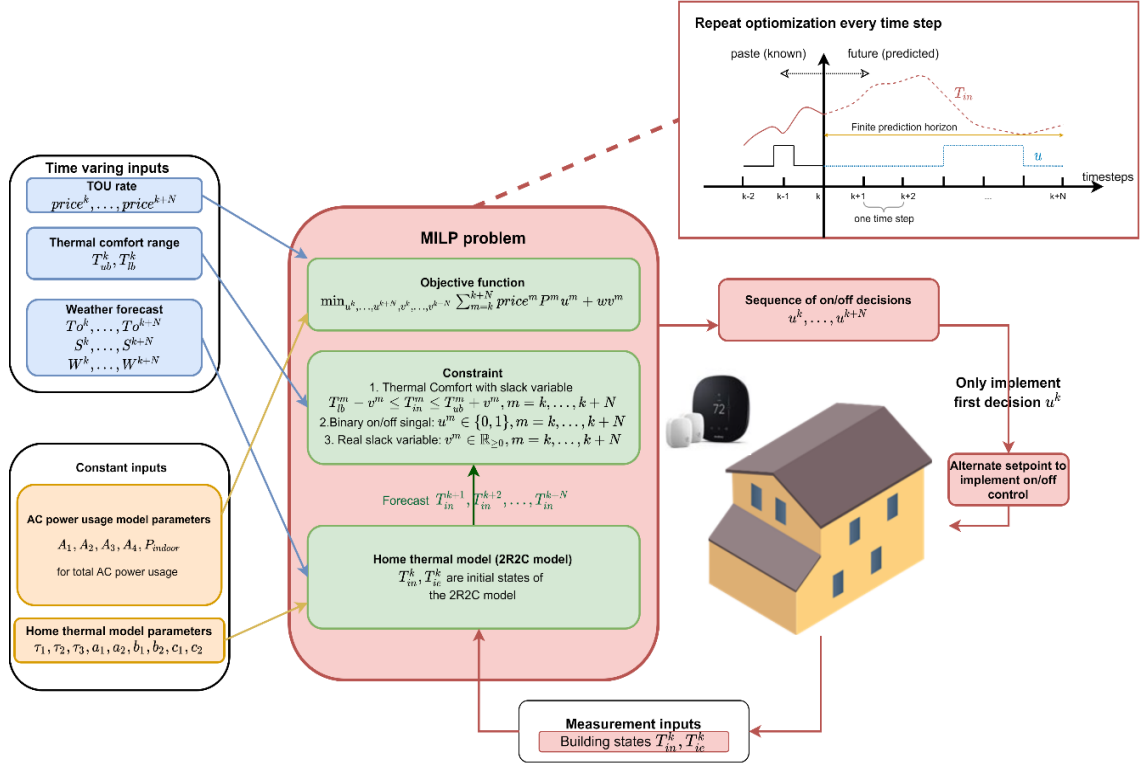


Figure 4.1 MPC agent.

4.2.3 Performance indicators for MPC agent

Two performance indicators are used to evaluate the performance of the MPC agent.

Based on (22), the 24-hour AC operation cost C is given by:

$$C = \frac{\Delta t}{\mathcal{J}} \sum_{m=1}^{N_{24h}} price^m P^m u^m, \quad (29)$$

where \mathcal{J} is a unit conversion factor chosen so that C has unit dollar, $m = 1$ represents the beginning of a 24-hour period, and $m = N_{24h}$ represents the end. With C , the *cost saving*

\mathcal{J} —the first performance indicator—is defined as:

$$\mathcal{J} = \frac{C_{baseline} - C_{mpc}}{C_{baseline}} \times 100\%, \quad (30)$$

where C_{mpc} is the 24-hour AC operation cost with the MPC agent, and $C_{baseline}$ is the cost without (i.e., the AC is under normal operation).

Although cost saving \mathcal{J} is a useful indicator of the MPC agent's effectiveness, it does not fully reflect the latter's load-shifting ability. To address this limitation, among the metrics for evaluating energy flexibility [14,63,78,121], the *flexibility factor* FF —the second performance indicator—is adopted and defined as:

$$FF = \frac{\int_{L_{low}} P dt - \int_{L_{high}} P dt}{\int_{L_{low}} P dt + \int_{L_{high}} P dt} \quad (31)$$

where L_{low} (L_{high}) represents the off-peak (on-peak) hours and P is the total AC power usage first defined in (23). Note that FF is always between -1 and 1, where -1 means all energy is consumed during on-peak hours, while 1 means it is consumed during off-peak hours. Therefore, flexibility factor FF quantifies the load-shifting effectiveness of the MPC agent.

4.3 MPC-based Precooling in Real Homes

4.3.1 *Cloud-based platform for data management and control*

The proposed cloud-based platform is shown in Figure 4.2. The platform is developed using Python and supports several functions of IoT devices including data communication, alert system, and remote control. To run the MPC-based precooling test on the platform, three types of datasets are required: the parameter identification dataset, the real-time control dataset, and the post-analysis dataset. Table 4.1 lists, for each type of dataset, its purposes, components, data collection approaches, and data sources. In this paper, a uniform time step Δt of 5

minutes is used for all three types of datasets. In addition, the optimization horizon of 12 hours (i.e., 144 time steps) is chosen, which means that at each time step, 12-hours-ahead forecasts of weather and the TOU rate are required. To collect and preprocess data of different types, several Python modules have been developed. Moreover, the open-source time series database InfluxDB 2.0 [122] is used for data storage and communication, and Grafana v7.4 [123] is used for data display and user interface.

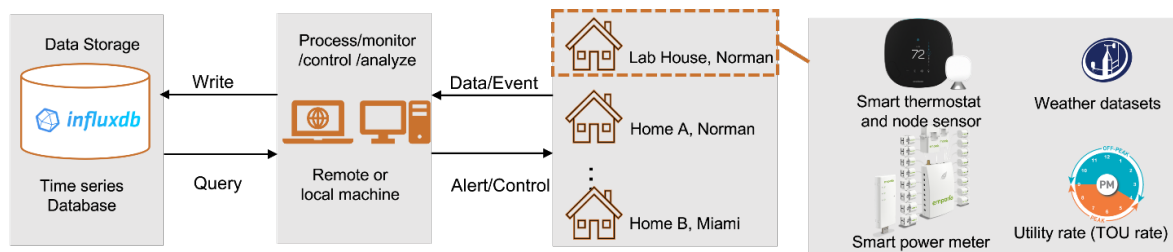


Figure 4.2 Proposed cloud-based platform for data management and control.

Table 4.1 Three Types of Dataset

Types	Purposes	Components	Approaches	Data Sources
Parameter identification dataset	Train 2R2C model	Indoor air and interior wall surface temperatures; AC on/off signals	Query from database	Smart thermostat; node sensor
		AC power usage	Query from database	Smart power meter [124]
		Historical weather dataset	Scrap from website	Mesonet [39]; [40]; and weather services [117]
Real-time control dataset	Initial states for home thermal model	Indoor air and interior wall surface temperatures	Query from thermostat API	Smart thermostat; node sensor

	Home thermal model inputs for MILP formulation	Weather forecast; TOU rate	Query from weather API; Query from utility	Solcast [116] and weather services [117]; utilities [125–127]
	Thermal comfort inputs for MILP formulation	Upper and lower bounds on indoor air temperature	Homeowners' previous preferences or current thermal comfort requests	Historical setpoints; private correspondence
Post-analysis dataset	Other information for analysis	AC mode; away time	Query from smart thermostat API	Smart thermostat

4.3.2 *Implementation setup*

In this section, we introduce the experimental setup for the MPC-based precooling test, including the participating test homes, and the temperature and power measurement setup in those homes.

4.3.2.1 *Introduction of eleven (11) participating test homes*

The test homes include ten single-family homes in Norman, Oklahoma, and one single-family home in Miami, Florida. To ensure that a diverse set of test homes were chosen, a survey was carried out for each home to gather its building characteristics including its floor area, building age, building structure material, AC unit condition, and whether the homeowner had signed up for a TOU rate. The survey result is tabulated in Table 4.2 Except for what we refer to as the Lab House near the University of Oklahoma, the test homes were labeled alphabetically based on the order in which they were registered. Figure 4.3 shows the

distribution of floor areas, count plot of AC tonnage, building ages, and AC unit service time for the 11 test homes. As can be seen from Figure 4.3 the test homes were diversely distributed in terms of building and AC characteristics.

Table 4.2 Test home information.

Home	Floor Area	Building Age	Building Material	Year of AC	# AC Units	AC Tonnage	TOU Rate	Smart Power Meter
Home A	2400	10-20 years	Wood Frame	1-3 years	1	5	-	Yes
Home B	2560	10-20 years	Concrete	1-3 years	2	3 (up)/4 (down)	-	Yes
Home C	3500	Above 20 years	Wood Frame	0-1 years	2	2.5 (up)/4 (down)	Yes	-
Home D	2500	10-20 years	Wood Frame	8-15 years	1	3.5	-	-
Home E	1540	1-5 years	Wood Frame	1-3 years	1	2.5	-	-
Home F*	1812	10-20 years	Wood Frame	1-3 years	1	-	-	-
Home G	1800	Above 20 years	Wood Frame	1-3 years	1	4	-	-
Home H**	3000	Above 20 years	Wood Frame	8-15 years	1	-	-	-
Home I	2600	Above 20 years	Unknown	0-1 years	1	5	-	Yes
Home J	2150	10-20 years	Wood Frame	8-15 years	1	2.5	Yes	Yes
Lab House	1658	Above 20 years	Wood Frame	1-3 years	1	3.5	-	Yes

* This home withdrew from participation when it was sold in June 2022.

** This home withdrew from participation due to thermostat incompatibility.

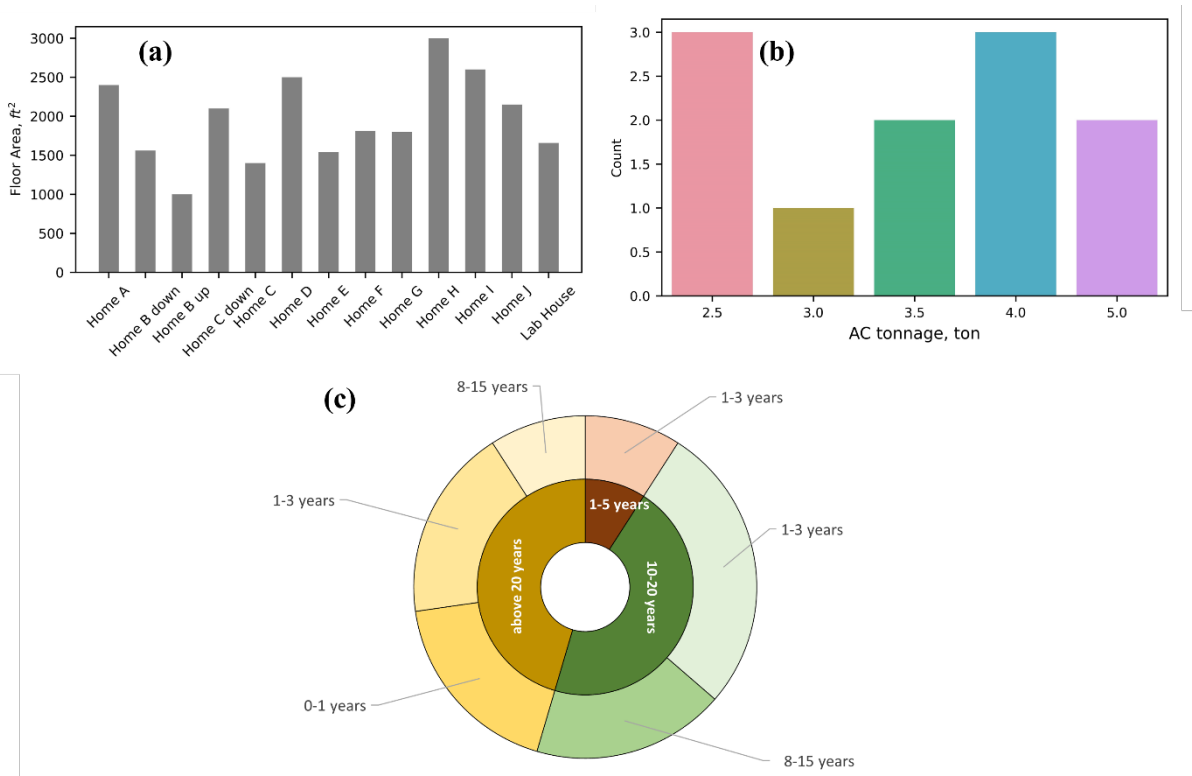


Figure 4.3 Test homes with diverse characteristics. (a) Building floor area distribution. (b) Count plot of AC tonnage. (c) Building ages (inner) and AC unit service time (outer).

Although there were 11 test homes, two withdrew from participation shortly after the field test began (Home F was sold and Home H had thermostat incompatibility issues). In addition, during the test period, only two participants (Home C and Home J) enrolled in TOU programs. For the rest who did not enroll, the MPC-based precooling test carried out anyway assuming that they did. At where the homes were located, three different TOU programs (i.e., rate structures) were available: Oklahoma Gas & Electric (OG&E) Smart Hours [125] where on-peak hours run from 2 pm to 7 pm, Oklahoma Electric Corporate (OEC) TOU rate [126] where on-peak hours run from 3 pm to 7 pm, and Florida Power and Light (FPL) TOU rate [127] where on-peak hours run from 12 pm to 9 pm. With the sponsorship of the research, all the

homes were installed with smart thermostats and node sensors, among which five (Home A, Home B, Home I, Home J, and the Lab House) also had smart power meters installed for collecting indoor and outdoor power measurements. For those houses without smart power meters, the total AC power usage P in (23) was assumed to be a constant depending on the AC model.

The upper and lower temperature bounds in the MPC-based precooling test were determined based on observed homeowner preferences during normal operation and on what they considered to be acceptable during the MPC-based precooling test. Figure 4.4 presents the Gaussianized setpoint distribution of several homes during normal operation (black) and during the test (red). Note that a 90°F setpoint means the homeowner manually turned off the AC. Also note that because no one lived in the Lab House, its temperature bounds can be freely chosen. As such, the bounds were set to different values to allow us to examine their impact.

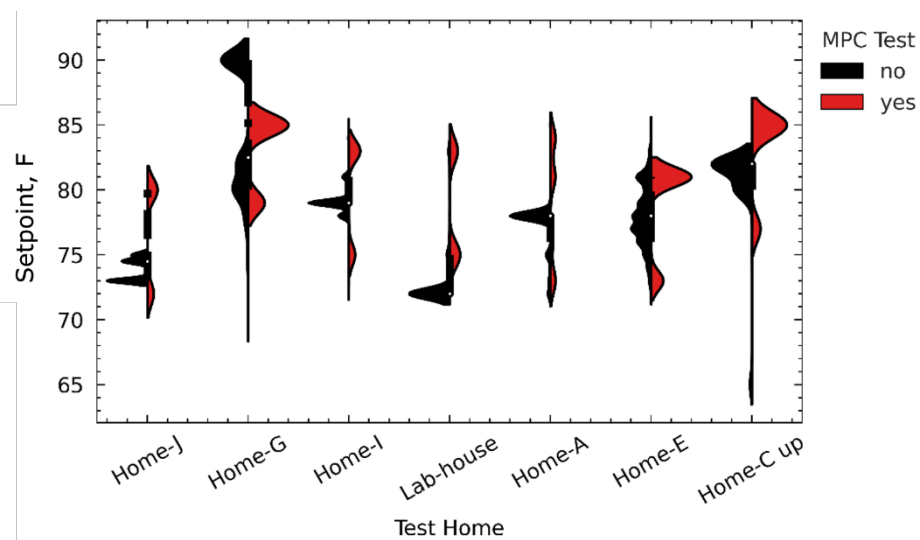


Figure 4.4 Setpoint distribution in several test homes.

4.3.2.2 *Temperature and power measurements*

Figure 4.5 shows the temperature and power measurement setup at some of the test homes. Specifically, Figure 4.5(a) shows a smart thermostat for tracking the indoor air temperature and communicating with the cloud-based platform, Figure 4.5(b) shows a node sensor for measuring the interior wall surface temperature, and Figure 4.5(c) shows a smart power meter for measuring the total energy usage of each circuit inside the electric panel. The latter makes it possible to distinguish between the energy consumed by indoor and outdoor AC units.

In the 2R2C model described by (32) and (33), the interior wall surface temperature is one of the two states that reflects the heat transfer between indoor and outdoor environment. Accurate measurement of such temperature—which depends heavily on the sensor location—is therefore important. The node sensor is preferably mounted on a wall with no direct solar, no wind, and no shading. Thus, north-facing walls were selected because solar impact from the sun was lower compared to south- and west-facing walls, and east-facing walls for some of the test homes were chosen due to their structure. The walls near a diffuser, the kitchen, and the windows were avoided to not expose the sensor to unwanted disturbances.

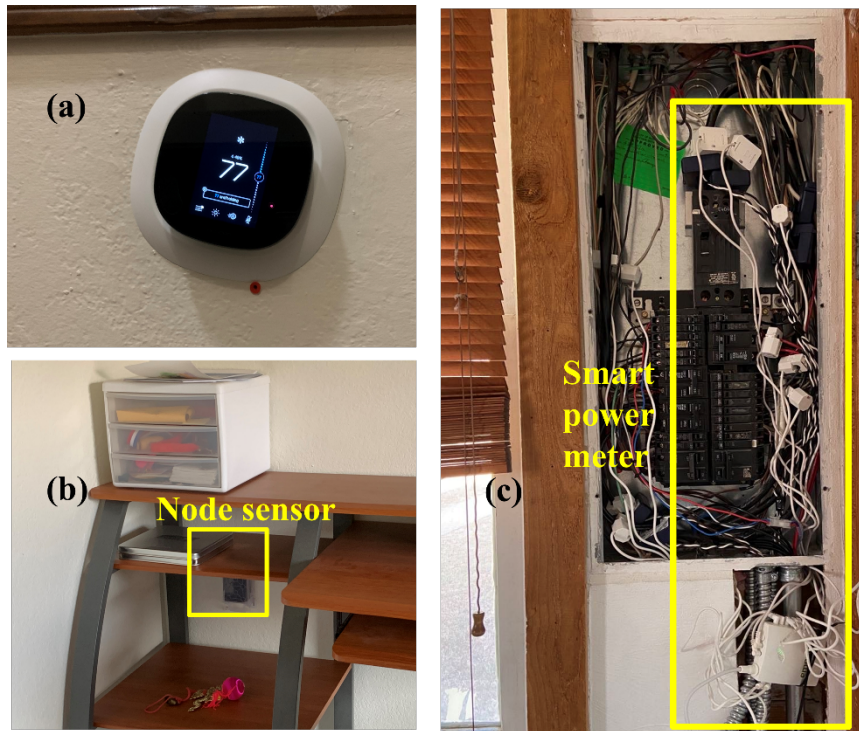


Figure 4.5 Equipment installed at test homes. (a) Smart thermostat. (b) Node sensor. (c) Smart power meter.

4.3.3 MPC precooling test filed test results

In this section, a discussion of the MPC agent performance at one test home is presented first. Then, by using results from multiple test homes, assessment of the MPC agent through with and without performance comparison and energy flexibility comparison are demonstrated. The dominant factors that impacted the MPC-based precooling test are analyzed at the end.

4.3.3.1 MPC performance assessment at one test home

From July to October 2022, a set of MPC-based precooling tests were conducted at different times at the nine test homes including the Lab House. Each test lasted five to 20 days depending on the homeowner's schedule, except for the test at Home G, which was terminated

prematurely at the homeowner’s request. This section describes the MPC agent performance at one test home—the Lab House.

Figure 4.6 shows seven consecutive days (from August 4 to August 10, 2022) of test results at the Lab House. Specifically, Figure 4.6(a) depicts the indoor air temperature T_{in} (blue) and interior wall surface temperature T_{ie} (red) along with the thermal comfort range T_{ub} and T_{lb} (black). Figure 4.6 (b) displays the TOU rate and number of seconds in a 5-minute interval during which the AC was on. Figure 4.6 (c) shows the actual weather conditions (solar irradiation S , wind speed W , and outdoor air temperature T_o). Also shown in the figure are red-shaded areas representing the on-peak hours. Analyzing Figure 4.6 , several observations can be made. First, T_{in} mostly stayed within T_{ub} and T_{lb} but did occasionally exceed T_{ub} . When T_{in} exceeded T_{ub} , the slack variables in (22) imposed a penalty on the objective function. If this happened during the on-peak hours, the penalty might be small compared to the high electricity cost and hence might not cause the AC to turn on immediately. Second, T_{ie} changed much more slowly compared to T_{in} . This was because the building structure was a much larger thermal storage compared to indoor air. Such a difference also justified the use of the 2R2C model. Third, the MPC agent was able to take advantage of the cooler early morning air to precool the space with higher efficiency. More importantly, unlike rule-based precooling control which has a fixed precooling schedule regardless of weather, the MPC agent was capable of adapting its schedule to weather, TOU rates, etc. For example, precooling began at 4:40 am on August 4 and 5:50 am on August 10. Fourth, despite the

precooling, T_{in} often reached T_{ub} before the on-peak hours ended, forcing the AC to be on even though that was costly. These observations show that the MPC agent behaved as intended. Finally, we remark that the reason T_{in} reached as high as 83°F on August 10 was that the MPC agent went offline unexpectedly.

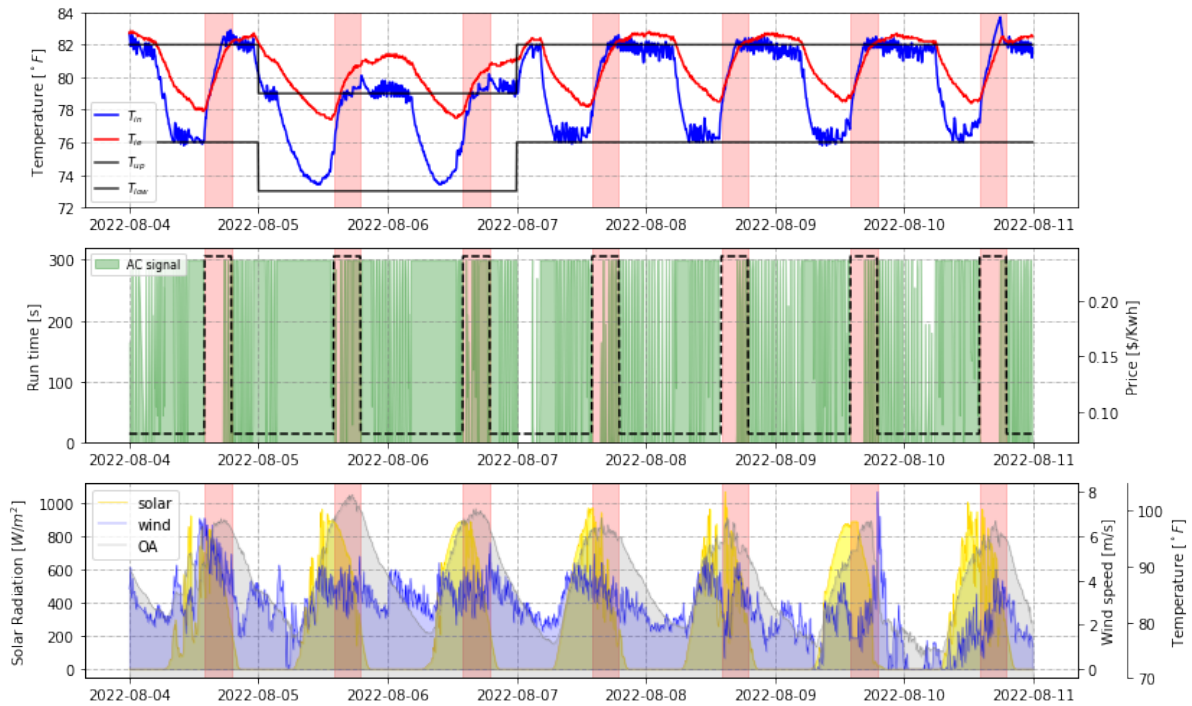


Figure 4.6 Seven consecutive days of MPC-based precooling test results. (a) Smart thermostat and node sensor measurements. (b) AC signals and TOU rate. (c) Actual weather conditions.

4.3.3.2 MPC performance assessment at multiple test homes

In this section, the result with and without the MPC agent are compared and then the building's energy flexibility is assessed by comparing the flexibility factors across multiple homes. Since the weather is a dominant input variable for AC operation, to eliminate the weather influence, a modified matrix profile-based weather clustering algorithm is introduced

to assess the MPC performance in different weather conditions.

A. Weather clustering

To allow fair comparisons without complicated calculation, and filter out the abnormal weather conditions, a modified weather clustering algorithm is developed by combining Matrix Profile-based algorithms discussed in a series of studies [128–130]. To cluster the long (four months, 5 minute interval), time-series (focusing on clustering subsequences rather than a single point), and multi-dimension (outdoor air temperature, solar, and humidity) data, the Snippets algorithm [130] is used to find the three most representative days of the test period. Then, days exhibiting high pattern similarity with the representative days are selected using the modified algorithm which considers solar irradiation, outdoor air, and relative humidity. The details of the modified algorithm can be seen in Appendix A.

Figure 4.7 shows the weather clustering results for Norman (Subplot (a)) and Miami (Subplot (b)). The bold curves show the outdoor temperatures of three representative days, the hot (red), mild (orange), and cold (blue) days found by the Snippets algorithm from July to October. The dashed curves are the 24-hour time series weather data associated with the three representative days. The solar irradiation and relative humidity are also considered, but not shown in the figures. The grey curves are the rest of the days that are not included in any cluster, which are labeled as extreme days. For the data in Norman, the layer between each cluster is clear, which means the data within each group are similar to each other. Since Miami

experienced a severe thunderstorm during the test period, its cold cluster includes some abnormal weather change periods compared with the hot and mild clusters. Hence, each day in the cold cluster has a large difference, but a similar shape.

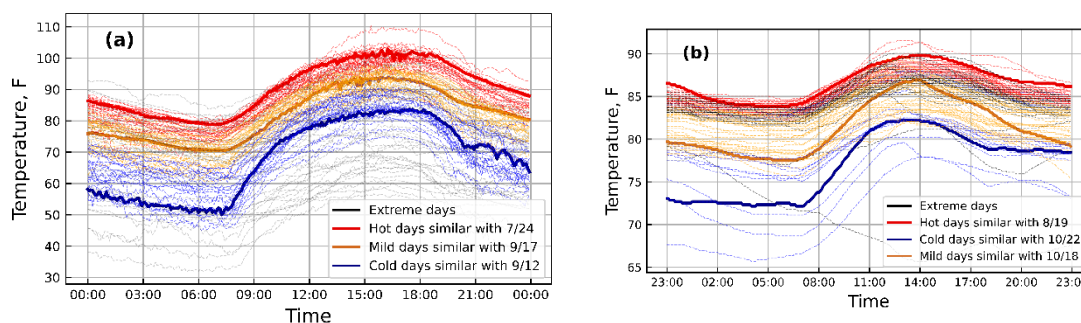


Figure 4.7 Weather clustering results in (a) Norman and (b) Miami from July to October

B. Performance with and without MPC agent

Figure 4.8 shows the breakdown of the average daily AC system cost by on-peak and off-peak hours, comparing MPC agent performance to normal operation on hot summer days. The test homes are listed in descending order of total cost reduction. The homes which did not have any MPC test on hot summer days are excluded. The average daily total costs are listed on the top of each bar. The percentage reduction of total cost and on-peak hour cost between the normal and MPC operation are listed underneath. Across different homes, the average cost savings of using the MPC agent on hot summer days range from 28.72% to 51.306% compared to normal operations. In all cases, the cost savings are contributed by the savings during on-peak hours (cost savings during on-peak hours range from 36.174% to 68.436%) indicating the load shifting effectiveness of the MPC agent.

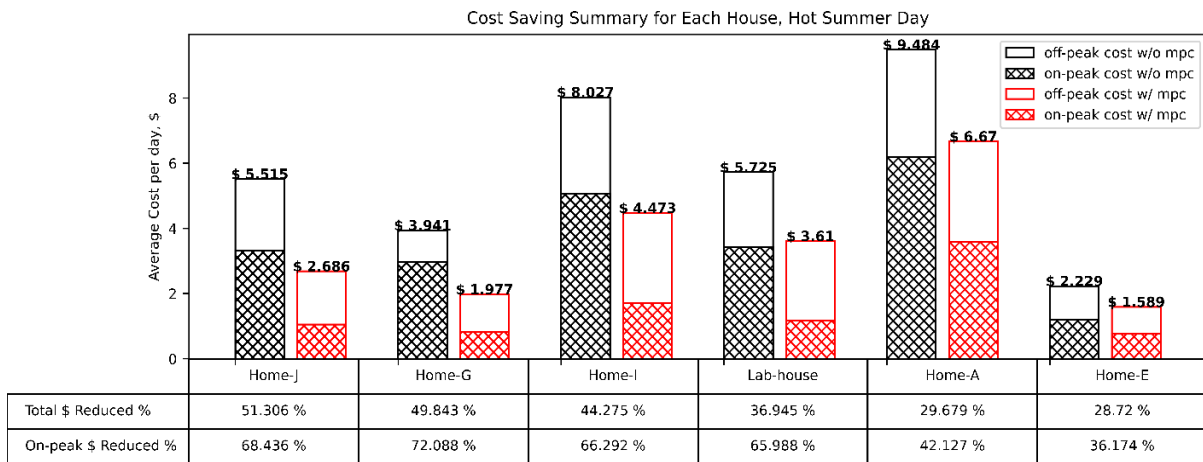


Figure 4.8 Cost saving summary for each home on hot summer days for normal operation days and MPC operation days.

Figure 4.9 is the daily average cost summary of the mild summer days. The legend in Figure 4.8 also applies to this figure. Although significant savings were obtained during on-peak hours, indicating MPC load shifting effectiveness, the total cost savings including on-peak and off-peak hours show a wider range compared with the results for hot summer days, from 6.737% to 60.318% while the main floor of Home C shows negative savings. The smaller saving percentages occur to the three AC units which have very small baseline costs and therefore the savings can be easily concealed by variations of home activities and operation changes. For example, Home C main floor (Home-C down in Figure 11) shows a negative saving. It is because the MPC agent was only allowed to work with a 2-degree temperature band (78-80°F, from 9:00 pm to 6:00 am), and a 3-degree temperature band (rest of the day, 78-81°F) were set. The narrow comfortable temperature band impedes the benefit of the precooling. In addition, the homeowner requested to switch indoor temperature comfort range

upper limit from 81°F to 80°F at 9 pm, leading to excessive AC operation at night, resulting in higher off-peak hour costs (the white space of the bar), and thus higher total costs.

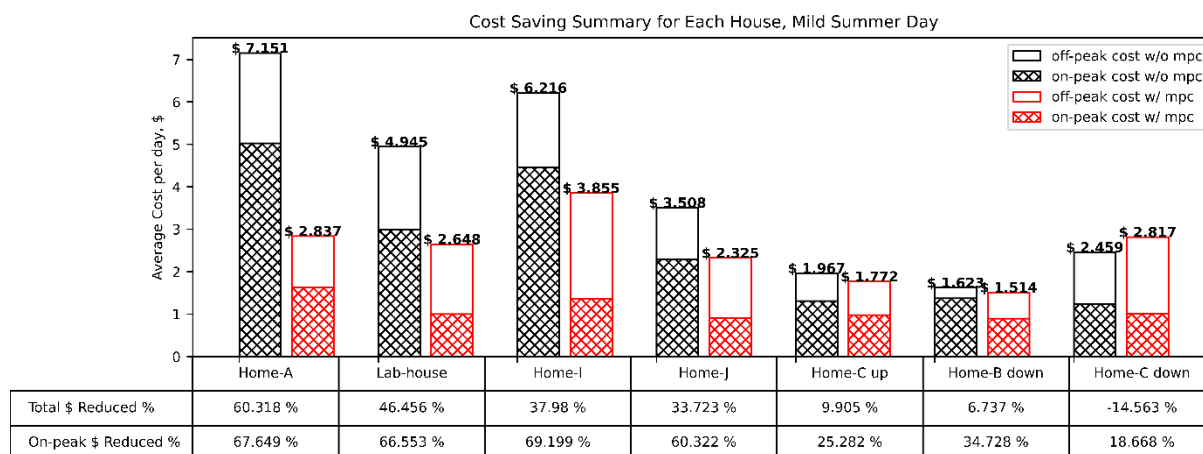


Figure 4.9 Cost saving summary for each home on mild summer days for normal operation days and MPC operation days.

Table 4.3 includes the daily average of total costs, the cost during on-peak hours, energy consumption, and the AC run time during on-peak hours (in percentage) with and without the MPC agent. Table 4.4 shows a detailed summary of the tests on mild summer days. Table 4.5 is a detailed summary of the tests on cold summer days. Despite the colder outdoor condition, cost savings still exist when MPC agents were used on the main floor of Home B and Home D. These dollar savings after MPC agent implementation not only further incentivize homeowners to participate in the DR programs, but also bring a positive impact on grid reliability, as large cost savings during on-peak hours indicates significant load shifting capacity.

Table 4.3 Cost and energy consumption summary for homes on hot summer days.

	Cost, \$		Energy, kWh		Peak Cost, \$		TOU run time, %	
	MPC	normal	MPC	normal	MPC	normal	MPC	normal
Home-A	6.67	9.49	53.50	66.93	3.59	6.19	59.46	93.84

Home-E	1.59	2.23	13.48	17.86	0.77	1.20	22.66	38.06
Home-G	1.98	3.94	17.81	24.53	0.83	2.97	20.92	68.99
Home-I	4.47	8.03	41.65	58.03	1.71	5.08	28.10	76.21
Home-J	2.68	5.51	24.82	41.23	1.05	3.33	29.42	92.41
Lab- house	3.61	5.73	35.40	42.99	1.17	3.43	34.47	92.26

Table 4.4 Cost and energy consumption summary for homes on mild summer days

	Cost, \$		Energy, kWh		Peak Cost, \$		TOU run time, %	
	MPC	normal	MPC	normal	MPC	normal	MPC	normal
Home-A	2.84	7.15	21.92	47.51	1.62	5.03	26.19	80.54
Home-B down	1.51	1.25	18.48	9.89	0.90	1.05	14.63	16.87
Home-C down	2.82	2.46	22.51	17.50	1.01	1.24	26.11	32.10
Home-C up	1.77	1.97	12.12	12.12	0.98	1.31	40.73	54.51
Home-I	3.85	6.03	36.76	39.70	1.37	4.29	23.67	74.92
Home-J	2.33	3.51	21.47	24.71	0.91	2.30	27.13	69.77
Lab- house	2.65	4.94	24.76	36.88	1.00	2.99	30.92	94.67

Table 4.5 Cost and energy consumption summary for homes on cold summer days.

	Cost, \$		Energy, kWh		Peak Cost, \$		TOU run time, %	
	MPC	normal	MPC	normal	MPC	normal	MPC	normal
Home-B down	0.19	0.41	2.95	2.59	0.07	0.38	1.26	6.33
Home-D	0.27	2.43	1.74	19.45	0.20	1.32	4.99	33.31

C. Impact analysis of energy flexibility

In this section, the flexibility factor is used to further assess load shifting potential, i.e., the demand flexibility, for different buildings and operation conditions under the DR programs

after installing the MPC agent. Figure 4.10 illustrates the distribution of flexibility factor FF during all MPC test periods for each home, where a positive median line indicates the MPC agent provide effective load shifting on all test homes. The test result of Home G is not included as the sample size is too small. Although the MPC tests in Home D demonstrate cost savings, its flexibility factor is heavily skewed with a much wider distribution compared with other homes, indicating there is a large proportion of energy consumption occurs during on-peak hours in Home D. This is because Home D's tests were conducted on cold days. In some test conducted on cold days, a few AC operations in the late afternoon including on-peak hours were effective to maintain acceptable thermal comfort, making $\int_{L_{low}} P dt < \int_{L_{high}} P dt$, and thus a negative FF . It indicates that it is not efficient to implement load shifting operations during cold days when AC demand is low.

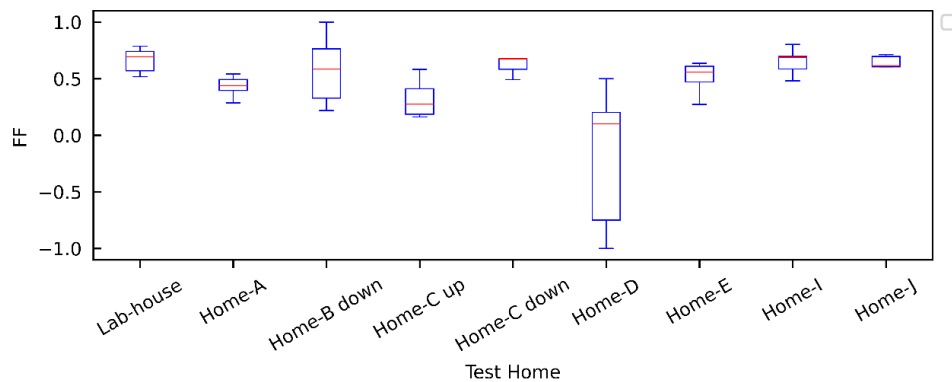


Figure 4.10 Flexibility factors during the entire test period except for extreme days

Therefore, to demonstrate the distribution of FF of Norman test homes versus their floor size, the results for cold days and extreme days (including Home D results) are excluded in

Subplot (a) in Figure 4.11. Notice that the floor area of the two-story homes is divided into two parts: the main floor and the upper floor. It can be seen in Subplot (a), as the floor size of the house increases, its flexibility factor tends to increase, and the variation of the flexibility reduces. A reasonable postulate is that smaller homes have smaller thermal mass, which makes the home more susceptible to external disturbances such as the outside environment, causing a rapid increase in indoor air temperature when the system enters on-peak hours. This trend holds true for other houses except for Home A (floor size 2400 ft^2), whose AC was confirmed with less capability to cool the space due to an undercharge fault, making the load shifting efficiency remains low. The AC's capacity impact on the flexibility factor is further verified in an additional experiment depicted in Subplot (c). If Home A is excluded and a threshold 2000 ft^2 is used, the average flexibility factor of larger houses is greater than that of smaller houses by 11.903%.

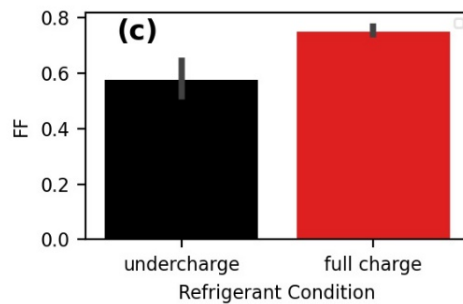
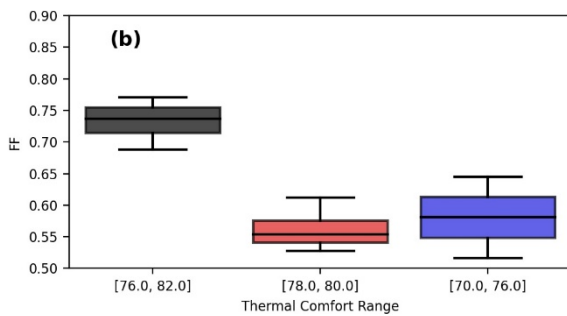
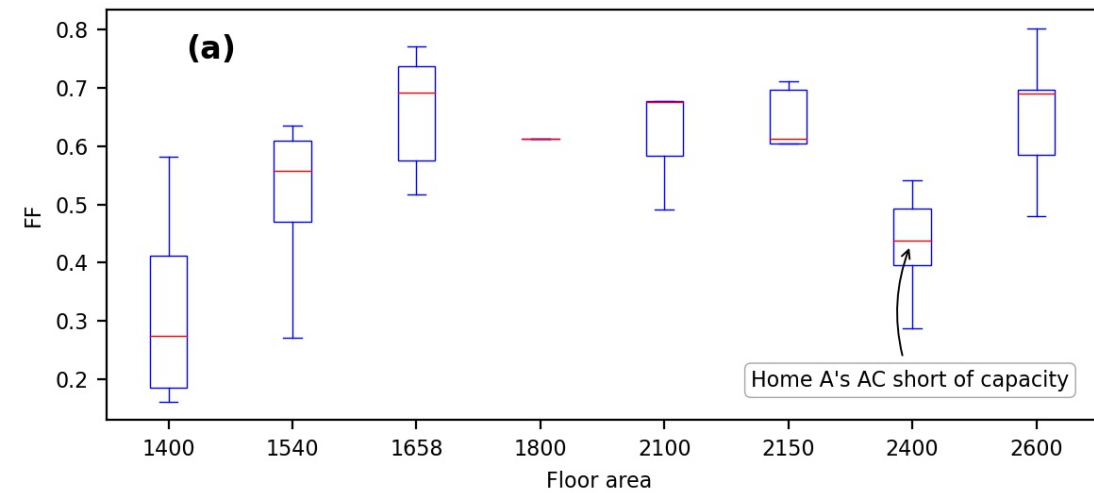


Figure 4.11 MPC agent performance assessment: (a) FF versus floor area in Norman test homes (b) distribution of FF for different temperature bands in the Lab House (c) average FF for different refrigerant charging conditions in the Lab House

Subplot (b) is the distribution of flexibility factors of the MPC agent for different thermal comfort temperature bands used in the Lab House. For a fair comparison, all test results in Subplot (b) are obtained using the data in mild summer day cluster. Three different thermal comfort temperature bands are tested. The tests with the temperature band [76°F, 82°F] provide the most flexibility. On the other hand, tests with the narrow temperature band [78°F, 80°F] show the least flexibility, since both precooling periods on off-peak hours and the free-floating

period (i.e., no or less AC operation period) during on-peak hours are all limited by the smaller band. Finally, the temperature band [70°F, 76°F] is associated with moderate flexibility due to its lower upper bound leading to more energy consumption during on-peak hours.

Although the tonnage of the AC might be suitable for each tested home, the fault and the refrigerant leakage may affect the output of the AC and hence affect the MPC agent results. To verify the flexibility of homes in this situation, an undercharge was purposely generated by withdrawing 30% refrigerant from the Lab House unit to create the MPC test under both full charge and undercharge conditions. Subplot (c) includes the average FF of the MPC agent in the Lab House under different charging conditions within the same temperature band [70°F, 76°F]. When the AC is in full charge, the average flexibility factor improves 23.164%, and the average total cost reduction is about 42.409%, which indicates the AC's ability to cool the space enhances the MPC agent's effectiveness on both load shifting and cost saving.

4.4 MPC-based Precooling at Virtual Testbed

Previous studies have successfully conducted the MPC-based precooling field test in nine real homes. Although in the field test study, it is easier to observe the impact factors that reflect the actual situation and are difficult to consider in the simulation study, field test study has a natural disadvantage: it is not possible to conduct multiple experiments for comparisons under the same condition. This limitation leads to the challenge of determining the impact of MPC hyperparameters. Moreover, due to the complexity of MPC-based precooling, it is natural to raise the question of whether one or more hyperparameters have a relatively larger impact so

development effort can be better allocated based on their impact size.

To answer this question, a virtual testbed is constructed to simulate a real Lab House reaction to the MPC-based precooling control. In addition, to see how MPC-based precooling will affect cost saving and demand flexibility in a virtual testbed, it is also important to know how the HVAC system will be performed under different MPC hyperparameters. The rest of the section is structured as follows. The development, calibration, and validation of the EnergyPlus co-simulation model to reflect the real-world smart thermostat response and on/off actuation of residential HVAC system is first introduced in Section 4.4.1.2. The choice of MPC hyperparameters is presented in Section 4.4.2.2 Finally, the MPC-based precooling results through the proposed virtual testbed are shown in Section.

4.4.1 Virtual Testbed Development

In this study, the virtual testbed of the Lab House in Norman, Oklahoma, is developed using the previous EnergyPlus model study [131]. Although the building characteristics of this model is investigated and calibrated in [131], modifications are needed since the model does not include an HVAC system. More importantly, refinement effort is also needed to correct the way that EnergyPlus models the cycling of direct expansion (DX) coils and the internal thermal mass which affects the indoor air fluctuation.

4.4.1.1 Residential building modelling

The Lab House information can be found in Chapter 3, where a SEER 14 heat pump is

installed. A residential DX system with a cooling coil has 9,000W gross rated cooling capacity, and 2.67 COP is modelled in EnergyPlus. The COP used in this EnergyPlus model, which is smaller than manufacturing rated COP, is chosen to take into account the system losses [132] and 30% refrigerant removal [133]. The energy model is shown in Figure 4.12.

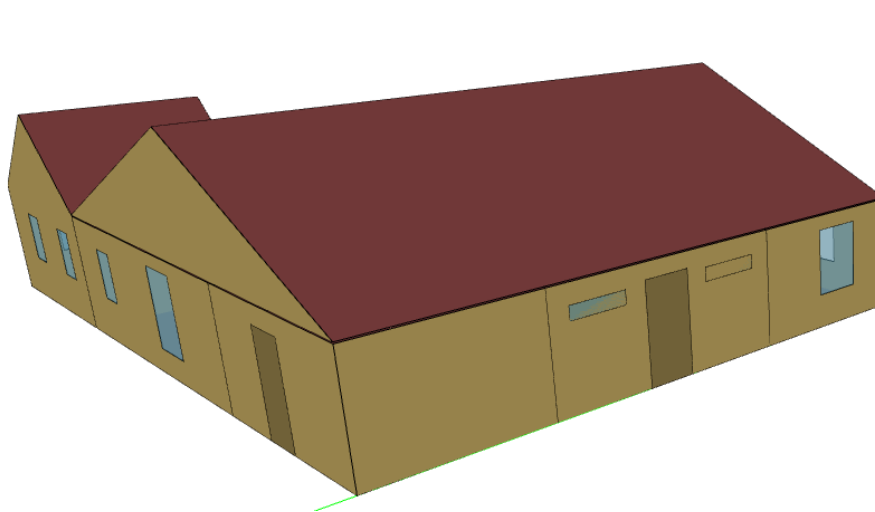


Figure 4.12 Energy model built for the Lab House in Norman, Oklahoma

4.4.1.2 EMS for on/off control with deadband

The refinement effort on correcting the way that EnergyPlus models the cycling of DX coils is needed because the typical way that EnergyPlus models the cycling of DX coils is based on the run time fraction (RTF) of the coil, which is the quotient of the ratio between the actual sensing cooling load and the capacity of the cooling coil (i.e., partial load ratio, PLR) and the compressor cycling efficiency losses (i.e., partial load fraction, PLF) [134]. In other words, the EnergyPlus provides either enough or full energy to maintain the thermostat setpoint perfectly [70,132]. This differs from the real-world operation in a typical residential building in the U.S.,

where a wider temperature range on cooling or heating setpoint are allowed to determine the HVAC system either fully on or fully off.

To mimic the on/off control with a deadband, an Energy Management System (EMS) syntax is added to the EnergyPlus model. EMS is a powerful and flexible control that can be directly interpreted by EnergyPlus by a small programming language called EnergyPlus Runtime Language (Erl), more detail can be found in [135]. In this study, the proposed Erl will assign the heat pump, which includes a DX cooling coil and a fan, to be fully on or off at the beginning of each time step depending on a pre-defined thermostat setpoint and deadband.

Table 4.6 EnergyPlus model variables for on/off control

Purpose	Name	Variable type	Actuated Component Unique Name
On off schedule on heat pump	cycle on off	Constant schedule (be 0 if not overwrite)	
Overwrite the on off schedule	AvailSCH_Overwrite	EMS: Actuator	cycle on off
Monitor Zone Air Temperature	Tin	EMS: Sensor	Zone air temperature
Monitor	Tset	EMS: Sensor	Zone Thermostat Cooling Setpoint Temperature

As shown in Table 4.6, two EMS objects called “*Sensor*” are created to monitor zone indoor air temperature (T_{in}) and thermostat cooling setpoint (T_{set}) at each timestep, and an EMS object called “*Actuator*” is created to overwrite the pre-defined “cycle on off” schedule on the heat pump. Then, an EMS object called “*ProgramCallingManager*” will call the program syntax shown in Figure 4.13 at the beginning of each timestep (i.e.,

BeginTimestepBeforePredictor), and the actuator will only overwrite the “cycle on off” into 1 (i.e., turn on the HVAC system) when the zone air temperature is above $T_{set} + T_{offset}$ (T_{offset} is given for modeling the deadband) until T_{in} reaches $T_{set} + 0.01$. Here, 0.01 is given to avoid runtime fraction calculation to start in Energyplus.

```

!- ===== ALL OBJECTS IN CLASS: ENERGYMANAGEMENTSYSTEM:PROGRAM =====
EnergyManagementSystem:Program,
  HVAC_uncontrolledloop_Supervision, !- Name
  SET Toffset = 0.5, !- Toffset is the deadband in degree Celsius
  IF Tin <= (Tset + Toffset) && AvailSCH_Overwrite == 0, !- Program Line 2
  SET AvailSCH_Overwrite = 0,
  ELSEIF Tin < (Tset+0.01),
  SET AvailSCH_Overwrite = 0,
  ELSE,
  SET AvailSCH_Overwrite = 1,
  ENDIF;

```

Figure 4.13 Proposed Erl for on/off control with deadband

4.4.1.3 *Internal thermal mass modification*

After modifying the model to mimic the on/off cycling behavior, an internal thermal mass also needs to be refined. Although an important heat transfer calculation component, the *InternalMass* object is applied to an interior wall, floor, and ceiling. There are still some underestimated internal masses that may cause a dramatic change in indoor air temperature. This issue has been mentioned in [70,132,136], and in each case, the suggested solution is to use *ZoneCapatianceMultiplier: ResearchSpeical* object, which serves as a correction factor for the zone (moist) air capacitance, to solve it. In this study, through trial and error by comparing EnergyPlus results and the smart thermostat data, the multiplier value 3 is selected.

4.4.1.4 EnergyPlus model validation

In this study, the proposed EnergyPlus model was run with a time step of 5 minutes, which corresponds with the smart thermostat granularity. Figure 4.14 demonstrates the indoor air temperature comparison between actual smart thermostat measurements and the EnergyPlus simulated results. The green dashed line in the figure represents the setpoint of the smart thermostat, and as can be seen from the solid red line (i.e., 73 °F) that the deadband is slightly less than 1 °F. Therefore, the deadband used in the proposed EMS is 0.9 °F (or 0.5 °C). The actual energy consumption obtained from the smart power meter and the simulated energy consumption obtained from the EnergyPlus are also listed in this figure. Although mismatches still exist especially in the late afternoon, the EnergyPlus simulated indoor air temperature matches smart thermostat measurement for the majority of the day, and the energy consumption is relatively close.

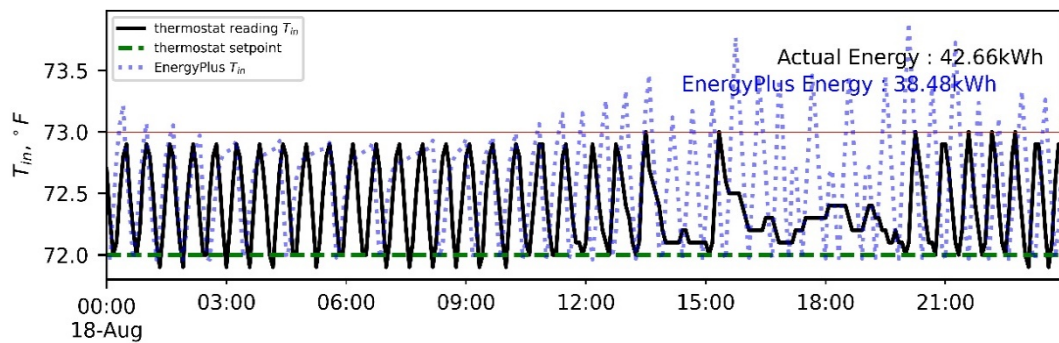


Figure 4.14 Indoor air temperature comparison between actual smart thermostat measurement and EnergyPlus simulation

Figure 4.15 shows the comparison between the simulated daily energy consumption obtained from EnergyPlus (red bar) and the actual daily energy consumption obtained from

the smart power meter (black) for five continuous days in August. The discrepancies between the simulated results and the actual measurements are listed above each pair of bars, where the largest discrepancy is 10.69 kWh and the smallest one is 1.71 kWh. The daily average outdoor air temperatures, as well as the daily maximum and daily minimum outdoor air temperatures, are listed as blue dots. It can be seen from the figure that the maximum discrepancy occurred when the minimum daily outdoor air temperature reached its lowest point during these five days. A reasonable suspension is when the outdoor weather condition has changed, for example, a cold night on August 19, or a cold day on August 21. The performance of the EnergyPlus model may not as good as the performance in a standard hot summer day (e.g., August 20).

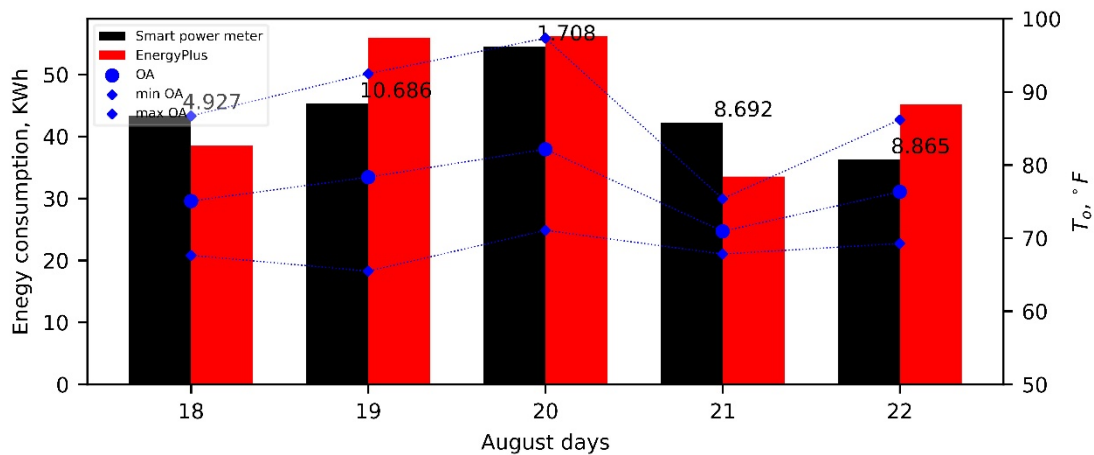


Figure 4.15 Daily energy consumption comparison with outdoor air temperature

4.4.2 Evaluate MPC-based precooling performance at the virtual testbed

Since the EnergyPlus model has been proven that it can provide a realistic indoor air temperature response and mimic the on/off cycle thermostat control with deadband, it can be useful for the MPC-based precooling control investigation. This section presents the MPC-

based precooling control workflow for the virtual testbed, the selection of a hyperparameter for MPC-base precooling control, and the test result analysis for different hyperparameters.

4.4.2.1 MPC-based precooling control workflow

To conduct the MPC-based precooling control using the virtual testbed, a Python script is used to process the required inputs, solve the same mixed-integer linear programming (MILP) problem mentioned in Section 4.2.1, and communicate with the EnergyPlus model. As shown in Figure 4.16, a similar workflow for the MPC agent mentioned in Section 4.2.2 are utilized. Note that three modifications are made for conducting the test in a virtual testbed. First, at each time step, the initial condition T_{in} and T_{ie} are obtained from zone air temperature and the north-facing inside face surface temperature instead of the thermostat and node sensor reading. Second, once the optimal on/off decision is made, different from the smart thermostat in real homes, this decision can be directly implemented by adjusting the actuator status at the virtual testbed through the Python Plugin connected with the EMS to turn on or off the heat pump. Third, the parameters of the model are identified by using the simulated dataset generated by EnergyPlus instead of the real measurements. The weather forecast, TOU rate, and thermal comfort range are the same datasets mentioned in Section 4.3.1.

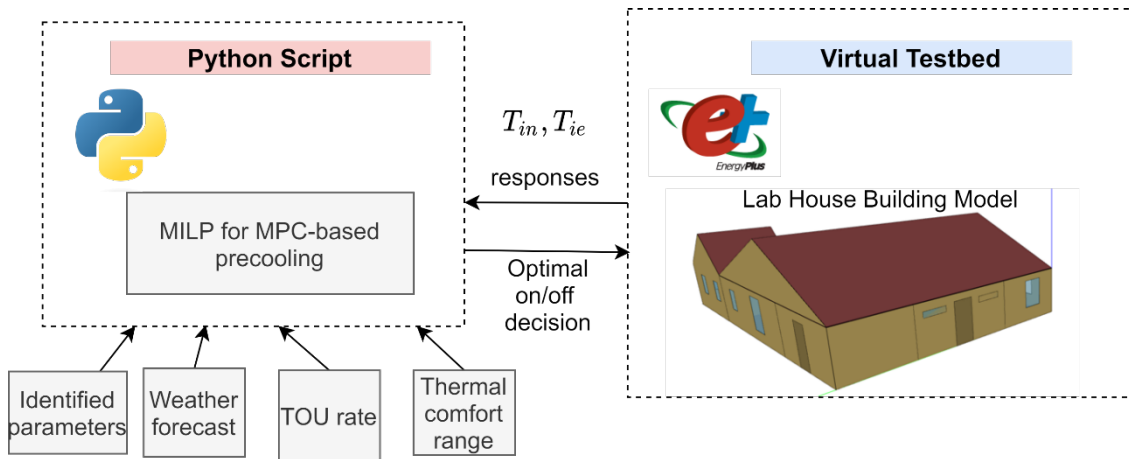


Figure 4.16 MPC-based precooling control through Python plugin and EnergyPlus virtual testbed.

4.4.2.2 MPC hyperparameter choice

Except for the factors that can greatly impact MPC-based precooling test results discussed in Section 4.3.3.2, the MPC hyperparameters, which are consistently chosen in field test results, may also have a different impact on the test results. Their impact is hard to test in the field due to the limitations of real homes but can be tested more conveniently in a virtual testbed. Among various hyperparameters, the model, optimization horizon, and use of forecasted weather are the impact factors discussed in this study. Apart from the concern of making the combinatorial space of hyperparameter too large, another reason for this choice is that some other hyperparameters, such as control sampling time, are less practical for real-world applications (as the smart thermostat typically updates at 5-minute intervals).

The prediction horizon, the usage of forecasted weather, and the model used in impact analysis are summarized in Table 4.7. Here the model used in the MPC-based precooling refers to the RC network model. The parameter identification of these types of models is discussed

in Chapter 3. To further investigate the impact of model accuracy on MPC results, except for the 2R2C model discussed in Chapter 3, a 1R1C model shown in (34), a 2R2C model with reduced AC impact by multiplying c_1 and c_2 by 0.9, and a 2R2C model with reduced thermal mass impact by multiplying τ_1 and τ_2 by 0.9 are also used in MPC-based precooling tests. Figure 4.17 demonstrates an example of a 2R2C model (red), a $2R2C_{reduce_{AC}}$ model, and a $2R2C_{reduce_{\tau}}$ model indoor air temperature and interior wall surface temperature simulation results versus EnergyPlus results.

$$\begin{aligned} \frac{dT_{in}}{dt} = & \frac{1}{\tau}(T_o - T_{in}) + \frac{1}{\tau}(a_1S + a_2S^2) + \frac{1}{\tau}(T_o - T_{in})(b_1W + b_2W^2) \\ & + \frac{1}{\tau}(c_1T_o + c_2T_o^2)u \end{aligned} \quad (34)$$

Although the discussion here did not include all the hyperparameters in the MPC agent, finding the optimal choice for the MPC hyperparameters can be a challenging and time-consuming task. It is beneficial to identify the hyperparameters that have the most significant impact on the precooling results and allocate the development resources accordingly. In this study, we conduct the sensitivity analysis and ablation study for these hyperparameters.

Table 4.7 MPC hyperparameters choice for impact analysis

Hyperparameters	Test values
Optimization horizon	2hr, 4hr, 6hr, 8hr,9hr,10hr,12hr
Forecasted weather	Prefect weather from weather services, Forecasted weather used in filed test, see Section 4.3.1.

Model	2R2C, 1R1C, $2R2C_{reduceAC}$, $2R2C_{reduce\tau}$
-------	---

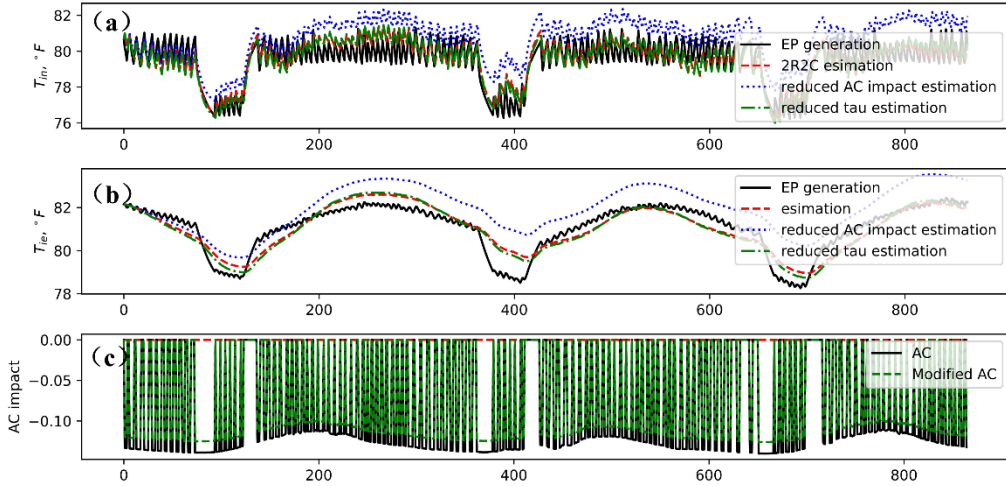


Figure 4.17 An example of a 2R2C model, a $2R2C_{reduceAC}$ model, and a $2R2C_{reduce\tau}$ model simulation results compared with the EnergyPlus outputs. (a) Indoor air temperature results. (b) Interior wall surface temperature results. (c) AC impact and reduced AC impact to the home thermal dynamics.

4.4.2.3 Test results and impact analysis for MPC hyperparameters

This section describes the MPC-based precooling test results at a virtual testbed. An example of an MPC-based precooling test on August 6, 2022, at a virtual testbed, is first demonstrated. Follows by the sensitivity analysis on the optimization horizon. Finally, the ablation study is shown and the dominant hyperparameters are discussed.

Figure 4.18 depicts the zone air temperature (red), north-facing interior wall surface temperature (blue), AC on/off signal (green), thermal comfort range (dashed gray), and outdoor air temperature (dashed blue) of the MPC-based precooling results using forecasted weather and 10-hour prediction horizon at virtual testbed. The shaded area represents the on-peak hour

periods. Similar observations mentioned in Section 4.3.3.1, can be found in this figure: T_{in} mostly stays within thermal comfort range and occasionally excess the T_{ub} and T_{lb} , and the slack variable may not causing the AC turned on immediately due to the consideration of high electricity costs; T_{ie} has much slower respond to the environment change which justifies the 2R2C model; and the MPC agent also takes advantage of cooler early morning air to precool the space with higher efficiency in this virtual testbed result.

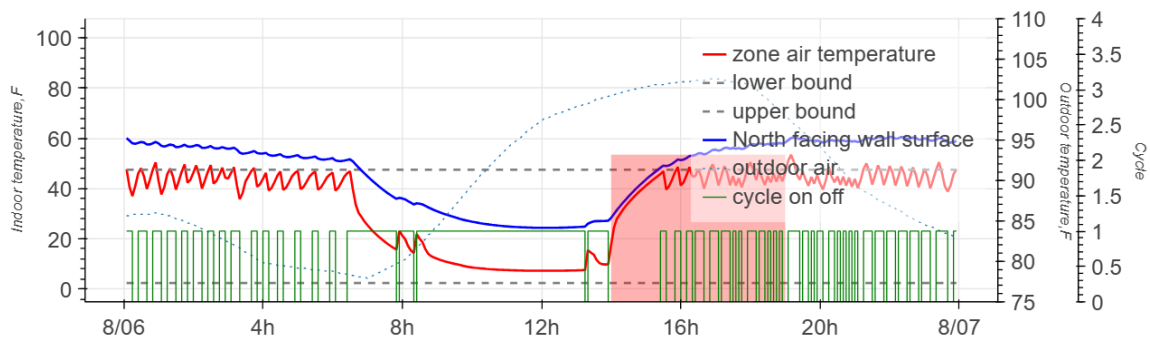


Figure 4.18 MPC-based precooling test result on August 6 at virtual testbed

In an MPC algorithm, the optimization horizon decides how long the MPC can “see” into the future at each time step to make decisions. In MPC-based precooling real-time control framework, choosing a long optimization horizon may also add computational burden and data collection cost since a longer horizon correspondence a longer forecast data and larger decision variable space. Figure 4.19 is the 24-hour AC operation cost with different choice of optimization horizons by using 2R2C model. As can be seen from Figure 4.19, under perfect weather condition, the larger optimization horizon, in most of the case, indicating lower cost. This is because using a longer optimization horizon provides more (perfect weather) information into the MILP problem, and hence the MPC agent can better anticipate changes in

the system and make better decision accordingly. If forecast weather dataset is used in MPC-based precooling, a longer optimization horizon, however, may not be an optimal choice. This is because a longer optimization horizon provides more uncertainty information into the MILP problem and making the MPC agent makes suboptimal decisions at each time step. Therefore, 6-hour horizon provides the best result.

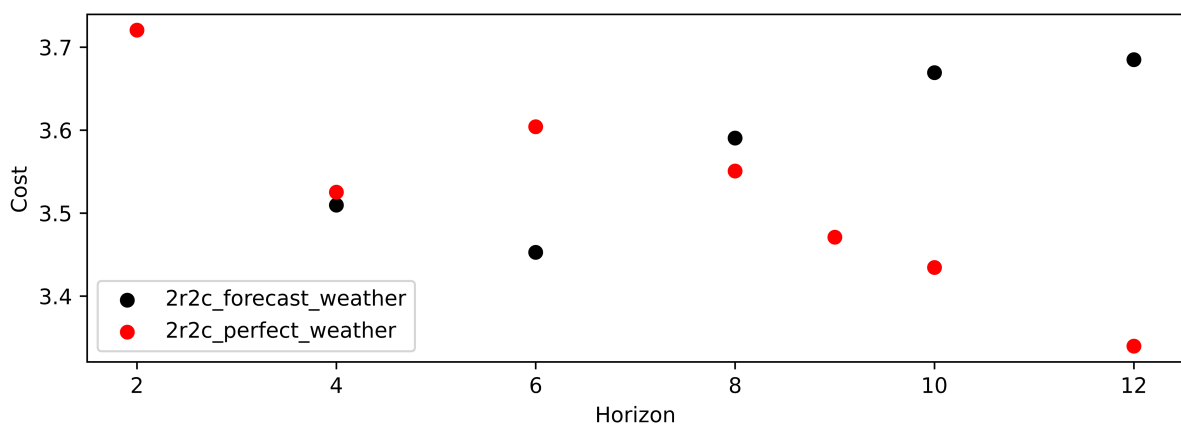


Figure 4.19 Sensitivity analysis on optimization horizon in terms of cost saving with forecasted weather and with perfect weather.

Since the good MPC-based precooling test is the result of the combined effort of all the hyperparameters to identify the contribution of each, an ablation study is performed. Starting from the ideal case, i.e., using a 12-hour optimization horizon, 2R2C model without any modification, and providing perfect weather conditions to the model, each ablation is obtained by removing one component (horizon, model, or weather dataset) from the ideal case. Figure 4.20 compares the accumulated 24-hour AC operation cost of the ideal case (solid green) to the

five ablations and the baseline case (dashed black), i.e., 9-hour optimization horizon, 1R1C model with forecast weather. As can be seen from the figure, switching from the perfect weather to a forecast weather dataset (red line) result in the highest increase in cost (except for the baseline case), and followed by replacing the 2R2C model with a 1R1C model (orange line), while only reducing the prediction horizon under perfect weather (blue line) causes the least increase in cost.

In addition to the ablation study, the comparison study from the baseline case to the ideal case is also performed. Specifically, starting with the baseline case, each case is obtained by replacing one component (horizon, model, or weather dataset) from the baseline case. Figure 4.21 shows a comparison of the accumulated 24-hour AC operation cost of the baseline case (solid black) to the five cases as well as the ideal case (dashed green). Similar to the ablation study, switching from the forecast weather to a perfect weather dataset (blue line) result in the highest cost reduction, and followed by replacing the 1R1C model with a 2R2C model (orange line), while only reducing the prediction horizon under forecast weather (purple line) causes less cost reduction.

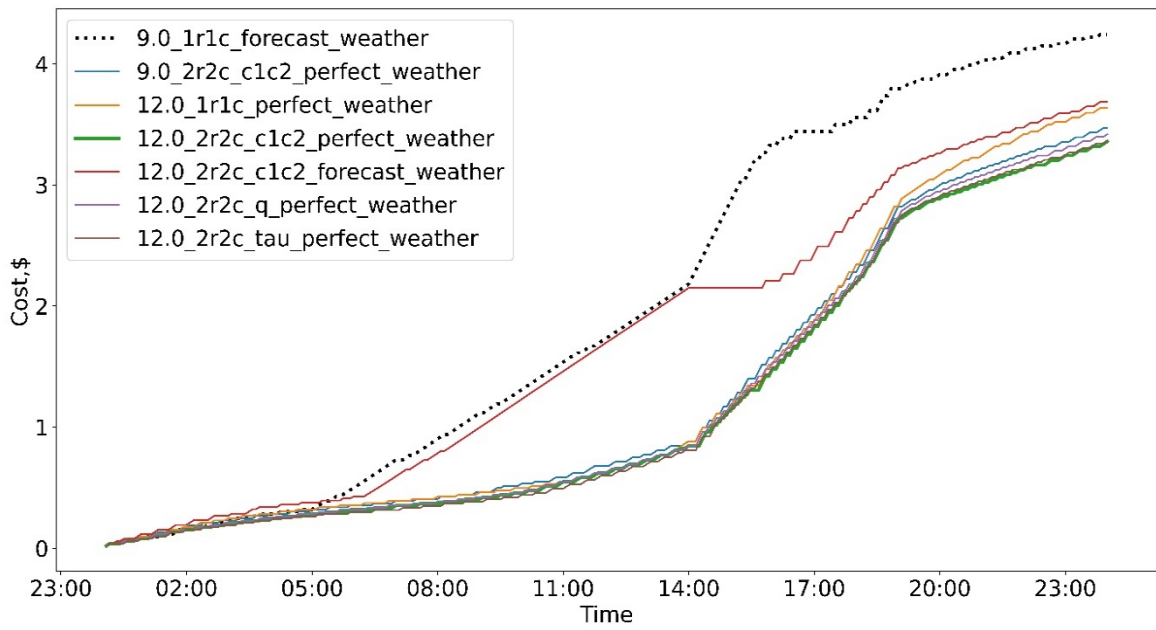


Figure 4.20 Ablation study using cumulative AC operation cost: compare ideal case (bold green) to baseline case (dashed black) and five different ablations.

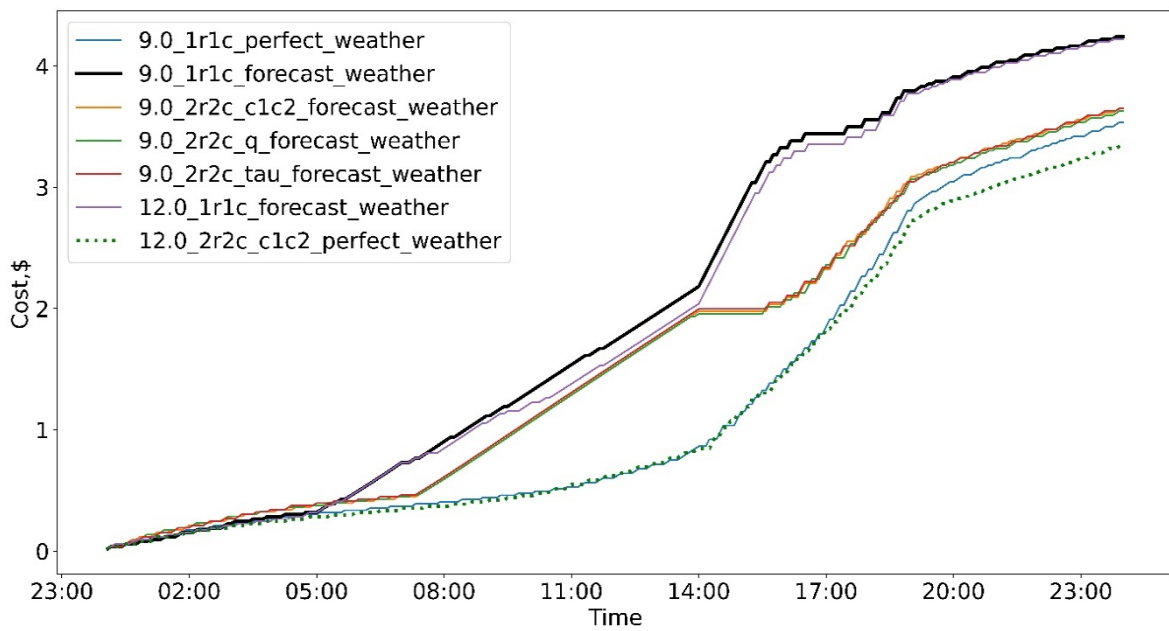


Figure 4.21 Performance comparison using cumulative AC operation cost: compare baseline case (bold black) to idea case (dashed green) and five different cases between them.

4.5 Summary

The MPC-based precooling agent is developed, implemented and analyzed using nine-home fielding testing with the help of IoT devices and cloud-based data platform in terms of the cost saving and energy flexibility potential in different conditions under a TOU rate. The performance of the agent is also studied using whole building modeling tool, i.e., EnergyPlus, and the Python scripts, in virtual testbed in terms of the impacts of hyperparameters.

In the field test study, the MPC agent is implemented in nine residential buildings and the performance of the agent is evaluated through four-month experiment data. Through experiments in multiple buildings, the robustness and effectiveness of the proposed MPC agent are proved. The distribution of the cost savings on hot summer days ranges from 28.720% to 51.306%, while on mild summer days ranges from -14.563% to 60.318% with a much wider range. For a smaller savings amount on mild days, the saving percentages are more sensitive to human activities and operation changes.

Through the comparison of the flexibility factors of the MPC agents, it is observed that homes with larger floor size generally have larger flexibility than smaller homes. Test results show that difference in flexibility can be as much as 11.903% for an average 775 square foot increase. Homes with similar characteristics may not achieve similar load shifting results, and two other major factors, the thermal comfort temperature band and AC's condition, also affect flexibility. Despite having the same upper and lower bound, a temperature band with higher average temperature results in larger flexibility when compared with a temperature band with

a lower average temperature. Diverse setpoint choices under similar weather conditions are also observed in Norman in normal operation, which further emphasizes the importance of research on personal thermal comfort preferences. Meanwhile, some AC faults that may not be easily detected in residential buildings' AC systems, such as low refrigerant levels, affect load-shifting capability. In the Lab House, 30% refrigerant leakage may lead to a 23.164% flexibility reduction and a 42.409% cost increment, which indicates the importance of automated fault detection research in residential buildings.

As with any field study, certain limitations should be acknowledged. Due to the risk aversion of some participants to the AC operation, the MPC agent tests could not be scheduled for too long, causing the full potential of the agent not to be fully realized. Such limitation motivates the development of the virtual testbed. In this study, an EnegyPlus virtual testbed and its control framework for co-simulation purposes is developed. The former is modified by using the EMS and Python script to mimic the on/off cycle in the majority of residential AC systems in the U.S. Through the indoor air temperature compared with the real thermostat measurement and the energy consumption comparison with the smart power meter measurements, the virtual testbed simulation result shows a quite closer behavior to the real Lab House, hence can be useful in impact analysis in this study.

In this study, the MPC hyperparameters' impact is studied at a virtual testbed by conducting sensitivity analysis and ablation study. Through the sensitivity analysis of optimization horizon, it is observed that a longer optimization horizon may bring a smaller cost

under perfect weather conditions. However, once the forecasted weather is used, such conclusion can be different, a relatively shorter optimization horizon—6 hour in our case study—may preferable due to the uncertainty exists in forecast weathers. The ablation study also validates this observation. Among all impact factor studies, the MPC-based precooling is most sensitive to the usage of forecast weather, followed by accuracy of the models, such as using 1R1C model instead of 2R2C, and the proper prediction horizon.

5 Connected Home Energy Management for Grid-Interactive Building

5.1 Introduction

Having developed the model for major flexibility resources in a smart home and proved the load shifting ability of an AC system through both field study and virtual testbed study, the cost saving and energy flexibility potential in a smart home is investigated through simulation in this chapter. Specifically, a scheduling algorithm coordinating thermal appliances (HVAC system and water heater), non-thermal appliances (clothes washer, dryer, and dishwasher), and a rooftop residential photovoltaic (PV) system, with the PV system being optionally coupled with the battery energy management system, are developed. The structure of this chapter is as follows. First, three types of PV systems that are common in residential buildings are introduced. Second, the mixed-integer linear programming (MILP) problems of the scheduling algorithm for smart homes for different types of PV systems are formulated.

To address the existing limitations and barriers in smart home energy management that user behavior is not fully considered, we have focused on an important aspect that has often been overlooked: user preference. Unlike previous studies, we have taken into account the preferences of users regarding non-thermal appliances. Our investigation is based on an analysis of the American Time Use Survey (ATUS) data, which provides valuable insights into American daily activities and appliance usage. By incorporating user preference in our scheduling algorithm, we aim to contribute to a more holistic understanding of home energy

management systems and better align them with the needs and preferences of users.

Building upon the aforementioned context, it is important to acknowledge the focus of this study does not involve selling power generated by solar systems back to the grid. This means excess energy generated by the PV is wasted if it cannot be used by the appliances or does not need to be stored in the battery, and the user will not get the credits from the grid. It is worth mentioning that such consideration is not unique to this study. For instance, a previous work [137] also ignored the selling excess solar energy options. The author in [137] indicated that feed-in tariffs are not widely available in the U.S., and when the battery is available to collect the majority of the excess PV power, the low feed-in tariff makes selling action less economical. There are three types of solar systems that are commonly used in residential buildings: grid-connected, grid-off, and hybrid systems, as shown in Figure 5.1.

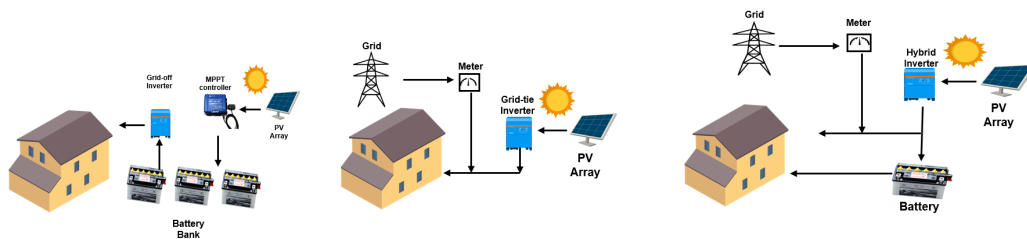


Figure 5.1 Three different types of PV Systems: Off-grid (left), Grid-tie (middle), and Hybrid System (right).

Figure 5.1 also demonstrates the power flows associated with three types of systems. As the name implies, the off-grid system is not connected to the grid, relying entirely on solar generation. In an off-grid system, the energy generated by the solar panels is stored in a battery

bank, which then supplies all the energy to the house when it is needed. The off-grid solar system price is much higher than the other two types, as they require a much larger energy storage system. While a charger controller is embedded into the system to optimize energy generation, this type of system is more frequently used in cabins, cottages, and even RVs (recreational vehicles), and it is especially useful for remote areas where access to the grid is limiting.

Grid-tie systems are more frequently used in residential construction. In this system, the smart home can draw electricity from the grid while also drawing electricity from the photovoltaic system. A grid-tie inverter will convert the solar-generated direct current (DC) to alternating current (AC). However, no battery is involved in this case. To further integrate the battery system with the photovoltaic system, a hybrid system integrates a battery with the photovoltaic system by using a hybrid inverter that not only converts DC to AC but also charges the battery is considered. Therefore, the smart home with a hybrid system obtains energy from the grid, directly from the photovoltaic panel, and from the battery. Both grid-tie and hybrid systems provide more energy flexibility and resilience to the buildings. In this study, only the grid-tie and the hybrid solar system are considered.

5.2 Smart Home Scheduling Formulation

In this section, the mathematical programming formulation for smart home optimal scheduling is introduced. First, a baseline model which considered the thermal and non-thermal appliances scheduling without any PV or battery involvement in the modeling is introduced.

The baseline model depicts an optimal scheduling approach to coordinate both thermal and non-thermal appliances. Following that, a scheduling model integrating the grid-tie and hybrid PV system are introduced, respectively. The first models considered the operation of smart home appliances when the PV panels directly provide some energy to the appliances through a grid-tie inverter, underscoring the immediate solar power utilization. The later model has an AC battery that can store extra PV generation and provide energy at a later time, stressing the advances in energy storage and flexible power supply. Lastly, based on the American Time Use Survey and real-time water usage data, the user preference models, which relies on statistical analysis and data mining technique, simulate household activities related to the timing of appliance use and water consumption are demonstrated.

5.2.1 Baseline model

To model the operation of each appliance, we use a series of binary variables to represent their on/off control signals throughout the scheduling period. Let $i \in \{I = 1, 2, \dots, N_{nt}\}$ denote the number of all non-thermal appliances in a smart home, $j \in J = \{1, 2, \dots, n_i\}$ denote sequential energy phase of each appliance i , and $k \in K = \{1, 2, \dots, m\}$ denotes each time slot of scheduling period, of which N_{nt} , n_i , and m are the number of non-thermal appliances, the energy phase of non-thermal appliance i , and the total time slots, respectively. In this baseline model, we first introduce a constraint applied to all appliances and scenarios—the electricity safety model, which is given by:

$$P_{\text{allappliances}}^k \leq P_{\text{limit}} \quad \forall k, \quad (35)$$

where the $P_{\text{allappliances}}^k$ is the total power usage of all appliances at time slot k , and P_{limit} is the maximum power usage. This constraint is rigorously maintained regardless of whether additional power flows into the home from the PV or battery system or not.

5.2.1.1 *Non-thermal appliances*

Non-thermal appliances are a group of appliances that can be turned on at any time of the day. However, once they are turned on, the output and run time depend on their **energy phase**, which is a subtask in the operation of a non-thermal appliance that runs for a period and consumes a specified amount of energy. For instance, washing clothes is typically divided into three phases: washing, rinsing, and extraction. Hence, the clothes washer has three energy phases. Different energy phases of the same appliance may consume a different amount of power and complete operation at a different period [96].

The decision variable which represents the on/off signal for energy phase j for appliance i at each time slot k is x_{ij}^k . Besides, we use the auxiliary variables s_{ij}^k to aid the construction of constraints to regulate the on/off state of non-thermal appliances. Constraint (36) is used to regulate the operation time for each energy phase j of appliance i , where t_{ij} operation time for the energy phase, which depends on the appliance's specification.

$$\sum_{k=1}^m x_{ij}^k = t_{ij} \quad \forall i, j, \quad (36)$$

To further regulate the operation of energy phases, the following constraints, which

utilize the binary nature of x_{ij}^k and its auxiliary variable s_{ij}^k , are used.

$$x_{ij}^k + s_{ij}^k \leq 1 \quad \forall i, j, k, \quad (37)$$

$$x_{ij}^{k-1} - x_{ij}^k \leq s_{ij}^k \quad \forall i, j, k = 2, \dots, m, \quad (38)$$

$$s_{ij}^{k-1} \leq s_{ij}^k \quad \forall i, j, k = 2, \dots, m, \quad (39)$$

$$x_{ij}^k \leq s_{ij-1}^k \quad \forall i, j = 2, \dots, n_i, k = 1, \dots, m, \quad (40)$$

$$0 \leq \sum_{k=1}^m (s_{ij-1}^k - (x_{ij}^k + s_{ij}^k)) \leq D_{ij} \quad \forall i, j = 2, \dots, m, \quad (41)$$

$$x_{mj_0}^k \leq s_{nj-1}^k \quad \forall k. \quad (42)$$

where constraints (37) and (38) force the auxiliary variable s_{ij}^k to be 1 once energy phase - j of appliance i turns off (i.e., x_{ij}^k to be 0 and x_{ij}^{k-1} to be 1). Based on constraints (37) and (38), constraint (39) ensures each energy phase of the appliance only executed once. The combination of these three constraints along with the constraint (36) ensures the energy phase j is uninterruptable. Similarly, constraint (40) ensures energy phase j turns on only after the previous energy phase turns off. Constraint (41) regulates the time delay between energy phase j and $j - 1$ should be no less than time slots D_{ij} . Moreover, constraint (42) ensures that the first energy phase j_0 of appliance m turns on only after the last energy phase j_{-1} of appliance n , which aims to mimic real-life scenarios such that the clothes washer not running after the dryer. Finally, by using a series of binary sequences TP_i^k , the user time preference constraint on each appliance i is shown in (43). In this study, the user time preference, which specifies the time interval (referred as U) within which appliance i should be operated, is

assumed to be provided by the user. given by:

$$x_{ij}^k \leq TP_1^k \quad \forall i, j, k. \quad (43)$$

The choice of this non-thermal appliance usage preferred timing is discussed in Section 5.2.4. In addition, in [96], the power consumption of each appliance is one of the problem's decision variables, implying that once the appliances are turned on, the algorithm can determine the power consumption of each appliance at each time slot. However, controlling the exact power usage for some of the appliances is a hard operation that may lead to some mechanical issues. We only consider manipulating the on and off for each appliance in this study. As a result, we only control the on/off of the non-thermal appliances.

5.2.1.2 HVAC system

For an HVAC system in the smart home, the on/off signal for each time slot u_{hvac}^k is a binary variable whose output depends on the home thermal dynamics. To accurately capture the home thermal dynamics, a simplified second-order thermal network model which is introduced in Chapter 3 is adopted, as expressed by

$$\frac{T_{ie}^{k+1} - T_{ie}^k}{\Delta t} = \frac{1}{\tau_1} (T_o^k - T_{ie}^k) + \frac{1}{\tau_2} (T_{in}^k - T_{ie}^k), \quad (44)$$

$$\begin{aligned} \frac{T_{in}^{k+1} - T_{in}^k}{\Delta t} = & \frac{1}{\tau_3} (T_{ie}^k - T_{in}^k) + \frac{1}{\tau_3} (a_1 S^k + a_2 (S^k)^2) \\ & + \frac{1}{\tau_3} (T_o^k - T_{in}^k) (b_1 W^k + b_2 (W^k)^2) \\ & + \frac{1}{\tau_3} (c_1 T_o^k + c_2 (T_o^k)^2) u_{hvac}^k, \quad \text{for all } k \end{aligned} \quad (45)$$

where k denotes time slot, T_{ie}^k is the interior wall surface temperature, T_{in}^k is the

indoor air temperature, T_o^k is the outdoor air temperature, S^k is the solar irradiation, W^k is the wind speed, and $u_{hvac}^k \in \{0,1\}$ is the AC on/off signal. Note that (46) and (47) contain a total of nine unknown parameters: τ_1 is the time constant of the building envelope, τ_2 is a coefficient associated with the thermal resistance of air and thermal capacitance of the building envelope, τ_3 is the time constant of the indoor air, and $a_1, a_2, b_1, b_2, c_1, c_2$ are coefficients representing the impact of solar irradiation, wind speed, and AC output, respectively. Also note that in general, the AC output mainly depends on both T_o^k and indoor air web bulb temperature. However, the latter has relatively small variation compared to T_o^k and, thus, is neglected.

The indoor air temperature is regulated by the following constraint during the scheduling period, as expressed by

$$\begin{aligned} \underline{T}_{hvac} \leq T_{in}^k &= \frac{\Delta t}{\tau_3} (T_{ie}^{k-1} - T_{in}^{k-1}) + \frac{\Delta t}{\tau_3} (a_1 S^{k-1} + a_2 (S^{k-1})^2) + \\ &\frac{\Delta t}{\tau_3} (T_o^{k-1} - T_{in}^{k-1}) (b_1 W^{k-1} + b_2 (W^{k-1})^2) + \frac{\Delta t}{\tau_3} (c_1 T_o^{k-1} + c_2 (T_o^{k-1})^2) u^{k-1} + \quad (48) \\ T_{in}^{k-1} &\leq \bar{T}_{hvac} \quad \text{for } k = 1, 2, \dots, m-1, \end{aligned}$$

where \underline{T}_{hvac} and \bar{T}_{hvac} are the acceptable lower and upper temperature bounds on indoor air temperature T_{in} .

5.2.1.3 Water heater system and its coupling effect

Generally, water heaters have a higher power rating compared to other electrical appliances in home. Furthermore, hot water is utilized in many other appliances: for instance, clothes washer withdraws hot water from the hot water tank. Consequently, accurate prediction of water heater behavior, considering these interconnected thermal impacts (i.e., coupling

effect), is crucial for the efficient operation of all appliances. The RC network concept introduced in the HVAC system is also widely used in the water heater modelling. In addition, the physical model can be easily modified to model the coupling effects between the water heater and other appliances. In this study, the modelling method in [105] is adopted and modified to capture the dynamics of the hot water inside the tank.

For the electric water heater, the on/off signal at each time slot u_{wh}^k is a binary decision variable whose output is depending on its thermodynamic behavior. It is assumed that once the clothes washer is in the washing or rinsing phase, i.e., $x_{cwj}^k = 1$ for $j = \text{washing or rinsing}$, the hot water is withdrawn from the tank, and the tank is replenished with the same amount of city cold water. Additionally, we assume the water heater is exposed to the indoor air temperature, and the changing indoor air temperature T_{in} influences the thermal leakage of the tank. The modified model is discretized as follow:

$$\begin{aligned} \frac{T_{wh}^{k+1} - T_{wh}^k}{\Delta t} &= \frac{1}{\alpha} \left(\frac{A}{R} (T_{room} - T_{wh}^k) \right) \\ &+ \frac{1}{V} \left(\sum_{cwj}^{n_{cw}} x_{cwj}^k W_{cwj} + W_{wh}^k \right) (T_{inlet} - T_{wh}^k) \\ &+ \frac{1}{\alpha} (P_{wh} u_{wh}^k) \quad \text{for } k = 1, 2, \dots, m - 1. \end{aligned} \quad (49)$$

where T_{wh}^k is the temperature of the water inside the tank at time slot k ; α is the lumped parameter equal to $\rho C_p V$; ρ , C_p , V , A , and R are the density of water, the specific heat of water, the tank volume, surface area, and its thermal resistance, respectively; In this study, we assume that once the hot water is withdrawn, cold water with temperature T_{inlet} will enter the

tank, and T_{room} represents the air temperature of the room that the water heater exposure to, which is a relatively constant value since the water heater usually installed in a utility room instead of living space. W_{cwj} represents the water usage for clothes washer at energy phase j and W_{wh} represents a typical water usage in a home, which is further discussed 5.2.4. In addition, like the HVAC system, the water temperature is also regulated within a thermal comfort temperature range, \underline{T}_{wh} and \bar{T}_{wh} which are assumed to be given:

$$\underline{T}_{wh} \leq T_{wh}^k \leq \bar{T}_{wh} \quad \text{for } k = 1, 2, \dots, m - 1. \quad (50)$$

5.2.2 Grid-tie PV system

To incorporate the PV panel and the home appliances, we added two more decision variables into the model. P_{pv}^k denotes the power from the photovoltaic panel array to the home appliances at each time slot, while P_{grid}^k denotes the power the homeowner must buy from the grid. In this study, the author assumes the homeowner did not get credits from the grid, implying the homeowner will not sell their energy to the grid. As a result, the grid-tied PV panel model becomes:

$$\min_{\text{cost}} = \sum_{k=1}^m c^k P_{grid}^k \quad (51)$$

$$P_{grid}^k + P_{pv}^k = \left(\sum_{i=1}^{N_{nt}} \sum_{j=1}^{nt} (x_{ij}^k P_{ij}) + u_{wh}^k P_{wh} + u_{hvac}^k P_{hvac}^k \right) \forall k \quad (52)$$

$$0 \leq P_{pv}^k \leq P_{max}^k \forall k \quad (53)$$

$$P_{grid}^k \geq 0 \forall k \quad (54)$$

Constraint (52) denotes the energy usage of the appliances comes from either the direct energy from PV or the grid. Constraint (53) restrict the minimum and maximum amount of energy that PV system can generates. Here, the P_{max}^k is generated by using PV power generation forecast model described in Chapter 3. Note since the energy losses from the inverter and wires has already considered in PV power generation forecast model, the author did not put efficiency parameter in the constraints. Combines the Constraint (54) and the left-hand side in Constraint (52), if the PV panel generates more energy than the home appliances needed, the excess part will be wasted. In fact, the following inequality constraint is considered: $0 \leq P_{pv}^k \leq \min \{P_{max}^k, (\sum_{i=1}^{N_{nt}} \sum_{j=1}^{nt} (x_{ij}^k P_{ij}) + u_{wh}^k P_{wh} + u_{hvac}^k P_{hvac}^k)\}$. Note, since the PV panel generates energy, the utility bill calculation is different from baseline model. In this model, the cost depends on the P_{grid}^k and the utility rate in different time step (Constraint (51)). Here we assume the power flow for each equipment is constant during each time slot, hence c^k is not only the utility rate (\$/kWh), but the utility rate per time interval.

5.2.3 Hybrid PV system with AC battery

The second integration model in this study is the hybrid PV system. Here, we will assume that a hybrid inverter is available to transfer energy from the PV system to the home or to (with-)draw energy from the AC battery. In this case, at each time step k , the home appliances have three distinct energy sources: the direct PV panel generation $P_{pv_{toloa}}^k$, the energy buying from the grid P_{grid}^k , and the possible battery discharge P_{dis}^k . The following model illustrates such power flow between the house (all appliances), the PV, the battery, and the grid:

$$\min f_{cost} = \sum_{k=1}^m c^k P_{grid}^k \quad (55)$$

$$P_{pv_{toloa}}^k + \eta_d P_{dis}^k + P_{grid}^k = \left(\sum_{i=1}^{N_{nt}} \sum_{j=1}^{nt} (x_{ij}^k P_{ij}) + u_{wh}^k P_{wh} + u_{hvaca}^k P_{hvaca}^k \right) \forall k \quad (56)$$

$$E_b^k - E_b^{k-1} = (P_{charge}^k - P_{dis}^k) \Delta t \quad \forall k = 2, 3, \dots, m \quad (57)$$

$$P_{charge}^k / \eta_c + P_{pv_{toloa}}^k = P_{pv}^k \quad \forall k \quad (58)$$

$$SOC_{min} \leq \frac{E_b^k}{C_b} \leq SOC_{max} \quad \forall k \quad (59)$$

$$0 \leq P_{charge}^k \leq L_c C^k \quad \forall k \quad (60)$$

$$0 \leq P_{dis}^k \leq L_d D^k \quad \forall k \quad (61)$$

$$E_b^0 = SOC_{min} * C_b \quad (62)$$

$$0 \leq P_{pv}^k \leq P_{max}^k \quad \forall k \quad (63)$$

$$C^k + D^k \leq 1 \forall k \quad (64)$$

$$C^k, D^k \in \{0,1\} \forall k \quad (65)$$

In this model, the objective function is minimizing the electricity cost (Constraint (55)), which is determined by the utility rate per time step and the energy drawn from the grid. Constraint (56) refers to the balance of the power flow of the smart home, where the energy comes from PV, battery and the grid. Constraint (57) represents the energy balance inside the battery, where E_b^k is the amount of energy left in the battery at the end of time slot k and Δt is the time interval for each time step, and P_{charge}^k is the PV panel's power which stores to the battery. Note that it might be harmful to the battery start with zero energy at the beginning of the day, therefore, we assumed that the initial state of the battery is not empty (Constraint (62)). Constraint (58) indicates the total power flow from PV eventually go in two directions: one part of the power goes to the house load; another part goes to the battery system. Besides, the total power flow from the PV panel to the battery and the appliances should not exceed the estimated maximum amount of PV generation, which is limited in Constraint (63). Constraint (59) restricted the state of charge of the battery should be in a certain range, which helps extend the battery life cycle. The battery (dis-)charge typically has a maximum power flow and continuous power flow, with the maximum power flow typically being a high value that lasts only 3 to 10 seconds. Thus, in Constraint (60) and (61), L_c, L_d indicates the continuous power flow within the time step k . Finally, two binaries C^k, D^k are introduced to prevent the charging and discharging process happening simultaneously (Constraint (64), Constraint (65)).

5.2.4 *Occupancy behavior and user preference model*

Residential buildings serve as habitats for homeowners, providing comfort, safety, and the backdrop for daily life. Their primary function isn't just to provide physical shelter, but rather to facilitate an environment conducive to human wellbeing and satisfaction. Besides, as mentioned in Chapter 1, many researchers indicated that user might have intended to maintain their appliance usage routine even if they enrolled in DR programs. In this context, the human element becomes an important factor in the research of smart home energy management. In this study, the proposed models and management strategies can be designed to be not just technically efficient, but also practically adaptable and respectful of the users' routines and lifestyles.

5.2.4.1 *User time preference on appliance usage*

In this research, the American Time Use Survey (ATUS) is used to build a user preference appliance usage model. The ATUS is a high-quality dataset, which focuses on recording homeowners' time allocation across daily activities, collected annually by the U.S. Bureau of Labor Statistics [138]. Figure 5.2 demonstrates an example of the ATUS dataset. In this dataset, 'TUCASEID' represents a unique ID for different households, while 'TUACTIVITY_N' characterizes various activities including food preparation and laundry. The "TUSTARTTIM" and "TUSTOPTIM" further indicate the start and stop time of each activity. The ATUS data contains a wealth of information, including the homeowner information and work time. More details can be found in [138].

TUCASEID	TUACTIVITY N	TUSTARTTIM	TUSTOPTIME
20210101020210	1	04:00:00	07:00:00
20210101020210	2	07:00:00	07:30:00
20210101020210	3	07:30:00	08:00:00
20210101020210	4	08:00:00	12:00:00
20210101020210	5	12:00:00	13:30:00
20210101020210	6	13:30:00	17:30:00
20210101020210	7	17:30:00	18:00:00
20210101020210	8	18:00:00	19:00:00
20210101020210	9	19:00:00	21:00:00
20210101020210	10	21:00:00	04:00:00

Figure 5.2 An example of ATUS dataset.

A substantial amount of data has been analyzed for the purpose of this study, including the activity time allocation of 228,455 households (over 4,440,000 records). The dataset spans nearly two decades, i.e., from 2003 to 2021. The analysis of a large amount of the ATUS data set allows for a thorough exploration of American appliance usage patterns and contributes to the development of an effective smart home energy management model.

While the ATUS provides a high-quality dataset, additional work has to be done to transfer its information into the useful information can be used in proposed scheduling framework. In this study, a Probability Density Function (PDF) is generated through statical analysis for kitchen cleaning and laundry. To generate the PDF of these activities, the 24-hour scheduling period was divided into 10-minute intervals. The start and stop times of each activity were examined to calculate the number of ongoing activities within each interval. As a result, PDFs were generated to represent the frequency of activity occurrence across the different time slots. Figure 5.3 demonstrates generated PDFs for kitchen clean-up and laundry activities. Notices that kitchen cleans up, although it may not precisely equal dishwasher use time, the

approximation is sufficient for this study—estimate a user’s time preference towards using the dishwasher.

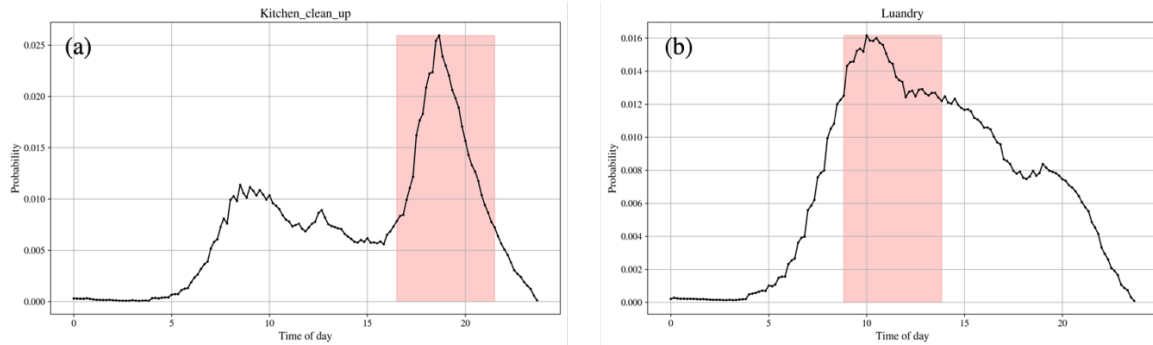


Figure 5.3 Example of the probability density function of (a) Kitchen clean up and (b) Laundry.

To transfer the PDFs to the user’s time preference, a sliding window is employed. The sliding window scans across the probability density function, capturing the maximum area under the curve within its span. The red-shaded areas in Figure 5.3 illustrate scenarios where 5-hour windows generate the highest areas under the curve (PDF curves) for kitchen and laundry activities. To capture the diverse user preferences regarding appliance use time, sliding windows of various lengths are used, representing a range of relaxed to stringent time constraints. For example, larger windows might signify more relaxed preferences, whereas narrower windows could suggest that users might be stricter on appliance use time.

5.2.4.2 *Water usage profile*

To generate a typical water usage profile, a real-world hot water usage profile from [139] is used. The dataset was collected in a single-family house in 2018 by Raspberry Pi. Although the data only captures the time and hot water flow rate, it is still extensive. To extract a typical

hot water usage pattern, a data mining approach—matrix profiling of data snippets—is utilized. The time series data snippets method has already been used in a more complicated clustering task in Chapter 4, it will not be reiterated here. As shown in Figure 5.4, the raw data spans from January to August. The focus of this research is the cooling season, only data from July and August are included in the calculations.

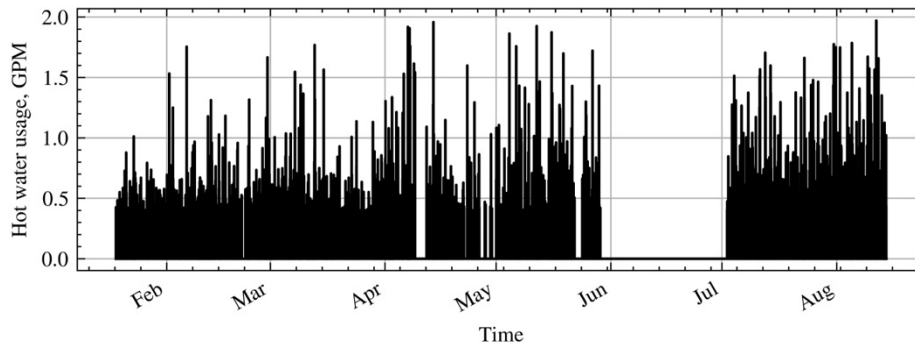


Figure 5.4 Hot water usage data collected from a single-family water heater [139].

5.2.5 User-preferred thermal comfort range

To quantify thermal comfort and account for temperature changes, the following constraints are used when multi-objective optimization is considered.

$$T_{hvac} - T_{in}^k \leq -M_1 z_1^k \quad \forall k, \quad (66)$$

$$-\bar{T}_{hvac} + T_{in}^k \leq -M_2 z_2^k \quad \forall k, \quad (67)$$

$$z_1^k, z_2^k \in \{0,1\} \quad \forall k, \quad (68)$$

where M_1 (M_2) is a constant number that indicates the distance between the acceptable lower (upper) temperature bound and the preferred lower (upper) temperature bound; z_1^k and z_2^k are two sets of binary decision variables to distinguish between preferred and acceptable

temperature ranges. To better understand how z_1^k and z_2^k work, the concept of reward $z_1^k + z_2^k$ is considered. Three examples with different rewards are shown in Figure 5.5, where the x-axis represents different time slot, and the y-axis is the indoor air temperature. As illustrated in the figure, z_1^k and z_2^k will change their values based on where the indoor air temperature drops over the different temperature ranges of time slot k . If the indoor air temperature is between \underline{T}_{hvac} and $\underline{T}_{hvac} + M_1$, but not within the preferred range, the agent will get a normal reward, i.e., $z_1^k = 0, z_2^k = 1$ in this scenario, as shown in Example 1. Similarly, if indoor air temperature is between \bar{T}_{hvac} and $\bar{T}_{hvac} - M_2$, the reward is still the same, and $z_1^k = 1, z_2^k = 0$ as shown in Example 2. Once the temperature falls within the preferred range, the agent gets the highest reward, i.e., $z_1^k + z_2^k = 2$, as shown in Example 3. In this study, it is assumed that both the acceptable and preferred lower and upper temperature bounds are given (i.e., \underline{T}_{hvac} , \bar{T}_{hvac} , M_1 , and M_2 are given).

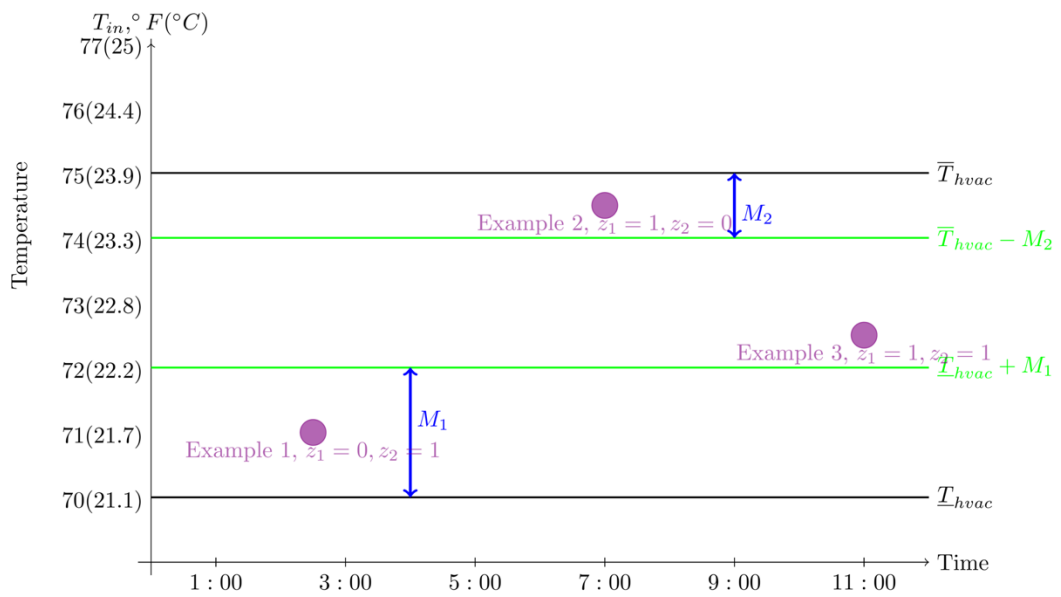


Figure 5.5 Comfort reward examples.

The first objective considered in this study is the cost of electricity, and the second is thermal comfort. The electricity cost is inclusive of all non-thermal appliance operating costs $\sum_{k=1}^m c^k \left(\sum_{i=1}^{N_{nt}} \sum_{j=1}^{n_i} x_{ij}^k P_{ij}^k \right)$, the water heater costs $\sum_{k=1}^m u_{wh} P_{wh}$, and the HVAC system costs $\sum_{k=1}^m u_{hvac} P_{hvac}$.

Besides the electricity cost, we want to maximize the total comfort index $\sum_{k=1}^m z_1^k + z_2^k$. The total comfort index ranges between m and $2m$ since the total number of time slots is m . For example, the optimization agent will earn m points if all temperatures exceed the preferred temperature range, or it may earn $2m$ points if all temperatures remain within the preferred temperature range. Maximizing this objective function can be interpreted as maximizing thermal comfort or maximizing the reward that the agent can obtain over the scheduling period. Moreover, cost and comfort objectives are measured and quantified in different scales. In this study, the goal programming technique is used for resolving the uneven weight problem.

In goal programming, two pairs of deviation variables d_1 and d_2 are introduced. Instead of using either cost or comfort as objective function, with goal programming, the objective function becomes:

$$\min z = \sum_{i=1}^2 w_i (d_i^+ + d_i^-), \quad (69)$$

where w_1 and w_2 are the weights corresponding with electricity cost and thermal

comfort, respectively. To strike a balance between cost and comfort, we add the following constraints to the optimization problem:

$$\sum_{i=1}^2 w_i = 1, \quad (70)$$

$$\frac{f_{\text{cost}}}{\text{target cost}} - d_1^+ + d_1^- = 1, \quad (71)$$

$$\frac{f_{\text{comfort}}}{\text{target comfort}} - d_2^+ + d_2^- = 1, \quad (72)$$

$$d_i^+ \times d_i^- = 0 \quad \forall i, \quad (73)$$

$$d_i^+, d_i^- \geq 0 \quad \forall i. \quad (74)$$

To ensure effective control over the weights of each objective, constraint (70) regulates the sum of weights to be 1. To make sure that the distance between the objective on cost (comfort) and the target cost (target comfort) is normalized, constraint (18) (constraint (72)) is included. The target cost (target comfort) is the objective by solving the single-objective optimization problem. Finally, constraints (73) and (74) indicate that d^+ (d^-) will be zero if the other one is non-negative. By combining (69)–(74), the weights in the objective function represent the user's intention on cost and comfort because the two objectives are normalized and the sum of weights is constrained.

5.3 Case Study

This section provides information and specification for the simulation study. In the simulation, the scheduling period of each appliance is 24 hours with a 10-minute time interval. Hence, if $k \in K = \{1, 2, \dots, m\}$ denotes each time slot of the scheduling period, then $m =$

144. Three non-thermal appliances, namely a clothes washer, a clothes dryer, and a dishwasher, and two thermal appliances, an HVAC system, an electric water heater, a PV panel, and a home battery are considered in the test problem. The water heater and the HVAC system have behaviors that are determined by the governing equations discussed in Chapter 3. Since the data collection, parameter identification, and validation results of the HVAC system and water heater system are introduced in the previous chapter, their parameter identification results are not reiterated in this chapter.

The specifications for the water heater are adopted from [105] and listed in Table 5.1. The left-hand side of the table contains information about the water heater's thermal properties, ambient temperature, inlet water temperature, and power consumption. To ensure the homeowner's comfort, we also implement the constraint that the water heater's temperature must remain within a specified range. In this case study, the water temperature inside the tank must be between $113^{\circ}F$ and $140^{\circ}F$. For the HVAC system, the temperature is allowed to be float between $76^{\circ}F$ to $82^{\circ}F$.

Table 5.1 Other appliances specifications

<i>Water Heater Specifications</i>			<i>Type</i>	<i>Phase</i>	<i>Power (W)</i>	<i>Water (m³/s)</i>
Notation	Value	Unit	Washer	Wash	2000	0.000035
ρ	997.77	kg/m ³		Rinse	2000	0.000026
C_p	4182	kJ/kg°C		Extraction	800	0
$C = \rho C_p V$	1.24 * 10 ⁶	J/°C	Dish washer	Fill & sense	250	0
$G = A/R$	8.12	W/°C		Preheat & wash	1300	0
T_{room}	25	°C		Wash	250	0

T_{inlet}	15	$^{\circ}C$		Partial fill	250	0
T_{wh}^0	62	$^{\circ}C$		Heated rinse	1300	0
P_{wh}	4500	W		Final rinse	250	0
			Dryer	Drying	5000	0

Table 5.1 also lists the specifications of all non-thermal appliances considered in the model. The right-hand side of the table contains information about non-thermal appliances' energy consumption. Three non-thermal appliances were considered: the dishwasher, the clothes washer, and the dryer. Each of them has a unique energy phase, which may consume a different amount of energy. Since the water heater is coupled with the clothes washer; thus, an additional column representing the water consumption of non-thermal appliances is included. Additionally, to ensure safe electricity usage and to prevent too many appliances from running concurrently, the scheduling model includes a constraint on total power flow. In each of the following results, we set $P_{limit} = 8000W$.

For the PV panel, we choose the rooftop PV installed in a real residential building collected from NREL's open dataset [140]. The specification of the PV panel can be seen in Table 5.2. It is assuming the inverter used in both grid-tie and hybrid PV systems has enough capacity—the inverter does not pose a bottleneck for the interaction between the PV panel and the battery. The meteorological data for the PV power generation forecast was obtained from the weather service website [141]. Moreover, interpolation was applied to align the meteorological data with the time interval used in this case study.

Table 5.2 PV panel specifications for test location

Location	PV Size	Module Type	Array Type	Array Azimuth
-----------------	----------------	--------------------	-------------------	----------------------

Latitude: 30.07°N				
Longitude: 90.09°W	5.76 kW	Monocrystalline	Fixed (no tracking)	180°
Array Tilt	Inverter Efficiency			
15°	Up to 96%			

The specification of the battery system is listed in Table 5.3. The specifications for the battery were approximated based on those of a model commonly used in residential energy storage systems, which is a lithium-ion battery [142]. The battery capacity plays an important role in the power flow of the battery. The battery's maximum (dis-)charging power is the maximum power available for a brief period, hence we will not use that directly in the model. Instead, the continuous power (dis-)charging output, L_c, L_d , is used in our model. Additionally, the (dis-)charging efficiency represents the energy lost from the PV panels to the battery. In fact, L_c, L_d is not constant during the time of the day. Instead, L_c, L_d can be affected by many factors. For example, In [143], the author indicated that the solar battery continuous charging power could be correlated with the solar irradiation as well. In this study, this value is based on the rated (continuous) power output and power factor listed in the datasheet in [142].

Table 5.3 Battery system specifications

Notation	Description	Value
C_b	Battery Capacity	3 kWh
L_c, L_d	Continuous (dis-)charging power	1280 W
η_c, η_d	Battery (dis-)charging efficiency	0.9
SOC_{min}, SOC_{max}	Minimum and Maximum State of Charge	20%, 80%

In this case study, the TOU rate (\$0.05 from 11:00am to 1:50pm, \$0.18 from 2:00pm to 7:00pm, and \$0.03 for the reminder) and the user's preferred time range for running the specific

appliances are considered in all cases. Using sliding windows of 3, 5, and 8 hours across the PDFs yields different time preferences termed strict, moderate, and flexible time preferences, as shown in Table 5.4. In addition, the time delay between each phase of appliances in this case study should be no more than 40 minutes.

Table 5.4 User time preference interval used in case study

Activity	3-hour, strict		5-hour, moderate		8-hour, flexible	
	Start	Stop	Start	Stop	Start	Stop
Kitchen cleans up	5:30pm	8:30pm	4:30pm	9:30pm	1:40pm	9:40pm
Laundry	9:00am	12:00pm	8:50am	1:50pm	8:20am	4:20pm

Figure 5.6 depicts the results of hot water usage pattern learning. Figure 5.6(a) presents raw hot water usage data from July 1 to August 15, where the red subsequence corresponds to the 24-hour hot water profile illustrated in Figure 5.6(b). According to the results from the data snippet analysis, the hot water profile in Figure 5.6(b) is the most representative profile of these days—it has the highest similarity with other 24-hour subsequences.

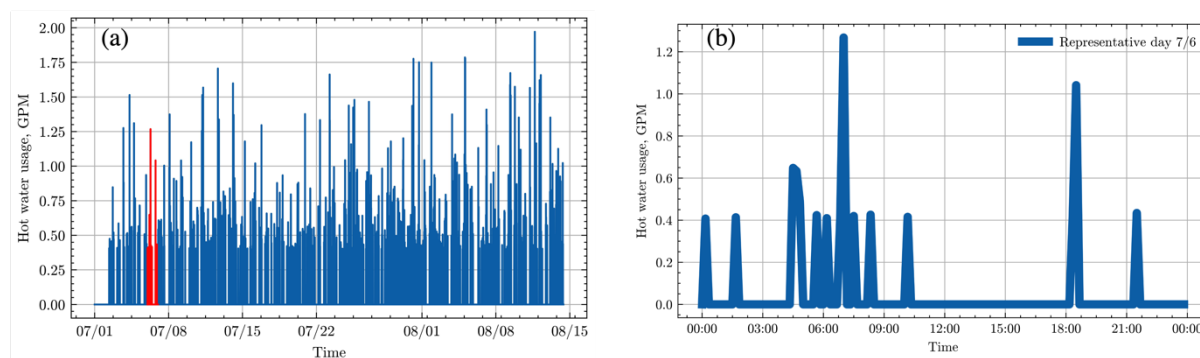


Figure 5.6 Hot water usage profiles (a) hot water usage profile from July to August (b) most representative hot water usage 24-hour profile.

The simulations carried out on a desktop with an AMD 12-core CPU and 32GB of memory with the Python-based optimization package Pyomo [144–146] are demonstrated. The optimal scheduling is solved by Gurobi [147].

5.4 Result and Discussion

The optimization results will be classified into three categories in this section: baseline operation, which is the optimal scheduling without a PV or battery system; Grid-tie system operation, which is the optimal scheduling with a PV system but no battery; and Hybrid system operation, which is the optimal scheduling with both a PV and a battery system. These optimization results, which are reported first in the following section, are generated by solving single-objective optimization problems that aim to minimize costs. Then, it follows with the impact analysis of different factors such as user's time preference. Although the primary focus of this study is to build a scheduling framework for the homeowners, with their decision mainly driven by cost and comfort, cost saving, energy flexibility, and PV self-consumption across three categories will still be reported to provide more insights into different stakeholders. Finally, the multi-objective optimization results will be reported, which seeks to minimize cost while maximizing the time that temperature remains within the user's preferred thermal comfort range, without exceeding the acceptable thermal comfort range.

5.4.1 *Three optimal operations*

The 24-hour optimal scheduling result for the baseline operation, using data from August 7 and assuming a flexible user time preference, is shown in Figure 5.7. The following

simulations in this subsection also apply the same weather condition and user time preferences. The total operating cost for all appliances is \$4.99, which is also the cost of electricity brought from the grid, considering the home does not generate any power itself. On a hot summer day like August 7, the AC (the second subplot on the right-hand side) needs to compensate for additional cooling load due to the hot and humid outdoor conditions. Also, to minimize the operation cost, the AC pre-cools the space to avoid expensive operation during mid-peak and on-peak hours. These two reasons result in frequent AC operation in the morning and before the mid-peak hour. However, the AC cannot completely avoid the on-peak hour, leading to frequent operation while indoor air temperature oscillates around the upper bound of the thermal comfort range. The water heater (the first subplot on the right-hand side) pre-heats the water in the tank and allows the water temperature to go down when the price is high. In baseline operation, non-thermal appliances (subplots on the left-hand side) strictly adhere to the user time preference and avoid the mid-peak and on-peak hour (gray area in the subplots). The last subplot on the left-hand side is the total power buy from the grid (i.e., power consumption summation of all appliances in baseline operation) and the TOU rate.

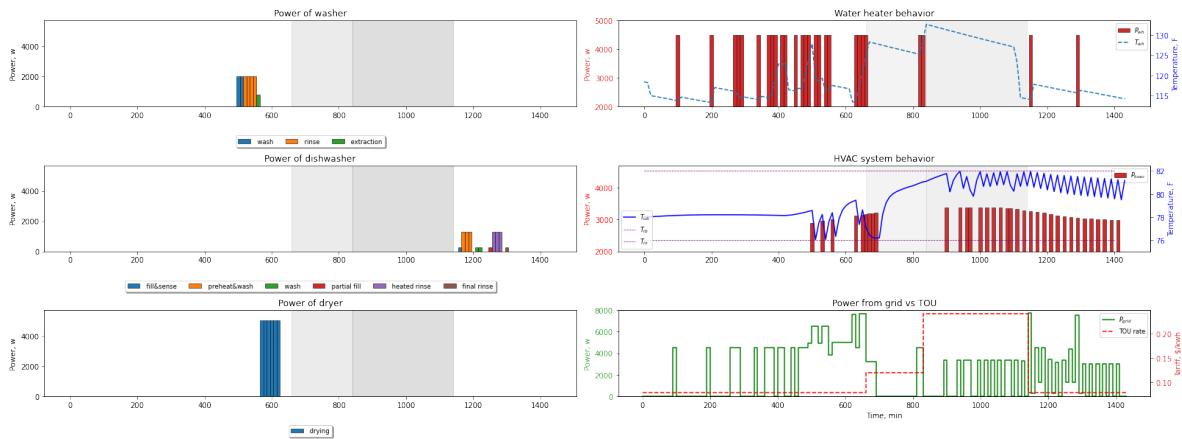


Figure 5.7 Baseline operation of five appliances on August 7.

The five appliances' optimal operation for the Grid-tie system is listed below. With the addition of a PV panel, the total cost is \$2.13, which reduces 57.31% compared with the baseline operation. The water heater and the HVAC system still pre-heat/pre-cool the water and the space air. Notice that in comparison to the baseline operation, there is a noticeable increment in energy consumption of the water heater and the HVAC system during mid-peak hours and on-peak hours. Similarly, the dishwasher is also shifted to the on-peak hour. This outcome is primarily due to that a large part of home load during mid-peak and on-peak is compensated by the PV panel power generation—the “free” power compared with the power from the grid. Thus, appliances aim to utilize PV power generation rather than use the power from the grid, even if during off-peak hours. However, PV power generation has some limitations in terms of timing: its generation might not be large enough in the morning and in the late afternoon, and its generation has overlap with the mid-peak and on-peak hours. When a certain appliance is switched on during mid-peak and on-peak hours, and the PV cannot

completely compensate its load, will lead to a cost increment. Hence, not all the PV power generation, even during mid-peak and on-peak will be allocated to the home load, which leads to a potential PV power waste. More details can be found in Section 5.4.2. The last subplot in Figure 5.8 is the power from the grid and TOU rate. It is obvious that the grid demand during the mid-peak and on-peak is much lower than the baseline cases. In this operation, power from the grid is not zero meaning the PV panel cannot fully compensate the home load.

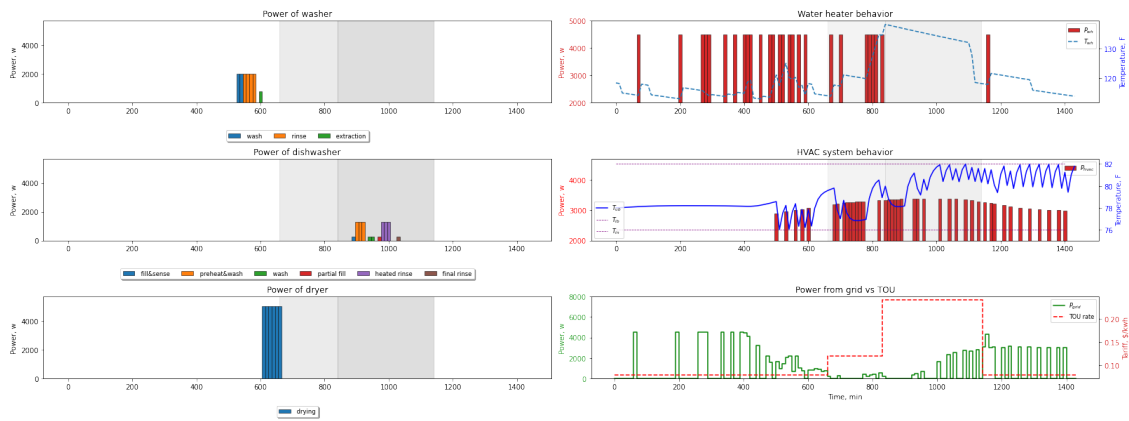


Figure 5.8 Grid-tie system operation of five appliances on August 7.

Figure 5.9 demonstrates the operations of five appliances in a Hybrid system operation when a medium size battery is used. In this case, the 24-hour operation cost is \$1.53 which 69.33% reduction is made compared to the baseline operation. Similar to the Grid-tie system, the water heater and HVAC system operation during mid-peak hours and on-peak hours increases. Furthermore, different from the Grid-tie system, the dryer is shifted to the mid-peak hours rather than operated in off-peak hours. In this case study, dryer consumes around 5000W, which is larger than the PV power generation from 1pm to 1:50pm (the duration of dryer

operation). The rest of the power are meet by the discharging of battery, which means the operation is “free” (i.e., no need to buy electricity from the grid from 1pm to 1:50pm in this operation). Also shown in the last subplot in Figure 5.9, most of load on the mid-peak and on-peak can be compensated by the optimal operation of PV and battery.

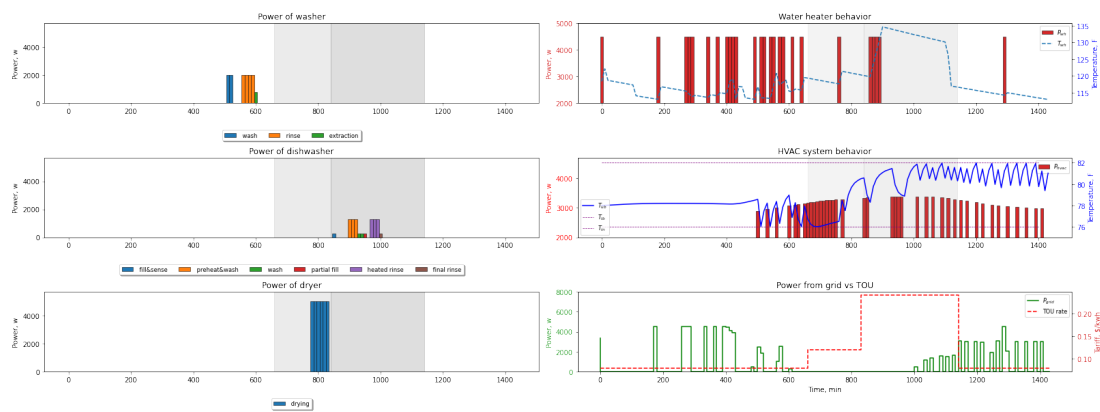


Figure 5.9 Hybrid system operation of five appliances on August 7.

Figure 5.10(a) summarizes the electricity from grid in three different cases. It is obvious that in the Grid-tie system and Hybrid system, the need to buy electricity from grid is reduced significantly. With a battery installed, the load during the mid-peak and on-peak hours reduces further. Figure 5.10(b) indicates battery (dis-)charging states for each time slot. Notices to clearly show the battery (dis-)charging, the charging power is drawn in a green, positive curve, while the discharging power is drawn as a red, negative curve. The blue dashed line is the state of charge of the battery, i.e., the percentage of the capacity still available in the battery. The battery discharges frequently between 12:00pm to 7:00pm (on-peak hour) to compensate for the load, which significantly reduces the cost.

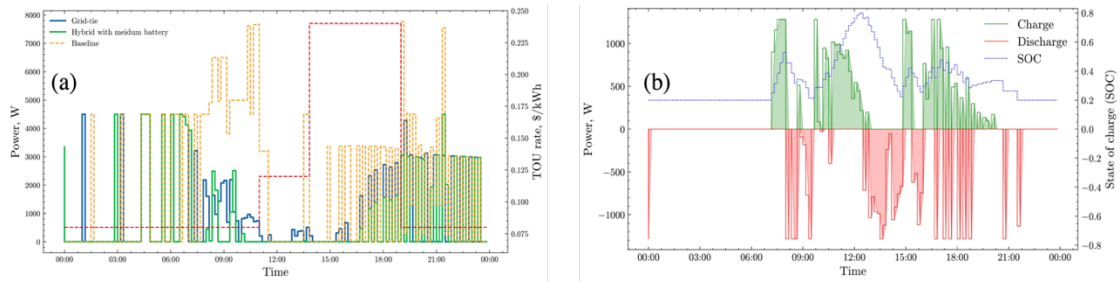


Figure 5.10 Power flow comparison (a) total power buy from grid for three different operations (b) battery (dis-)charging profiles.

5.4.2 Energy flexibility provided by different system operations

Figure 5.11(a)–(d) illustrate power flow for different scenarios on August 7, reflecting optimal scheduling under flexible user time preference. These scenarios include the Grid-tie system (i.e., no battery), and Hybrid system use a small battery, a medium battery, and a large battery, respectively. The larger battery has capacity of 5kWh, and the continuous (dis-)charging limitation is 2700W; the smaller battery refers to a battery that has to 1 kWh capacity, and the continuous (dis-)charging limitation is 900W. The green step plot represents the total power consumption from home appliances (home load), the yellow step plot shows the power contributed by the PV to the home load, and the red step plot is the battery discharge power. The blue step plot, marked with negative values, symbolizes the power that could be generated by PV but was not utilized for either home load or battery charging (i.e., the waste power).

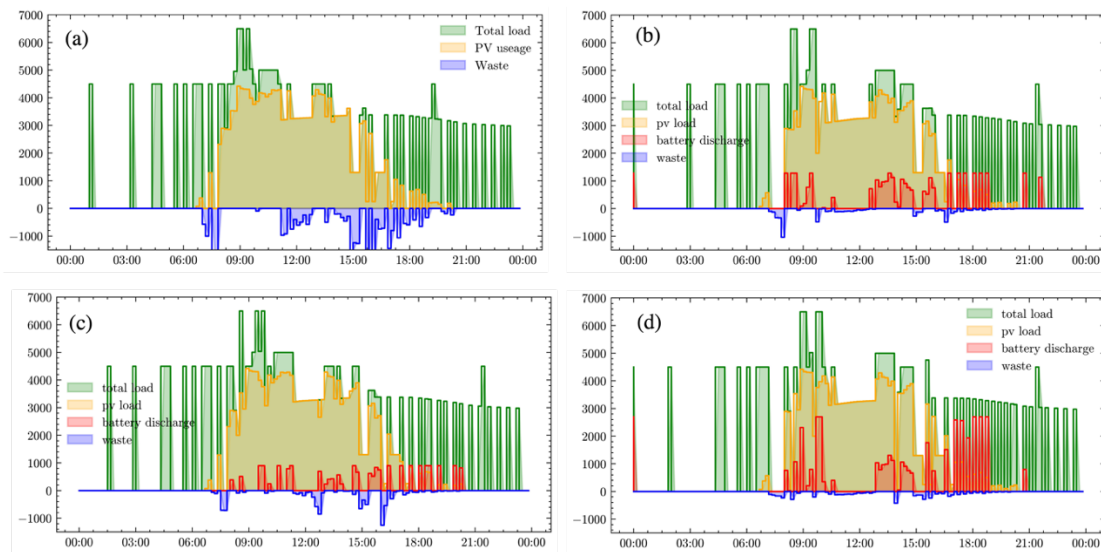


Figure 5.11 Power flow for (a) Grid-tie system (b) Hybrid System with a medium battery (c) Hybrid System with a small battery (d) Hybrid System with a large battery

As shown in Figure 5.11, the Grid-tie system has the largest waste energy, while the Hybrid system with a large battery has almost zero waste energy. Similar to our previous observation in Section 5.4.1, due to the limitation in terms of magnitude and timing (higher PV power generation occurs overlap with mid-peak and on-peak hours) of PV power generation, it is possible that no home load occurred at some time slot while the PV could generate more energy—leading waste of PV power generation. This limitation can be compensated by using a battery as shown in Figure 5.11(b)–(d), which store the amount of energy generated by PV is and not be able to allocate the to home load.

The less waste of PV power generation means the more “free” and clean energy are used. To quantify the benefit of PV power usage and the load-shifting ability in different system operation scenarios, three additional performance indices are used in this study, as shown in

Figure 5.12. For on-site PV panel generation and consumption, two metrics are popular and

usually used together: PV *self-consumption* (SC) quantifies the degree to which a PV panel generation is consumed by the home load, and PV *self-sufficiency* (SS) quantifies the degree that PV panel generation is sufficient to meet the total energy needs of the home [14]. Specifically, the self-consumption (SC) is defined by:

$$SC = \frac{\int_{24hour} P_{pv}}{\int_{24hour} P_{max}} \quad (75)$$

where $\int_{24hour} P_{pv}$ represents the total PV energy consumed by the home load (consumed by home load directly and battery if it exists), and $\int_{24hour} P_{max}$ represents maximum daily PV generation (including waste part). Self-sufficiency (SS) is defined by:

$$SS = \frac{\int_{24hour} P_{pv}}{\int_{24hour} P_{total\ load}} \quad (76)$$

where $\int_{24hour} P_{total\ load}$ represents the total energy needed in the home. Both SC and SS are between 0 to 1, and the low values indicate the mismatch between PV generation and home consumption [148]. As show in Figure 5.12, the Grid-tie system has the lowest PV panel generation and consumption performance, as well as the lowest flexibility factor. Compared with the Grid-tie system, adding a large, a medium, and a small battery will improve the 13.95%, 12.67%, and 10.47% in self-consumption; 17.73%, 14.29%, and 11.88% in self-sufficiency, and 38.41%, 20.91%, and 17.64% in flexibility factor, respectively.

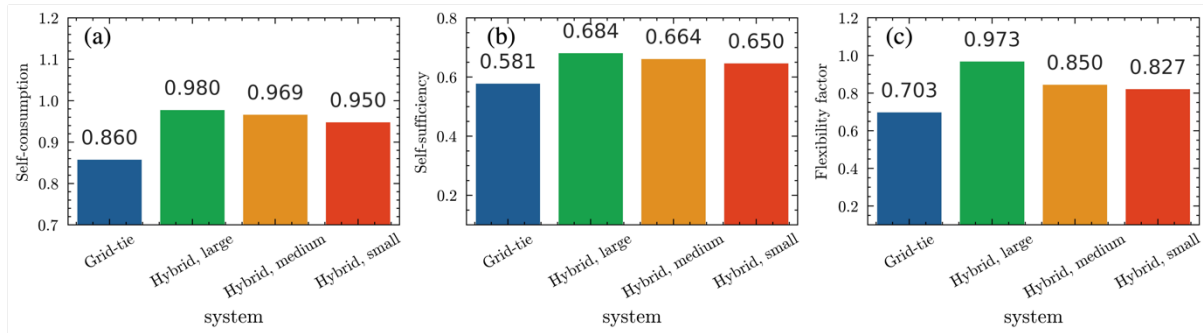


Figure 5.12 Performance indices for different system operations. (a) PV self-consumption (b) PV self-sufficiency (c) Flexibility factor.

5.4.3 Impact of user time preference and weather condition

To further understand the impact of weather conditions, battery size, and user's time preference on total operation cost, 30 scenarios are generated and optimization results in terms of cost are listed in Figure 5.13. The cost list on the top of each scenario is the 24-hour home appliance operation costs given a moderate user time preference. For the weather conditions, August 7, 2022, is chosen, which belongs to a hot summer day cluster mentioned in Chapter 4. Similarly, a day belonging to a mild summer day, September 8 is chosen. The machine learning model is also used to forecast PV power generation using the weather data of September 8. Meanwhile, to maintain consistency, other variables such as the thermal comfort range for AC remained unchanged. The battery specifications are consistent with those presented in Section 5.4.2, and the user time preference interval can be seen in Table 5.4.

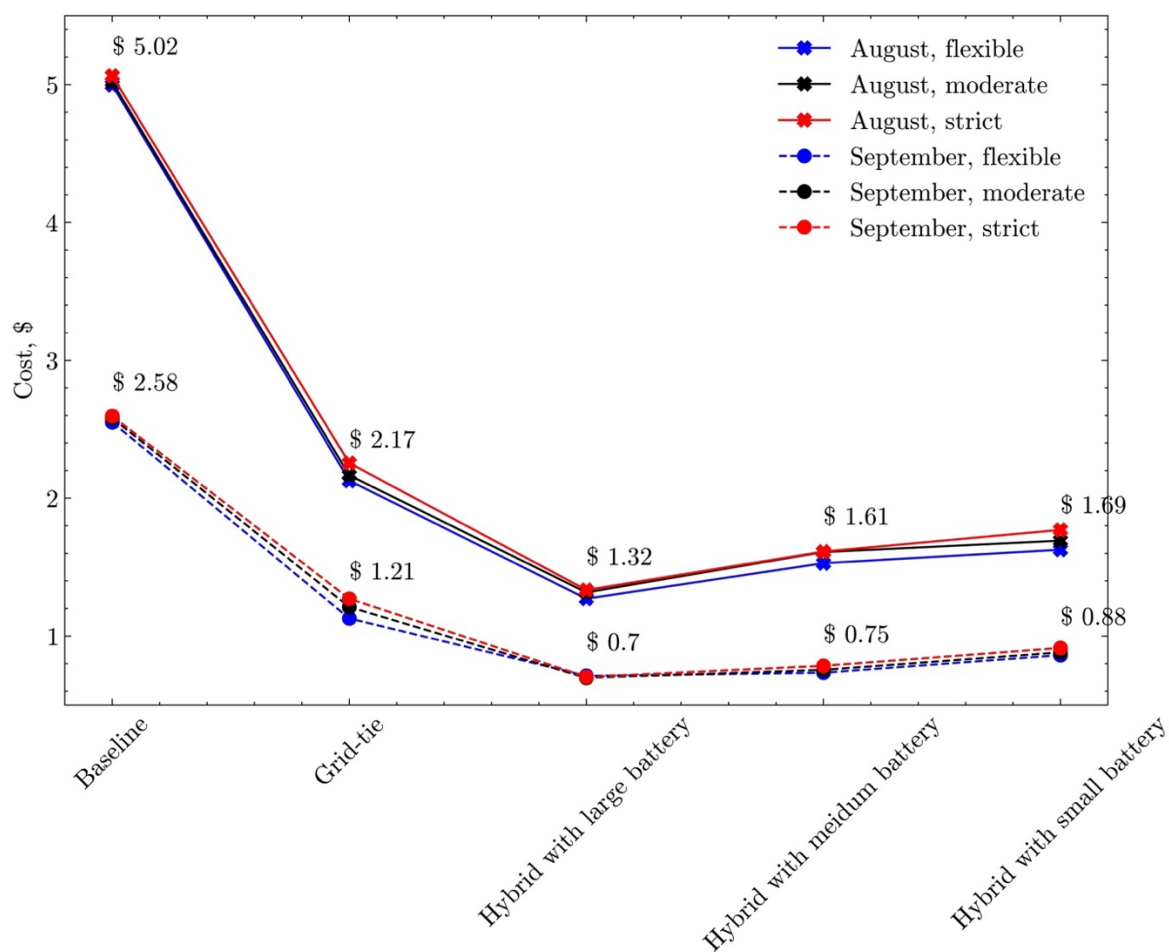


Figure 5.13 Cost-saving variation in terms of different weather conditions, user preference, and operation scenarios.

It can be observed in Figure 5.13 that the weather condition has a dominant impact across different operation scenarios. In baseline cases, a 48.61% cost reduction can be seen if the weather switches from a mild summer day to a hot summer day, and in a Hybrid system with a large battery, a 46.21% cost reduction can be seen. Although the user's time preference has less impact on cost, it can be observed from the figure that the user's time preference impact gets larger when a PV panel is involved but no battery is given (Grid-tie system) or the battery cannot provide enough energy flexibility (Hybrid system with small battery). This result ties to

the inherent limitation of PV power generation (also discussed in previous sections), and the characteristics of non-thermal appliances operation. Non-thermal appliances cannot have a large inactive period like the AC free-floating period, and they must maintain the sequential operation of energy phases. When a strict time limit is given, non-thermal appliances may struggle to take full advantage of PV power generation. As shown in Figure 5.14, with strict user preference, the dishwasher has to run during the late afternoon (the first energy phase starts at 6:40pm and the last energy phase ends at 8:30pm), where the PV power generation is quite limited at that time; with a flexible user time preference, dishwasher is shifted to the on-peak hour to consume more PV power generation. This energy consumption (from non-thermal appliances) and generation (from PV) mismatch can be solved by providing a proper size battery.

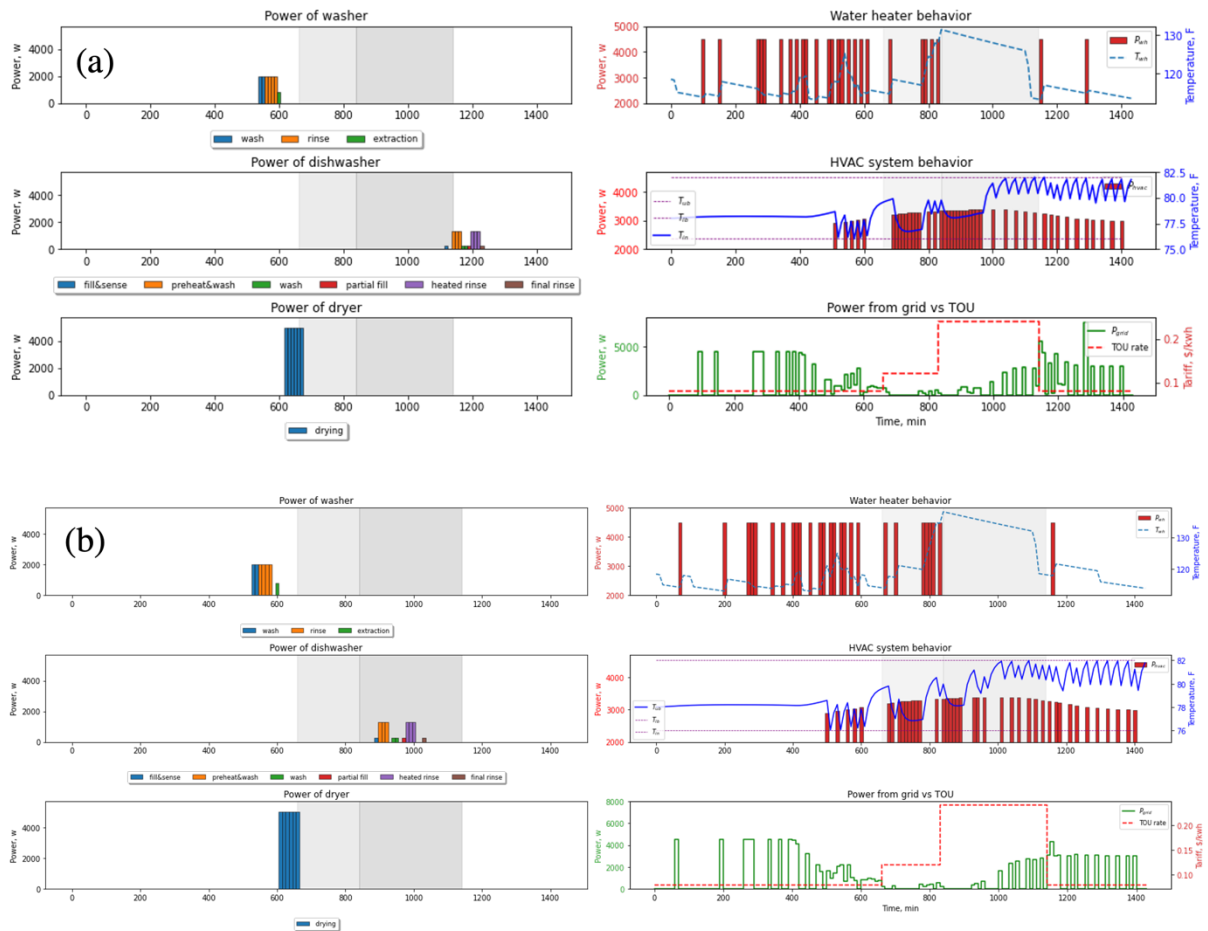


Figure 5.14 User preference impact on appliance operation. (a) Grid-tie system appliance operation given strict user time preference. (b) Grid-tie system appliance operation given flexible user time preference.

5.4.4 Cost and comfort trade-off

In this section, the trade-off between cost and thermal comfort is discussed. The scheduling result of the smart home appliances is conducted by using the Hybrid system operation with a medium size battery on August 7. The user time preference is the flexible scenario, and the other parameters are kept the same as in the previous cases. Figure 5.14(a)–(b) demonstrates 24-hour AC behavior using different weights in the multi-objective optimization framework mentioned in 5.2.5, while the other appliances' operation is not shown.

The light purple bars in all subplots represent the AC power usage, which is a summation of outdoor and indoor unit power usage. The red curve in Figure 5.14 is the indoor air temperature and the green shaded area represents the preferred temperature range. Note that such a preferred temperature range can be given by the user or derived from some comfort standards such as the PMV index. Here we assume that users prefer a uniform temperature range 78°F–80°F and accept a uniform temperature range 76°F–82°F throughout the scheduling period. The three vertical dashed lines divide the scheduling time into four segments, with "shoulder" or the mid-peak hours starting in the morning and the on-peak hour starting in the afternoon.

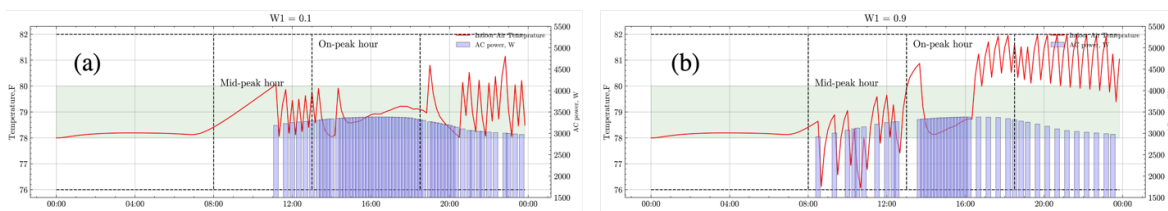


Figure 5.15 Multi-objective optimization results where HVAC system behavior shown only. (a) choose w_1 equals 0.1 (b) choose w_1 equals 0.9.

Two subplots clearly demonstrate how the AC operation change when different weights applied to the optimal scheduling. When the smaller weight, e.g., w_1 equals 0.1, is applied to the scheduling, the system prefers to maximize thermal comfort instead of saving more money. Consequently, the AC system tends to run for extended hours to maintain the temperature between the preferred thermal comfort range. On the other hand, if a larger weight, w_1 equals 0.9, is applied to the scheduling, the system tends to be cost-effective rather than maintain preferred thermal comfort. In that case, the AC precooling effect is obvious. In Figure 16(b),

AC tends to operate to bring the indoor air temperature into a lower level when the price is

relatively low (mid-peak hours), and also operate more when the PV power generation is large between 1:40pm to 4:10pm. Figure 5.16 shows the all appliance's operations when equal weights on cost and comfort are used. In this case, the 24-hour operation cost is \$1.67 while the time that the indoor air temperature maintain in the preferred thermal comfort range (called *comfort*) is 18.17hr. Compared with the scenario that only minimizes the cost, where the operation cost is \$1.52 and the comfort is 12.33hr, the case shown in Figure 5.16, only a 9.15% cost increment extends 47.36% comfort.

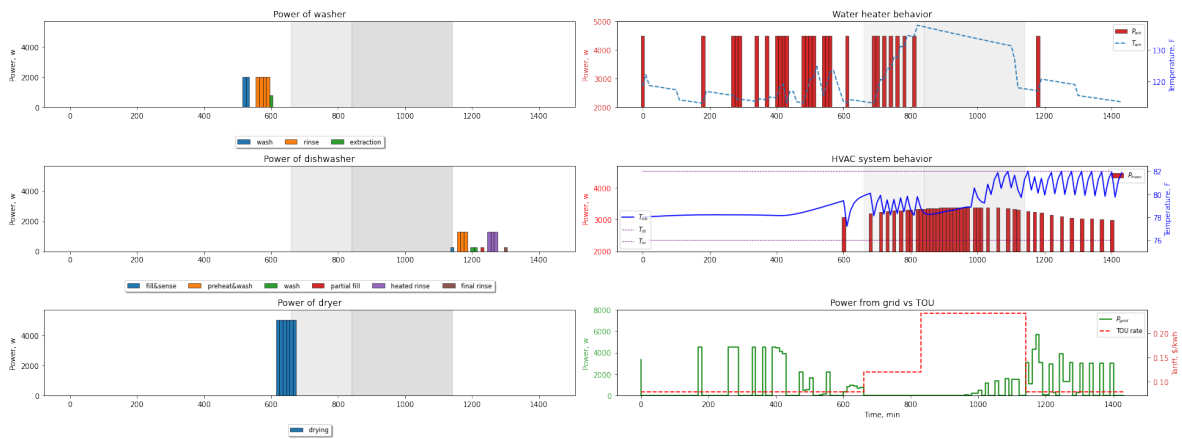


Figure 5.16 Scheduling result with the multi-objective framework with equal weights on cost and comfort (w_1 equals 0.5)

Figure 5.17 shows the cost and comfort trade-off plot–Pareto front which showcases the different combinations of cost and comfort achieved by varying w_1 . It provides insights into the relationship between cost and comfort and enables homeowners to make informed choices based on their preferences for cost or thermal comfort.

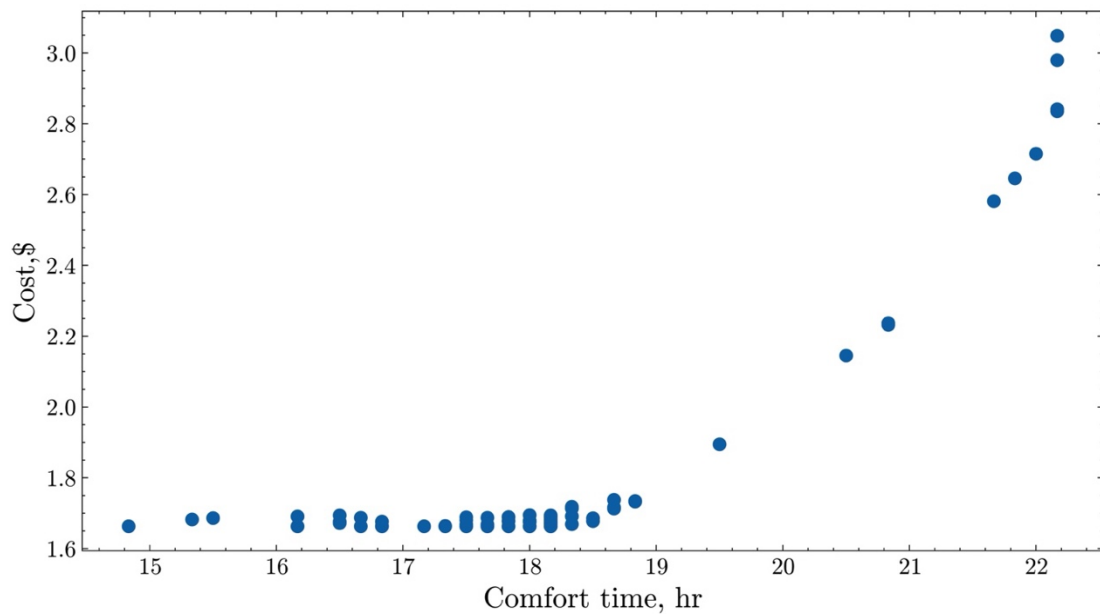


Figure 5.17 Cost and comfort trade-off.

5.5 Summary

In this study, by using mixed-integer linear programming, reduced-order RC-based thermal model, statistics analysis on user's behavior, and multi-objective framework, a household energy management system that integrates thermal, non-thermal appliances, PV panels, and battery systems are developed, simulated, and analyzed. Different from previous studies that hardly considered the user's preference on time of use appliances and water usage, by analyzing the American Time of Use Survey across over three decades and Snippets of the most representative water usage pattern, the proposed framework provides a more realistic simulation scenario. In addition, different from previous studies that rely on either a single temperature range or an ideal temperature setpoint, by providing the homeowner with acceptable and preferred temperature ranges, the framework enables a user-defined trade-off

between electricity cost and thermal comfort. The electricity bill was minimized under the TOU rate by scheduling the operation of each appliance, the energy flow from the solar panel to the appliances, and the (dis-)charge action of the battery.

From the simulation result, compared with the baseline scenarios that no PV and battery installed, the Grid-tie system generated 57.31% of cost savings, while 69.33% can be achieved if a Hybrid system with a medium size battery is included. By providing three more performance indices, the simulation results are evaluated in terms of energy flexibility. It is observed that due to the inherent limitation of the PV system, the Grid-tie system has up to 13.95%, 17.73%, and 38.41% reduction in self-consumption, self-sufficiency, and flexibility factor compared with the Hybrid system with a large battery. Although the Grid-tie system does not include a battery, the scheduling incorporates optimization for the HVAC system and water heater, two other flexible resources in the home. Since these two appliances are the biggest energy “eaters” in homes, they make significant energy flexibility in terms of load shifting. However, the flexibility they provide differs from the rapid flexibility demonstrated by the home battery. The home battery not only encase cost reduction by improving PV self-consumption and self-sufficiency, but it also offers greater flexibility and stabilizes the cost even when users have strict use time preferences (i.e., low energy awareness). Finally, the proposed multi-objective framework results in a 5.84-hour extension in comfort time with only a 9.15% electricity cost increment compared with the minimizing-cost-only strategy.

As with simulation studies, the results reported in this study have limitations. The 2R2C model, which is used to regulate the HVAC system operation, did not consider the internal heat gain. In our case study, because the HVAC system model is built upon the data collected from an unoccupied laboratory house and because the internal heat gain generated by the appliances is relatively small, ignoring the internal heat gain is deemed acceptable. Therefore, the impact of user occupancy, which is associated with internal heat gain, is not discussed in this study. In addition, it is assumed that the user-acceptable and preferred temperature ranges on the HVAC system and user time preference are given. Although the proposed framework can integrate with the given preference easily, deriving user thermal preference and time preference based on historical data is a focus area for the next phase study.

6 Conclusion, Limitations and Future Studies

6.1 Conclusion of this study

This study focused on the development of an energy management framework for coordinating multiple load-flexible resources in residential buildings under demand response programs. The framework embeds three major components. First, it utilizes learning-based gray-box and black-box models for predicting the behavior of the HVAC system, water heaters, and PV power generation. Secondly, it establishes the optimal operation of the HVAC system. Lastly, it focuses on user-centered optimization for the smart Home Energy Management system integrating multiple appliances. By combining these three major components, the proposed framework lays a foundation for a practical solution to transform a residential building into an energy-efficient, cost effective, and grid-interactive connected home. The validity of the framework is substantiated through simulation and field test results, highlighting its feasibility for real-world applications.

Specifically, physical-based gray-box models for the HVAC system and water heater system, as well as a black-box model for the PV power generation are developed in this study. By utilizing these models, the amount of input required from homeowners are minimum. This approach eliminates the need for explicit knowledge of building constructions, water heater thermal properties, or specific PV characteristic, which are typically challenge for homeowner to obtain. Furthermore, the HVAC system model, as a control-oriented model, has been

enriched with adding novel physical-based constraints during the parameter identification procedure. This enhancement significantly increases the model applicability and effectiveness in predictive control. Therefore, the proposed models make the energy management in residential building feasible.

In addition, an MPC-based precooling agent has been developed and implemented in nine test homes using smart thermostats. The successful tests indicate that the proposed MPC agent is a feasible solution for implementing optimal operations in residential building. However, it is also important to acknowledge that the successful control through MPC agent also relies on three different datasets for home thermal model training, real-time control implementation, and post-analysis. These datasets are obtained from the smart thermostat, the node sensor prepared with the thermostat, and various internet-accessible data resources such as weather station web services. While a smart power meter is not mandatory for the MPC agent, it helps build more accurate power consumption predictions for the control purpose. Therefore, to conduct optimal control, both smart thermostats and a robust data management platform to organize several different data sources and enables remote control are necessary.

Through the impact analysis using field test results and an ablation study using co-simulation results, various impact factors including building characteristic and MPC characteristics are discussed. For the building characteristic, factors such as floor size of the home, user's thermal comfort range, and the AC's performance to cool the space prominently affect the optimal precooling results. Regarding MPC characteristic, forecasted weather usage

in our case study plays an important role to impact precooling results.

The home energy management framework proposed in this study, which integrates the thermal appliances, non-thermal appliances, and PV battery system, demonstrates that optimal operation can be applied in a smart home environment to minimize electricity cost and meanwhile incorporate user's preferences. More importantly, through a statistical analysis of the American Time Use Survey and data mining of the water usage profile, user appliance user pattern and water usage behavior can be integrated into the energy management framework. Through the multi-objective optimization, the homeowner has the flexibility to not only choose their acceptable and preferred indoor air temperature range, but also determine their trade-off between cost and comfort by selecting the weights on each goal, thus enabling to achieve both benefits based on their intension.

Finally, four performance indices—cost saving, flexibility factor, self-consumption, and self-sufficiency—are utilized to quantify the optimization results in residential buildings. As mentioned in the introduction, the financial factor remains a major motivations for homeowners participating in demand response programs. Hence, the cost-saving generally present most valuable insight to the homeowners. The flexibility factor signifies the load-shifting ability of a residential building and provides valuable insight to stakeholders such as grid operators. On the other hand, self-consumption and self-sufficiency are performance indicators typically used to quantify the performance of distributed generation resources such as PV systems. A higher value of these two performance indices denotes an increased use of cleaner and “free” energy

during operation. This aspect holds interest both homeowners and grid operators.

6.2 Limitations

6.2.1 *Internal heat gain and occupancy*

Internal gain and occupancy schedules from individuals are not considered in our optimization framework. This is not only because the heat generated by normal human activity generally has a relatively smaller impact compared to influences from outdoor air, solar irradiation, AC cooling, and infiltration in a residential building. Additionally, the current occupancy sensor technology in residential building is relatively simple. For example, the occupancy detection embedded in the thermostat and node sensor used in this study relied on infrared technology, which means it fails to detect an occupied signal unless the homeowner physically passes by the sensor. Consequently, if the thermostat or sensor is located in a less traffic area in a home, such measurement is likely to be inaccurate. Inaccurate occupancy measurement may lead to less optimal operation, and it might even bring discomfort to homeowners. In addition, different from the carbon dioxide (CO₂) sensors, which are frequently used in commercial buildings, the occupancy sensor currently used in residential buildings lacks a headcount feature. The heat gain from electrical equipment is also ignored in this study. The electrical equipment included in this study is a clothes washer, dryer, and dishwasher, whose impact on the thermal dynamics remains relatively small.

6.2.2 *Growth of the problem size*

The *curse of dimensionality* is a common bottleneck of the optimization scheduling algorithm. In [149], the author explained how the problem size can be growing super linearly with an increased number of scheduling time steps and the number of devices using diagrams. The authors also reviewed recent publications in terms of techniques used to reduce the computational burden in home energy management system. It can be observed that the current solution primarily focuses on reducing the scheduling resolution or horizon, decreasing the number of devices, or combining both approaches. Similar to [92], this study delivers an upper time limit of 500 seconds for solving the optimal solution to the solver. This represents the trade-off between achieving optimal scheduling results and managing the computational burden. With the rapid development in the field of operational research and the emergence of new algorithms, there may be more potential solutions to this issue.

6.2.3 Thermal comfort

In the impact study, it is assumed that user's thermal comfort range is known before the MPC agent begins precooling or the home energy management framework starts to control the appliances. While several thermal comfort models such as PMV and PDD [150,151] can be embedded in the energy management framework, two limitations needed to be taken into account. First, the PMV and PDD indices mainly focus on a group of people's thermal comfort rather than on individual thermal comfort. As shown in the field test study, homeowners lived in houses with similar size, locations, and construction may have different thermal comfort ranges. This suggests that a PMV or PDD index might not be appropriate for a smart home

environment, where usually only two or three individuals determines the HVAC system setpoint. The second limitation is that thermal comfort should be considered as a transient model, as discussed in various publications such as [152–154]. In fact, we have observed that some homeowners may return home and set the AC to a lower setpoint than usual. One significant explanation for this behavior is that the change in thermal comfort is a transient process, not only determined by current environment (current room air temperature if homeowner is at home), but also influenced by the previous environment (previous hot outdoor environment if homeowner come back from outside). Homeowners returning from a very hot outdoor environment might require a lower temperature to remove the stored heat—this could affect the operation of the HVAC system as well. Therefore, there is a need to integrate a personal thermal comfort model, which considered transient, non-uniform effects in our proposed energy management framework in the future.

6.3 Future Studies

6.3.1 Meta-model to quickly evaluate energy flexibility potential of a home

Despite the impact analysis has concluded dominate impact factors that affect the optimal operation results of the HEMS, there is still a lack of a meta-model that can quickly evaluate load shifting potential of a home. Once optimizations have been conducted in large and diverse enough residential buildings, it would be possible to establish a meta-model that connects building and homeowner characteristics to cost saving and energy flexibility under a specific demand response program. Such a meta-model would allow the homeowner to estimate their

average precooling potential prior to implementing the energy management framework, and it would also facilitate policymakers or grid operators to quantify the energy flexibility of a group of buildings with ease. Although, there are already some researchers discussed about such a meta-model, developing the meta-model requires a large and diverse set of residential buildings. In this regard, some representative building stock, such as ResStock [155], can serve as important tools.

6.3.2 *Data-driven occupancy behavior model*

As mentioned in Section 6.2.3, a personal thermal comfort model, which considered a transient change in human skin and core temperature, is needed. Although individual differences have been considered in various physical-based transient thermal comfort modeling, most of these differences are primarily represented by variables such as human body mass or blood flow rate. This does not sufficiently encompass differences in personal preferences. In this regard, a hybrid model to predict dynamic, individual thermal comfort might also be necessary for future work.

Apart from the thermal comfort, a data-driven model to predict the times when occupants arrive and leave home could also significantly enhance the widespread of energy management framework. An important application of this would be integrating the Electrical Vehicle (EV) charging schedules into the energy management framework, which largely depend on the homeowner's arrival and departure home times.

6.3.3 *Stochastic optimizations consider uncertainty*

In the energy management framework, several resources involve uncertainties, including the prediction such as the PV-power generation, forecasted weather conditions, and the user's appliance usage time. As these uncertainties increase, it becomes worthwhile to conduct robust optimization techniques, such that stochastic optimization, to coordinate multiple load-flexible resources. However, stochastic optimization often depends on maximizing the expectation of the objective across different scenarios, which can require substantial computational resources due to multiple uncertainties resources exists. Therefore, it is important to utilize or develop an algorithm capable reducing the number of potential scenarios derived from these multiple sources of uncertainty. Ideally, this algorithm should identify the most representative scenario, the worse scenario, and the best scenario, thereby balancing computation efficiency and with the quality of optimization.

This material is based upon work supported by the U.S. Department of Energy’s Office of Energy Efficiency and Renewable Energy (EERE) under the Building Technologies Office Award Number DE-EE0008697.

Appendix

Appendix A: Matrix Profile-based Weather Clustering Algorithm

In this paper, we combine the matrix profile-based data snippets from [130,156] (for finding the most representative subsequence) with the MASS algorithm from [156,157] (for a fast pattern discovery search) and introduce our own modification to group similar weather patterns into clusters. Figure A. 1 displays a flowchart that describes the modified algorithm, which contains three stages. The first stage is the data processing, whereby we first query historical weather datasets from a local weather station. We then resample the data at 5-minute intervals, remove duplicate and missing data, and run different data imputations (linear or polynomial interpolation) to improve data quality. The second stage is the time series snippets stage, whereby we use the matrix profile-based data snippets in [130] to identify three representative days—hot, mild, and cold days—from July to October. The third stage is the clustering stage, whereby we first calculate the distance profiles of outdoor air temperature, solar irradiation, and relative humidity between each representative day and each 24-hour subsequence. To speed up calculation, the MASS algorithm from [157] is utilized. We then calculate the normalized weighted average of the aforementioned distance profiles and find the

top K days that are most similar to the representative days. The latter is accomplished by selecting the top K smallest points of the new distance profiles. In this paper, K is selected to be 30. Since outdoor air dictates the heat transfer between indoor indoor and outdoor enviornment, the weights assigned to the outdoor air temerperature, solar irradiation, and relative humidity are 0.6, 0.3, and 0.1. If a data point resides in multiple clusters, it is moved to the cluster from which it has the smallest new distance profile.

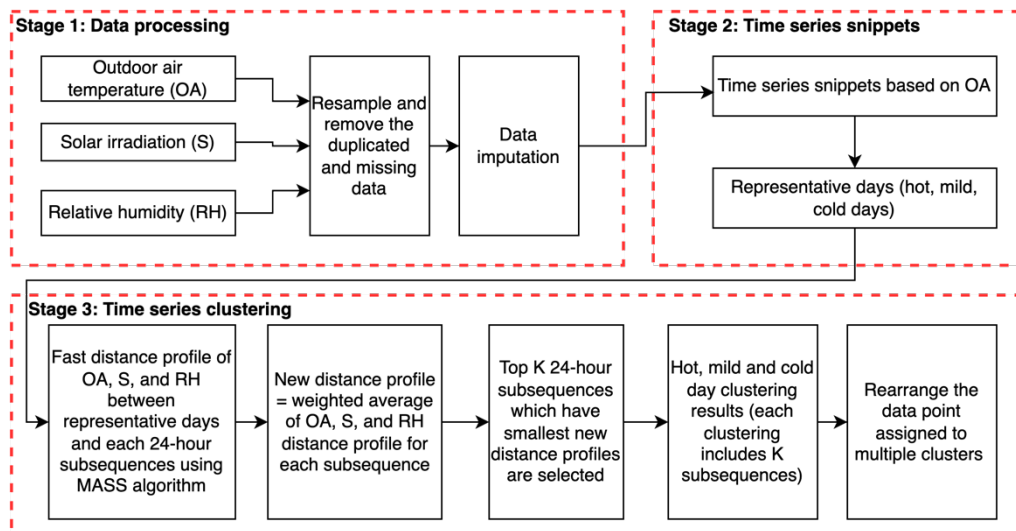


Figure A. 1 Three stages of the modified matrix profile-based weather clustering algorithm

References

- [1] U.S. Energy Information Administration (EIA), “Frequently Asked Questions (FAQs) What Is U.S. Electricity Generation by Energy Source?” [Online]. Available: <https://www.eia.gov/tools/faqs/faq.php>. [Accessed: 07-Mar-2023].
- [2] “Ten Years of Analyzing the Duck Chart: How an NREL Discovery in 2008 Is Helping Enable More Solar on the Grid Today” [Online]. Available: <https://www.nrel.gov/news/program/2018/10-years-duck-curve.html>. [Accessed: 08-Feb-2023].
- [3] “California ISO - Renewables and Emissions Reports” [Online]. Available: <https://www.caiso.com/market/Pages/ReportsBulletins/RenewablesReporting.aspx>. [Accessed: 08-Feb-2023].
- [4] Sinsel, S. R., Riemke, R. L., and Hoffmann, V. H., 2020, “Challenges and Solution Technologies for the Integration of Variable Renewable Energy Sources—a Review,” *renewable energy*, **145**, pp. 2271–2285.
- [5] Meyabadi, A. F., and Deihimi, M. H., 2017, “A Review of Demand-Side Management: Reconsidering Theoretical Framework,” *Renewable and Sustainable Energy Reviews*, **80**, pp. 367–379.
- [6] U.S. Energy Information Administration (EIA), 2022, *December 2022 Monthly Energy Review*, OE/EIA-0035.
- [7] U.S. Energy Information Administration (EIA), 2022, *Annual Energy Outlook 2022 with Projections to 2050*, AEO2022 Narrative, Washington, D.C., USA.
- [8] “Residential Energy Consumption Survey (RECS) - Data - U.S. Energy Information Administration (EIA)” [Online]. Available: <https://www.eia.gov/consumption/residential/data/2015/>. [Accessed: 18-Feb-2021].
- [9] “Reports on Demand Response and Advanced Metering | Federal Energy Regulatory Commission” [Online]. Available: <https://www.ferc.gov/power-sales-and-markets/demand-response/reports-demand-response-and-advanced-metering>. [Accessed: 23-Mar-2023].
- [10] “Demand Response and Time-Variable Pricing Programs,” Energy.gov [Online]. Available: <https://www.energy.gov/eere/femp/demand-response-and-time-variable-pricing-programs>. [Accessed: 31-Dec-2022].
- [11] Neukomm, M., Nubbe, V., and Fares, R., 2019, “Grid-Interactive Efficient Buildings Technical Report Series: Overview of Research Challenges and Gaps.”
- [12] Satchwell, A., Piette, M. A., Khandekar, A., Granderson, J., Frick, N. M., Hledik, R., Faruqui, A., Lam, L., Ross, S., Cohen, J., Wang, K., Urigwe, D., Delurey, D., Neukomm, M., and Nemtsov, D., 2021, *A National Roadmap for Grid-Interactive Efficient Buildings*, Lawrence Berkeley National Lab. (LBNL), Berkeley, CA (United States).
- [13] Chen, Y., Xu, P., Gu, J., Schmidt, F., and Li, W., 2018, “Measures to Improve Energy Demand Flexibility in Buildings for Demand Response (DR): A Review,” *Energy and* 159

- Buildings, **177**, pp. 125–139.
- [14] Li, H., Wang, Z., Hong, T., and Piette, M. A., 2021, “Energy Flexibility of Residential Buildings: A Systematic Review of Characterization and Quantification Methods and Applications,” *Advances in Applied Energy*, **3**, p. 100054.
- [15] “Use of Electricity - U.S. Energy Information Administration (EIA)” [Online]. Available: <https://www.eia.gov/energyexplained/electricity/use-of-electricity.php>. [Accessed: 04-Apr-2023].
- [16] Turner, W. J., Roux, J., and Walker, I. S., 2014, “Reducing Residential Peak Electricity Demand with Mechanical Pre-Cooling of Building Thermal Mass.”
- [17] McQuiston, F. C., Parker, J. D., and Spitler, J. D., 2005, *Heating, Ventilating, and Air Conditioning: Analysis and Design*, John Wiley & Sons, Hoboken, N.J.
- [18] Foucquier, A., Robert, S., Suard, F., Stéphan, L., and Jay, A., 2013, “State of the Art in Building Modelling and Energy Performances Prediction: A Review,” *Renewable and Sustainable Energy Reviews*, **23**, pp. 272–288.
- [19] Yao, Y., and Shekhar, D. K., 2021, “State of the Art Review on Model Predictive Control (MPC) in Heating Ventilation and Air-Conditioning (HVAC) Field,” *Building and Environment*, **200**, p. 107952.
- [20] Booyesen, M. J., Engelbrecht, J. A. A., Ritchie, M., Apperley, M., and Cloete, A. H., “How Much Energy Can Optimal Control of Domestic Water Heating Save?,” p. 22.
- [21] Naderi, S., Pignatta, G., Heslop, S., MacGill, I., and Chen, D., 2022, “Demand Response via Pre-Cooling and Solar Pre-Cooling: A Review,” *Energy and Buildings*, **272**, p. 112340.
- [22] Khezri, R., Mahmoudi, A., and Aki, H., 2022, “Optimal Planning of Solar Photovoltaic and Battery Storage Systems for Grid-Connected Residential Sector: Review, Challenges and New Perspectives,” *Renewable and Sustainable Energy Reviews*, **153**, p. 111763.
- [23] Severson, K. A., Attia, P. M., Jin, N., Perkins, N., Jiang, B., Yang, Z., Chen, M. H., Aykol, M., Herring, P. K., Fraggedakis, D., Bazant, M. Z., Harris, S. J., Chueh, W. C., and Braatz, R. D., 2019, “Data-Driven Prediction of Battery Cycle Life before Capacity Degradation,” *Nat Energy*, **4**(5), pp. 383–391.
- [24] Cai, J., Zhang, H., and Jin, X., 2019, “Aging-Aware Predictive Control of PV-Battery Assets in Buildings,” *Applied Energy*, **236**, pp. 478–488.
- [25] “Annual Electric Power Industry Report, Form EIA-861 Detailed Data Files” [Online]. Available: <https://www.eia.gov/electricity/data/eia861/>. [Accessed: 22-Mar-2023].
- [26] Reiss, P. C., and White, M. W., 2005, “Household Electricity Demand, Revisited,” *The Review of Economic Studies*, **72**(3), pp. 853–883.
- [27] Pritoni, M., Meier, A. K., Aragon, C., Perry, D., and Peffer, T., 2015, “Energy Efficiency and the Misuse of Programmable Thermostats: The Effectiveness of Crowdsourcing for Understanding Household Behavior,” *Energy Research & Social Science*, **8**, pp. 190–197.
- [28] Gyamfi, S., Krumdieck, S., and Urmee, T., 2013, “Residential Peak Electricity Demand Response—Highlights of Some Behavioural Issues,” *Renewable and Sustainable Energy Reviews*, **25**, pp. 71–77.

- [29] Nolan, S., and O'Malley, M., 2015, "Challenges and Barriers to Demand Response Deployment and Evaluation," *Applied Energy*, **152**, pp. 1–10.
- [30] Good, N., Ellis, K. A., and Mancarella, P., 2017, "Review and Classification of Barriers and Enablers of Demand Response in the Smart Grid," *Renewable and Sustainable Energy Reviews*, **72**, pp. 57–72.
- [31] Safdarian, A., Fotuhi-Firuzabad, M., and Lehtonen, M., 2019, "Demand Response from Residential Consumers: Potentials, Barriers, and Solutions," *Smart Grids and Their Communication Systems*, pp. 255–279.
- [32] Parrish, B., Heptonstall, P., Gross, R., and Sovacool, B. K., 2020, "A Systematic Review of Motivations, Enablers and Barriers for Consumer Engagement with Residential Demand Response," *Energy Policy*, **138**, p. 111221.
- [33] Walker, I. S., and Meier, A. K., 2008, "Residential Thermostats: Comfort Controls in California Homes," *Lawrence Berkeley National Laboratory*, **7**.
- [34] Rotondo, J., Johnson, R., Gonzales, N., Waranowski, A., Badger, C., Lange, N., Goldman, E., and Foster, R., 2016, *Overview of Existing and Future Residential Use Cases for Connected Thermostats*, Energetics Inc., Washington, DC (United States); Vermont Energy Investment
- [35] "Smart Thermostats & Smart Home Devices | Ecobee" [Online]. Available: <https://www.ecobee.com/en-us/>. [Accessed: 30-Mar-2023].
- [36] "Nest Temperature Sensor - Google Store" [Online]. Available: https://store.google.com/product/nest_temperature_sensor?hl=en-US&pli=1. [Accessed: 07-Apr-2023].
- [37] Sparr, B., and Maguire, J., 2021, "Increasing the Number of Installed Residential Heat Pump Water Heaters in the USA through Improved Technology and Utility Programs," *Current Sustainable/Renewable Energy Reports*, **8**, pp. 107–113.
- [38] "Monitor Your Water Use | Stop Leaks Fast | Flume 2 Smart Water Monitor," Flume [Online]. Available: <https://flumewater.com/product/>. [Accessed: 07-Apr-2023].
- [39] Doris, E., and Krasko, V., 2012, *Strategic Sequencing for State Distributed PV Policies: A Quantitative Analysis of Policy Impacts and Interactions*, National Renewable Energy Lab.(NREL), Golden, CO (United States).
- [40] Feldman, D., Ramasamy, V., Fu, R., Ramdas, A., Desai, J., and Margolis, R., 2021, *US Solar Photovoltaic System and Energy Storage Cost Benchmark (Q1 2020)*, National Renewable Energy Lab.(NREL), Golden, CO (United States).
- [41] "Powerwall," Tesla [Online]. Available: <https://www.tesla.com/powerwall>. [Accessed: 07-Apr-2023].
- [42] Inc, G. P. S., "Powering Your Home, Your Business, Your World." [Online]. Available: <https://www.generac.com>. [Accessed: 07-Apr-2023].
- [43] "Encharge 10, All-in-One Solar Battery for Your Solar Energy Storage Needs | Enphase Storage" [Online]. Available: <https://enphase.com/store/storage/iq-battery-10t>. [Accessed: 07-Apr-2023].

- [44] Kutrašnik, T., Mele, I., and Zelič, K., 2021, “Multi-Scale Modelling of Lithium-Ion Batteries: From Transport Phenomena to the Outbreak of Thermal Runaway,” *Energy Conversion and Management*, **236**, p. 114036.
- [45] Shabani, M., Wallin, F., Dahlquist, E., and Yan, J., 2022, “Techno-Economic Assessment of Battery Storage Integrated into a Grid-Connected and Solar-Powered Residential Building under Different Battery Ageing Models,” *Applied Energy*, **318**, p. 119166.
- [46] Minoli, D., Sohraby, K., and Occhiogrosso, B., 2017, “IoT Considerations, Requirements, and Architectures for Smart Buildings—Energy Optimization and next-Generation Building Management Systems,” *IEEE Internet of Things Journal*, **4**(1), pp. 269–283.
- [47] Bird, J., 2015, “Developing the Smarter Grid: The Role of Domestic and Small and Medium Enterprise Customers,” *Customer-Led Network Revolution*.
- [48] Nilsson, A., Bergstad, C. J., Thuvander, L., Andersson, D., Andersson, K., and Meiling, P., 2014, “Effects of Continuous Feedback on Households’ Electricity Consumption: Potentials and Barriers,” *Applied Energy*, **122**, pp. 17–23.
- [49] Simon, H. A., 1986, “Rationality in Psychology and Economics,” *Journal of Business*, pp. S209–S224.
- [50] Friis, F., and Christensen, T. H., 2016, “The Challenge of Time Shifting Energy Demand Practices: Insights from Denmark,” *Energy Research & Social Science*, **19**, pp. 124–133.
- [51] Keeney, K. R., and Braun, J. E., 1997, “Application of Building Precooling to Reduce Peak Cooling Requirements,” *ASHRAE transactions*, **103.1**, pp. 463–469.
- [52] Alimohammadisagvand, B., Jokisalo, J., and Sirén, K., 2018, “Comparison of Four Rule-Based Demand Response Control Algorithms in an Electrically and Heat Pump-Heated Residential Building,” *Applied Energy*, **209**, pp. 167–179.
- [53] Shi, H., Liu, J., and Chen, Q., 2019, “An RC-Network Approach for HVAC Precooling Optimization in Buildings,” *IEEE Trans. Sustain. Comput.*, pp. 1–1.
- [54] Kishore, R. A., Bianchi, M. V. A., Booten, C., Vidal, J., and Jackson, R., 2020, “Modulating Thermal Load through Lightweight Residential Building Walls Using Thermal Energy Storage and Controlled Precooling Strategy,” *Applied Thermal Engineering*, **180**, p. 115870.
- [55] Wang, J., Tang, C. Y., and Song, L., 2022, “Analysis of Precooling Optimization for Residential Buildings,” *Applied Energy*, **323**, p. 119574.
- [56] Vedullapalli, D. T., Hadidi, R., and Schroeder, B., 2019, “Combined HVAC and Battery Scheduling for Demand Response in a Building,” *IEEE Transactions on Industry Applications*, **55**(6), pp. 7008–7014.
- [57] Jiang, Z., Risbeck, M. J., Ramamurti, V., Murugesan, S., Amores, J., Zhang, C., Lee, Y. M., and Drees, K. H., 2021, “Building HVAC Control with Reinforcement Learning for Reduction of Energy Cost and Demand Charge,” *Energy and Buildings*, **239**, p. 110833.
- [58] Wang, Z., Chen, B., Li, H., and Hong, T., 2021, “AlphaBuilding ResCommunity: A Multi-Agent Virtual Testbed for Community-Level Load Coordination,” *Advances in Applied Energy*, **4**, p. 100061.

- [59] Tabares-Velasco, P. C., Speake, A., Harris, M., Newman, A., Vincent, T., and Lanahan, M., 2019, “A Modeling Framework for Optimization-Based Control of a Residential Building Thermostat for Time-of-Use Pricing,” *Applied Energy*, **242**, pp. 1346–1357.
- [60] Wang, J., Jiang, Y., Tang, C. Y., and Song, L., 2023, “Analysis of Predicted Mean Vote-Based Model Predictive Control for Residential HVAC Systems,” *Building and Environment*, **229**, p. 109952.
- [61] Afram, A., and Janabi-Sharifi, F., 2017, “Supervisory Model Predictive Controller (MPC) for Residential HVAC Systems: Implementation and Experimentation on Archetype Sustainable House in Toronto,” *Energy and Buildings*, **154**, pp. 268–282.
- [62] Baniasadi, A., Habibi, D., Bass, O., and Masoum, M. A. S., 2019, “Optimal Real-Time Residential Thermal Energy Management for Peak-Load Shifting With Experimental Verification,” *IEEE Transactions on Smart Grid*, **10**(5), pp. 5587–5599.
- [63] Finck, C., Li, R., and Zeiler, W., 2019, “Economic Model Predictive Control for Demand Flexibility of a Residential Building,” *Energy*, **176**, pp. 365–379.
- [64] Merema, B., Saelens, D., and Breesch, H., 2022, “Demonstration of an MPC Framework for All-Air Systems in Non-Residential Buildings,” *Building and Environment*, **217**, p. 109053.
- [65] Blum, D., Wang, Z., Weyandt, C., Kim, D., Wetter, M., Hong, T., and Piette, M. A., 2022, “Field Demonstration and Implementation Analysis of Model Predictive Control in an Office HVAC System,” *Applied Energy*, **318**, p. 119104.
- [66] Li, B., Wu, B., Peng, Y., and Cai, W., 2022, “Tube-Based Robust Model Predictive Control of Multi-Zone Demand-Controlled Ventilation Systems for Energy Saving and Indoor Air Quality,” *Applied Energy*, **307**, p. 118297.
- [67] Sutton, R. S., and Barto, A. G., 2018, *Reinforcement Learning: An Introduction*, MIT press.
- [68] Borrelli, F., Bemporad, A., and Morari, M., 2017, *Predictive Control for Linear and Hybrid Systems*, Cambridge University Press.
- [69] Camacho, E. F., 2007, *Model Predictive Control by Eduardo F. Camacho, Carlos Bordons Alba*.
- [70] Huchuk, B., Sanner, S., and O’Brien, W., 2021, “Development and Evaluation of Data-Driven Controls for Residential Smart Thermostats,” *Energy and Buildings*, **249**, p. 111201.
- [71] Kim, D., Wang, Z., Brugger, J., Blum, D., Wetter, M., Hong, T., and Piette, M. A., 2022, “Site Demonstration and Performance Evaluation of MPC for a Large Chiller Plant with TES for Renewable Energy Integration and Grid Decarbonization,” *Applied Energy*, **321**, p. 119343.
- [72] Lee, H., and Heo, Y., 2022, “Simplified Data-Driven Models for Model Predictive Control of Residential Buildings,” *Energy and Buildings*, **265**, p. 112067.
- [73] Xu, X., Wang, S., and Huang, G., 2010, “Robust MPC for Temperature Control of Air-Conditioning Systems Concerning on Constraints and Multitype Uncertainties,” *Building*

- Services Engineering Research and Technology, **31**(1), pp. 39–55.
- [74] Maasoumy, M., Razmara, M., Shahbakhti, M., and Vincentelli, A. S., 2014, “Handling Model Uncertainty in Model Predictive Control for Energy Efficient Buildings,” *Energy and Buildings*, **77**, pp. 377–392.
- [75] Afram, A., and Janabi-Sharifi, F., 2014, “Theory and Applications of HVAC Control Systems – A Review of Model Predictive Control (MPC),” *Building and Environment*, **72**, pp. 343–355.
- [76] Mayne, D. Q., Rawlings, J. B., Rao, C. V., and Sokaert, P. O. M., 2000, “Constrained Model Predictive Control: Stability and Optimality,” *Automatica*, **36**, pp. 789–814.
- [77] ENERNOC, 2009, *The Demand Response Baseline*, ENERNOC.
- [78] Zhang, L., Good, N., and Mancarella, P., 2019, “Building-to-Grid Flexibility: Modelling and Assessment Metrics for Residential Demand Response from Heat Pump Aggregations,” *Applied Energy*, **233–234**, pp. 709–723.
- [79] Wang, J., Tang, C. Y., and Song, L., 2020, “Design and Analysis of Optimal Pre-Cooling in Residential Buildings,” *Energy and Buildings*, **216**, p. 109951.
- [80] Joe, J., 2022, “Investigation on Pre-Cooling Potential of UFAD via Model-Based Predictive Control,” *Energy and Buildings*, **259**, p. 111898.
- [81] Lindelöf, D., Afshari, H., Alisafae, M., Biswas, J., Caban, M., Mocellin, X., and Viaene, J., 2015, “Field Tests of an Adaptive, Model-Predictive Heating Controller for Residential Buildings,” *Energy and Buildings*, **99**, pp. 292–302.
- [82] Risteska Stojkoska, B. L., and Trivodaliev, K. V., 2017, “A Review of Internet of Things for Smart Home: Challenges and Solutions,” *Journal of Cleaner Production*, **140**, pp. 1454–1464.
- [83] Vielma, J. P., 2015, “Mixed Integer Linear Programming Formulation Techniques,” *Siam Review*, **57**(1), pp. 3–57.
- [84] Godina, R., Rodrigues, E., Poursmaeil, E., Matias, J., and Catalão, J., 2018, “Model Predictive Control Home Energy Management and Optimization Strategy with Demand Response,” *Applied Sciences*, **8**(3), p. 408.
- [85] Izawa, A., and Fripp, M., 2018, “Multi-Objective Control of Air Conditioning Improves Cost, Comfort and System Energy Balance,” *Energies*, **11**(9), p. 2373.
- [86] Kampelis, N., Sifakis, N., Kolokotsa, D., Gobakis, K., Kalaitzakis, K., Isidori, D., and Cristalli, C., 2019, “HVAC Optimization Genetic Algorithm for Industrial Near-Zero-Energy Building Demand Response,” *Energies*, **12**(11), p. 2177.
- [87] Perez, K. X., Baldea, M., and Edgar, T. F., 2016, “Integrated HVAC Management and Optimal Scheduling of Smart Appliances for Community Peak Load Reduction,” *Energy and Buildings*, **123**, pp. 34–40.
- [88] Nagpal, H., Staino, A., and Basu, B., 2020, “Application of Predictive Control in Scheduling of Domestic Appliances,” *Applied Sciences*, **10**, p. 1627.
- [89] Chen, C., Wang, J., Heo, Y., and Kishore, S., 2013, “MPC-Based Appliance Scheduling for Residential Building Energy Management Controller,” *IEEE Transactions on Smart*

- Grid, **4**(3), pp. 1401–1410.
- [90] Lorestani, A., Aghaee, S. S., Gharehpetian, G. B., and Ardehali, M. M., 2017, “Energy Management in Smart Home Including PV Panel, Battery, Electric Heater with Integration of Plug-in Electric Vehicle,” *2017 Smart Grid Conference (SGC)*, pp. 1–7.
- [91] Iqbal, M. M., Sajjad, I. A., Nadeem Khan, M. F., Liaqat, R., Shah, M. A., and Muqeet, H. A., 2019, “Energy Management in Smart Homes with PV Generation, Energy Storage and Home to Grid Energy Exchange,” *2019 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*, pp. 1–7.
- [92] Lu, Q., Zhang, Z., and Lü, S., 2020, “Home Energy Management in Smart Households: Optimal Appliance Scheduling Model with Photovoltaic Energy Storage System,” *Energy Reports*, **6**, pp. 2450–2462.
- [93] Rocha, H. R. O., Honorato, I. H., Fiorotti, R., Celeste, W. C., Silvestre, L. J., and Silva, J. A. L., 2021, “An Artificial Intelligence Based Scheduling Algorithm for Demand-Side Energy Management in Smart Homes,” *Applied Energy*, **282**, p. 116145.
- [94] zheng, “An Integrated Smart Home Energy Management Model Based on a Pyramid Taxonomy for Residential Houses with Photovoltaic-Battery Systems | Elsevier Enhanced Reader” [Online]. Available: <https://reader.elsevier.com/reader/sd/pii/S0306261921005936?token=EAA5655C08C11384A0F10759BEDE85DEDA31A651FD1EF99BED411D29A0EAE53929AB18BFD4EC010C2FC1CD267149356E&originRegion=us-east-1&originCreation=20220403023848>. [Accessed: 02-Apr-2022].
- [95] Zheng, Z., Sun, Z., Pan, J., and Luo, X., 2021, “An Integrated Smart Home Energy Management Model Based on a Pyramid Taxonomy for Residential Houses with Photovoltaic-Battery Systems,” *Applied Energy*, **298**, p. 117159.
- [96] Sou, K. C., Weimer, J., Sandberg, H., and Johansson, K. H., 2011, “Scheduling Smart Home Appliances Using Mixed Integer Linear Programming,” *IEEE Conference on Decision and Control and European Control Conference*, IEEE, Orlando, FL, USA, pp. 5144–5149.
- [97] Veras, J., Silva, I., Pinheiro, P., Rabêlo, R., Veloso, A., Borges, F., and Rodrigues, J., 2018, “A Multi-Objective Demand Response Optimization Model for Scheduling Loads in a Home Energy Management System,” *Sensors*, **18**(10), p. 3207.
- [98] Yahia, Z., and Pradhan, A., 2020, “Multi-Objective Optimization of Household Appliance Scheduling Problem Considering Consumer Preference and Peak Load Reduction,” *Sustainable Cities and Society*, **55**, p. 102058.
- [99] Isermann, R., and Münchhof, M., 2011, *Identification of Dynamic Systems: An Introduction with Applications*, Springer.
- [100] Zhang, X., Rasmussen, C., Saelens, D., and Roels, S., 2022, “Time-Dependent Solar Aperture Estimation of a Building: Comparing Grey-Box and White-Box Approaches,” *Renewable and Sustainable Energy Reviews*, **161**, p. 112337.
- [101] Cai, J., and Braun, J., 2015, “Gray-Box Modeling of Multi-Stage Direct-Expansion

- (DX) Units to Enable Control System Optimization,” *ASHRAE Transactions*, **121**, pp. 203–216.
- [102] Kircher, K. J., and Max Zhang, K., 2015, “On the Lumped Capacitance Approximation Accuracy in RC Network Building Models,” *Energy and Buildings*, **108**, pp. 454–462.
- [103] Ogunsola, O. T., and Song, L., 2013, “Performance Analysis of a Simplified Model of Cooling Load for a Typical Office Building,” *ASME 2013 International Mechanical Engineering Congress and Exposition*.
- [104] Berthou, T., Stabat, P., Salvazet, R., and Marchio, D., 2014, “Development and Validation of a Gray Box Model to Predict Thermal Behavior of Occupied Office Buildings,” *Energy and Buildings*, **74**, pp. 91–100.
- [105] Shaad, M., Momeni, A., Diduch, C. P., Kaye, M., and Chang, L., 2012, “Parameter Identification of Thermal Models for Domestic Electric Water Heaters in a Direct Load Control Program,” *2012 25th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, IEEE, Montreal, QC, pp. 1–5.
- [106] Lakshmanan, V., Sæle, H., and Degefa, M. Z., 2021, “Electric Water Heater Flexibility Potential and Activation Impact in System Operator Perspective – Norwegian Scenario Case Study,” *Energy*, **236**, p. 121490.
- [107] Goh, C. H. K., and Apt, J., 2004, “Consumer Strategies for Controlling Electric Water Heaters under Dynamic Pricing,” Carnegie Mellon Electricity Industry Center Working Paper.
- [108] Lundberg, S. M., and Lee, S.-I., 2017, “A Unified Approach to Interpreting Model Predictions,” *Advances in Neural Information Processing Systems 30*, I. Guyon, U.V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, eds., Curran Associates, Inc., pp. 4765–4774.
- [109] Yu, X., and Ergan, S., 2021, “Key Variables in Determining Energy Shaving Capacity of Buildings during Demand Response Events,” *Journal of Construction Engineering and Management*, **147**(1), p. 04020146.
- [110] Scavuzzo, C. M., Scavuzzo, J. M., Campero, M. N., Anegagrie, M., Aramendia, A. A., Benito, A., and Periago, V., 2022, “Feature Importance: Opening a Soil-Transmitted Helminth Machine Learning Model via SHAP,” *Infectious Disease Modelling*, **7**(1), pp. 262–276.
- [111] Białek, J., Bujalski, W., Wojdan, K., Guzek, M., and Kurek, T., 2022, “Dataset Level Explanation of Heat Demand Forecasting Ann with Shap.”
- [112] Wang, Z., and Jia, L., 2020, “Short-Term Photovoltaic Power Generation Prediction Based on LightGBM-LSTM Model,” *2020 5th International Conference on Power and Renewable Energy (ICPRE)*, pp. 543–547.
- [113] Ltd, R. P., “Raspberry Pi,” Raspberry Pi [Online]. Available: <https://www.raspberrypi.com/>. [Accessed: 24-Apr-2023].
- [114] McPherson, R. A., Fiebrich, C. A., Crawford, K. C., Kilby, J. R., Grimsley, D. L., Martinez, J. E., Basara, J. B., Illston, B. G., Morris, D. A., Kloesel, K. A., Melvin, A. D.,

- Shrivastava, H., Wolfinbarger, J. M., Bostic, J. P., Demko, D. B., Elliott, R. L., Stadler, S. J., Carlson, J. D., and Sutherland, A. J., 2007, “Statewide Monitoring of the Mesoscale Environment: A Technical Update on the Oklahoma Mesonet,” *Journal of Atmospheric and Oceanic Technology*, **24**(3), pp. 301–321.
- [115] Brock, F. V., Crawford, K. C., Elliott, R. L., Cuperus, G. W., Stadler, S. J., Johnson, H. L., and Eilts, M. D., 1995, “The Oklahoma Mesonet: A Technical Overview,” *Journal of Atmospheric and Oceanic Technology*, **12**(1), pp. 5–19.
- [116] Solcast, 2018, “Solar Forecasting Data” [Online]. Available: <https://solcast.com>. [Accessed: 02-Mar-2023].
- [117] OpenWeatherMap, 2018, “Weather API - OpenWeatherMap” [Online]. Available: <https://openweathermap.org/api>. [Accessed: 27-Jan-2023].
- [118] “Rheem Performance 50 Gal. Tall 6 Year 4500/4500-Watt Elements Electric Tank Water Heater XE50T06ST45U1,” The Home Depot [Online]. Available: <https://www.homedepot.com/p/Rheem-Performance-50-Gal-Tall-6-Year-4500-4500-Watt-Elements-Electric-Tank-Water-Heater-XE50T06ST45U1/205810674>. [Accessed: 24-Apr-2023].
- [119] “Low Flow Polypropylene and TFE Liquid Flow Meters” [Online]. Available: https://in.omega.com/pptst/FPR301_302_303_304.html. [Accessed: 24-Apr-2023].
- [120] “Open Energy Data Initiative (OEDI),” OpenEI [Online]. Available: <https://data.openei.org/>. [Accessed: 30-Apr-2023].
- [121] Finck, C., Li, R., Kramer, R., and Zeiler, W., 2018, “Quantifying Demand Flexibility of Power-to-Heat and Thermal Energy Storage in the Control of Building Heating Systems,” *Applied Energy*, **209**, pp. 409–425.
- [122] InfluxDB, 2021, “InfluxDB OSS 2.0” [Online]. Available: <https://docs.influxdata.com/influxdb/v2.0/get-started/>. [Accessed: 27-Jan-2023].
- [123] Grafana, 2021, “Grafana: The Open Observability Platform,” Grafana Labs [Online]. Available: <https://grafana.com/>. [Accessed: 27-Jan-2023].
- [124] Emporia, 2022, “Emporia Vue: Gen 2 Whole Home Energy Monitor,” Emporia Energy [Online]. Available: <https://shop.emporiaenergy.com/products/gen-2-emporia-vue-whole-home-energy-monitor>. [Accessed: 26-Jan-2023].
- [125] Oklahoma Gas & Electric, 2022, “OG&E - SmartHours” [Online]. Available: <https://www.oge.com/wps/portal/ord/residential/pricing-options/smart-hours>. [Accessed: 27-Jan-2023].
- [126] Oklahoma Electric Corporate, 2022, “Time of Use Rate,” Oklahoma Electric Cooperative [Online]. Available: <https://okcoop.org/time-of-use-rate/>. [Accessed: 27-Jan-2023].
- [127] Florida Power & Light, 2022, “Residential Time of Use Rate” [Online]. Available: <https://www.fpl.com/rates/time-of-use.html>. [Accessed: 27-Jan-2023].
- [128] Yeh, C.-C. M., Kavantzias, N., and Keogh, E., 2017, “Matrix Profile VI: Meaningful Multidimensional Motif Discovery,” *2017 IEEE International Conference on Data*

- Mining (ICDM)*, IEEE, New Orleans, LA, pp. 565–574.
- [129] Gharghabi, S., Imani, S., Bagnall, A., Darvishzadeh, A., and Keogh, E., 2018, “Matrix Profile XII: MPdist: A Novel Time Series Distance Measure to Allow Data Mining in More Challenging Scenarios,” *2018 IEEE International Conference on Data Mining (ICDM)*, IEEE, Singapore, pp. 965–970.
- [130] Imani, S., Madrid, F., Ding, W., Crouter, S., and Keogh, E., 2018, “Matrix Profile XIII: Time Series Snippets: A New Primitive for Time Series Data Mining,” *2018 IEEE International Conference on Big Knowledge (ICBK)*, IEEE, Singapore, pp. 382–389.
- [131] Xie, L., 2022, “Development of Envelope Evaluation Benchmarks Using EnergyPlus and Data-Driven Thermal Model.”
- [132] Cetin, K. S., Fathollahzadeh, M. H., Kunwar, N., Do, H., and Tabares-Velasco, P. C., 2019, “Development and Validation of an HVAC on/off Controller in EnergyPlus for Energy Simulation of Residential and Small Commercial Buildings,” *Energy and Buildings*, **183**, pp. 467–483.
- [133] Domanski, P. A., Henderson, H., and Payne, W. V., 2014, “Sensitivity Analysis of Installation Faults on Heat Pump Performance,” NIST.
- [134] “EnergyPlus Version 22.2 Documents” [Online]. Available: <https://bigladdersoftware.com/epx/docs/22-2/>. [Accessed: 04-May-2023].
- [135] “Application Guide for EMS.”
- [136] Lee, S. H., and Hong, T., 2018, *Leveraging Zone Air Temperature Data to Improve Physics-Based Energy Simulation of Existing Buildings*, Lawrence Berkeley National Lab. (LBNL), Berkeley, CA (United States).
- [137] Zhang, H., Cai, J., Fang, K., Zhao, F., and Sutherland, J. W., 2017, “Operational Optimization of a Grid-Connected Factory with Onsite Photovoltaic and Battery Storage Systems,” *Applied Energy*, **205**, pp. 1538–1547.
- [138] “American Time Use Survey—ATUS 2003-2021 Multi-Year Microdata Files : U.S. Bureau of Labor Statistics” [Online]. Available: <https://www.bls.gov/tus/data/datafiles-0321.htm>. [Accessed: 19-Jun-2023].
- [139] Siegel, J., Das, A., Sun, Y., and Pratt, S., 2020, “Safe Energy Savings through Context-Aware Hot Water Demand Prediction,” *Engineering Applications of Artificial Intelligence*, **90**, p. 103481.
- [140] “AWS S3 Explorer for the Open Energy Data Initiative” [Online]. Available: https://data.openei.org/s3_viewer?bucket=oedi-data-lake&prefix=pvdaq%2Fcsv%2Fpvdata%2Fsystem_id%3D2%2F. [Accessed: 13-Apr-2023].
- [141] “Weather Data & Weather API | Visual Crossing” [Online]. Available: <https://www.visualcrossing.com/>. [Accessed: 21-Jun-2023].
- [142] “Enphase IQ Battery 3.36kwh ENCHARGE-3T,” Tandem Solar Systems.
- [143] Ando, Y., Oku, T., Yasuda, M., Ushijima, K., Matsuo, H., and Murozono, M., 2020, “Dependence of Electric Power Flow on Solar Radiation Power in Compact Photovoltaic

- System Containing SiC-Based Inverter with Spherical Si Solar Cells,” *Heliyon*, **6**, p. e03094.
- [144] Hart, W. E., Watson, J.-P., and Woodruff, D. L., 2011, “Pyomo: Modeling and Solving Mathematical Programs in Python,” *Math. Prog. Comp.*, **3**(3), p. 219.
- [145] Hart, W. E., Laird, C. D., Watson, J.-P., Woodruff, D. L., Hackebeil, G. A., Nicholson, B. L., and Sirola, J. D., 2017, *Pyomo—Optimization Modeling in Python*, Springer Science & Business Media.
- [146] Nicholson, B., Sirola, J. D., Watson, J.-P., Zavala, V. M., and Biegler, L. T., 2018, “Pyomo.Dae: A Modeling and Automatic Discretization Framework for Optimization with Differential and Algebraic Equations,” *Mathematical Programming Computation*, **10**(2), pp. 187–223.
- [147] Gurobi Optimization, L., 2021, “Gurobi Optimizer Reference Manual,” p. 969.
- [148] Ciocia, A., Amato, A., Di Leo, P., Fichera, S., Malgaroli, G., Spertino, F., and Tzanova, S., 2021, “Self-Consumption and Self-Sufficiency in Photovoltaic Systems: Effect of Grid Limitation and Storage Installation,” *Energies*, **14**(6), p. 1591.
- [149] Beaudin, M., and Zareipour, H., 2015, “Home Energy Management Systems: A Review of Modelling and Complexity,” *Renewable and Sustainable Energy Reviews*, **45**, pp. 318–335.
- [150] Fanger, P. O., 1972, “Thermal Comfort, Analysis and Application in Environmental Engineering.”
- [151] Fanger, P. O., 1973, “Assessment of Man’s Thermal Comfort in Practice,” *Br J Ind Med*, **30**(4), pp. 313–324.
- [152] Zhang, H., 2003, “Human Thermal Sensation and Comfort in Transient and Non-Uniform Thermal Environments.”
- [153] Jones, N. L., Chaires, I., and Goehring, A., “Detailed Thermal Comfort Analysis from Preliminary to Final Design,” Rome, Italy, pp. 2675–2682.
- [154] Takahashi, Y., Nomoto, A., Yoda, S., Hisayama, R., Ogata, M., Ozeki, Y., and Tanabe, S., 2021, “Thermoregulation Model JOS-3 with New Open Source Code,” *Energy and Buildings*, **231**, p. 110575.
- [155] Reyna, J., Wilson, E., Parker, A., Satre-Meloy, A., Egerter, A., Bianchi, C., Praprost, M., Speake, A., Liu, L., Horsey, R., and others, 2022, *US Building Stock Characterization Study: A National Typology for Decarbonizing US Buildings*, National Renewable Energy Lab.(NREL), Golden, CO (United States).
- [156] Yeh, C.-C. M., Zhu, Y., Ulanova, L., Begum, N., Ding, Y., Dau, H. A., Silva, D. F., Mueen, A., and Keogh, E., 2016, “Matrix Profile I: All Pairs Similarity Joins for Time Series: A Unifying View That Includes Motifs, Discords and Shapelets,” *2016 IEEE 16th International Conference on Data Mining (ICDM)*, IEEE, Barcelona, Spain, pp. 1317–1322.
- [157] Mueen, A., Zhing, S., Zhu, Y., Yeh, M., Kamgar, K., Viswanathan, K., Gupta, C., and Keogh, E., 2022, “The Fastest Similarity Search Algorithm for Time Series Subsequences

under Euclidean Distance.”