OPTIMAL DISTRIBUTED ENERGY RESOURCE CONTROL AND SCHEDULING IN A MICROGRID FRAMEWORK

by

Timothy M. Dodge

A thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Electrical Engineering

Approved:

Greg Droge, Ph.D. Major Professor	Mario Harper, Ph.D. Committee Member
Burak Sarsilmaz, Ph.D. Committee Member	D. Richard Cutler, Ph.D. Vice Provost of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2025

ABSTRACT

Optimal Distributed Energy Resource Control and Scheduling in a Microgrid Framework

by

Timothy M. Dodge, Master of Science

Utah State University, 2025

Major Professor: Greg Droge, Ph.D.
Department: Electrical and Computer Engineering

Advanced coordination within microgrids is essential for integrating renewable energy sources and distributed energy resources (DER) to enhance grid stability and operational efficiency. This research introduces a significant contribution through a hierarchical optimization framework employing receding-horizon control (RHC) for the effective coordination of electrical DERs, with a particular focus on energy storage systems (ESS) and electric vehicle (EV) chargers. This framework, leveraging load forecasting and a novel sequential optimization strategy, demonstrably achieves substantial improvements in microgrid operation, including simulated peak load reductions from 301 kW to approximately 150 kW in day-ahead tests and successful management of dynamic demand response signals by prioritizing objectives like peak power minimization, load profile smoothing, and state-of-charge management. Recognizing the critical role of thermal loads, this work further advances HVAC control methodologies for microgrids through two distinct, impactful approaches. A novel physics-based Resistor-Capacitor (RC) modeling methodology, featuring an innovative model incorporating water storage and a generalizable multizone graph formulation, demonstrates inherent model stability and optimal control capabilities, while also highlighting challenges in translating these optimal plans to simpler controllers. Concurrently, a data-driven approach utilizing linear system identification achieves promising

prediction precisions (FIT scores ranging from approximately 33% to over 98%) across diverse simulation environments. This data-driven optimization of HVAC set points reveals clear performance trade-offs, enabling, for instance, an approximate 20% peak power reduction alongside a 19% total energy reduction compared to baseline in the BESTEST Case 900 simulation. While this method shows strong potential, challenges in more complex, high-fidelity environments (such as the two-zone hydronic apartment with FIT scores as low as 33%) underscore the need for more advanced identification techniques. Although the full integration of these HVAC strategies represents an avenue for future work, this thesis provides a validated, comprehensive framework for electrical DER coordination and lays significant foundational groundwork in both physics-based and data-driven HVAC control suitable for advanced microgrid applications.

(199 pages)

PUBLIC ABSTRACT

Optimal Distributed Energy Resource Control and Scheduling in a Microgrid Framework

Timothy M. Dodge

As we use more renewable energy, such as solar power, and add new devices, such as electric vehicle chargers and battery storage, to our buildings, the management of electricity becomes more complex. These local energy sources and devices can form small "microgrids" that need careful coordination to work efficiently with the main power grid. This research developed a smart planning system to help manage these micro-grids. The system figures out the best times to use, store or charge different devices (such as batteries and EVs) to avoid costly, high electricity demand spikes and help stabilize the main power grid, especially when asked by the utility company. A major part of this work involved creating better ways to control building heating and cooling (HVAC) systems, which are big energy users. Two methods were developed: one learns from past energy use data to predict how temperature settings affect energy consumption, and another uses simplified physics models to understand how heat moves in a building. By intelligently scheduling all of these devices together, our aim is to make our energy use cheaper, more reliable, and better integrated with renewable resources.

Dedicated to my fiancée, Paige, for your enduring patience, support, and love throughout this journey.

ACKNOWLEDGMENTS

I wish to express my sincere gratitude to my advising professor, Dr. Greg Droge. Without his guidance and support, the completion of this work would not have been possible. When I began working in his lab, I could not have envisioned producing a body of work as expansive as that presented herein. Through my association with him, I have grown significantly, not only as an engineer but also as a communicator and writer.

I also extend my thanks to my committee members, Dr. Burak Sarsilmaz and Dr. Mario Harper. Their expertise provided valuable inspiration for this research and their guidance was instrumental in developing skills that greatly benefited this work.

Finally, I would like to thank my friends James Sweeden and Cody Grogan for their assistance with the analysis and development of the forecasting models outlined in the appendix.

Timothy Dodge

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYMS

Table 1: Acronyms

Acronym	**Definition**
AC	Alternating Current
AHU	Air Handling Unit
$ARIMAX$	Auto-Regressive Integrated Moving Average with Exogenous Input
$BOPTEST$	Building Optimization Performance Testbed
$CCCV$	Constant Current Constant Voltage
CNN	Convolutional Neural Network
$CONV-LSTM$	Convolutional Long Short Term Memory
CPP	Critical Peak Pricing
$DARE$	Discrete-time Algebraic Riccati Equation
DER	Distributed Energy Resources
DNN	Deep Neural Networks
EM	Energy Management
EMS	Energy Management System
ESS	Energy Storage System
EV	Electric Vehicle
FCU	Fan Control Unit
FLC	Fuzzy Logic Controller
$HVAC$	Heating, Ventilation and Air Conditioning
ISO	Independent System Operator
$LSTM$	Long Short Term Memory
$MDPs$	Markov Decision Processes
$MILP$	Mixed Integer Linear Program

Table 1 – continued from previous page

Acronym	Definition
$MINLP$	Mixed Integer Non-Linear Program
MLP	Multi-Layer Perceptron
$MOGA$	Multi-Objective Genetic Algorithm
MPC	Model Predictive Control
$NOCT$	Normal Operating Cell Temperature
$NSGA-III$	Constrained Non-Dominated Sorting Genetic Algorithm
$ODEs$	Ordinary Differential Equations
$OSQP$	Operator Splitting Quadratic Program (Solver)
$PDEs$	Partial Differential Equations
PID	Proportional Integral Derivative
PV	Photovoltaic
QP	Quadratic Programs
RC	Resistance-Capacitance
RHC	Receding Horizon Control
SOC	State Of Charge
STC	Standard Test Conditions
$3TC$	Third-Constant Model
TOU	Time-of-Use
$XGBoost$	Extreme Gradient Boosting

CHAPTER 1

INTRODUCTION

With the rapid advancement of renewable energy and power storage technologies, the need for localized and distributed grid services has become increasingly evident. This transition has led to the rise of microgrids, small-scale localized electrical systems that integrate various Distributed Energy Resources (DERs). Microgrids can operate autonomously or coordinate with larger electrical grids, offering a reliable and flexible electricity supply to specific regions or interconnected buildings. DERs, including renewable energy sources such as photovoltaic (PV) systems and wind turbines, and controllable loads such as heating, ventilation, and air conditioning (HVAC) systems, water heaters, energy storage systems (ESS) and electric vehicle (EV) chargers play a crucial role in decentralizing energy systems [7].

As microgrids become more prevalent, they provide significant advantages by increasing energy resilience and allowing more granular control over energy production and consumption. Utilities have started incentivizing electrical load management through these systems, recognizing their potential to enhance the overall robustness and flexibility of the larger grid. However, effective integration of DERs into microgrids remains a complex and critical challenge, which requires careful coordination of various energy sources and consumption patterns.

1.1 The Problem

The primary focus of this research is a comprehensive set of strategies for controlling and optimizing DER scheduling within microgrids. The central objective is to maximize the utilization of the main power grid while effectively responding to the demands of the utility (i.e., demand response). The use of the main power grid involves minimizing peak power consumption and flattening the overall power profile, which in turn reduces the load on the

grid and enhances its stability. Demand response strategies complement this by adjusting power consumption in real-time based on the utility company's needs, either lowering or increasing usage to support grid operations. This requires not only sophisticated forecasting and modeling, but also optimization strategies that can prioritize critical operational goals such as peak reduction while managing diverse resources, including complex thermal loads.

Maximizing the utilization of the main power grid offers two key benefits. First, reducing peak loads eases the burden on the grid during periods of high demand, helping prevent main power grid instability and power outages. Second, flattening power consumption improves the ability of the grid to predict energy use and generation of the microgrid, leading to more efficient power distribution and management. Demand response plays a crucial role in enhancing the flexibility and resilience of the grid. By dynamically modifying energy consumption in response to real-time grid conditions, demand response enables better alignment between supply and demand, ultimately improving the integration of renewable energy sources. Through a detailed examination of microgrid management, this research provides insight that enables more efficient and cost-effective energy distribution. By optimizing both the use of the grid and the response to demand, this work contributes to the overall sustainability and robustness of modern power systems.

The increasing reliance on renewable energy sources, such as solar and wind, introduces significant variability in energy supply due to their stochastic and intermittent nature. These fluctuations pose challenges to main power grid stability, with the "Duck Curve" phenomenon [8] exemplifying this problem. The "Duck Curve" reflects a mismatch between high solar power generation during midday, when demand is lower, and the sharp increase in demand during the evening, when solar generation declines. Managing this imbalance requires advanced strategies to integrate renewable energy into microgrids.

Energy storage systems are essential for managing the inherent variability of renewable energy sources within microgrids. These systems provide the flexibility to store surplus energy during periods of low demand or high generation and release it when necessary, thereby reducing reliance on traditional, often less sustainable, grid-based power.

HVAC systems are also an essential part of microgrid planning. According to the U.S. Energy Information Administration's Residential and Commercial Building Energy Consumption Survey, HVAC accounts for approximately 19% of electricity consumption for US households and 14% for US commercial buildings [9]. Planning of HVAC systems power consumption is challenging since thermal systems are complex and nonlinear. In addition, user comfort must also be considered since if users override smart controls, these plans can be invalidated.

Although minimizing power consumption is often seen as the key contribution of microgrids, the utility of the grid involves more nuanced goals. In addition to reducing peak demand, there are instances in which the grid prefers higher power utilization to balance supply and demand, stabilize the system, or optimize energy storage use. This research optimizes DER scheduling not only for power minimization, but also to enhance overall grid utility by addressing these broader objectives.

1.2 Previous Work

In the literature, the management of DERs within microgrids is typically divided into three primary categories: Forecasting, Energy Management (EM) Optimization, and Real-Time Control [10]. Each category addresses a distinct aspect of microgrid management, and research tends to focus heavily on one area while providing limited attention to others, resulting in a compartmentalized approach.

Forecasting is critical for managing microgrids, particularly for predicting renewable energy generation and consumption patterns. For example, [11] combines long-short-term memory (LSTM) neural networks to forecast power generation with a mixed integer linear programming (MILP) approach to control energy consumption. However, [11] focuses exclusively on ESS, neglecting other important DERs, such as solar inverters and EV chargers, which are becoming increasingly relevant in modern microgrids.

Energy Management Optimization involves optimizing and coordinating the load that each DER should provide or consume. In [12], a multi-objective optimization framework is developed to manage ESS, wind generators, PV systems, and EV chargers. Although this

approach offers a broader view of DER coordination, it lacks predictive capabilities and does not account for real-time variability, limiting its adaptability to the stochastic nature of renewable energy sources.

Another example of Energy Management Optimization is presented in [13], where the authors introduce an optimization-based model predictive control (MPC) strategy that specifically targets HVAC systems within a microgrid. The focus is on reducing emissions, minimizing operational costs and maintaining human comfort, but other loads, such as electric vehicle chargers and photovoltaic systems, are treated as uncontrolled variables. This reduces system flexibility and limits the coordination of all DERs within the microgrid.

In contrast, [14] develops a single-objective MPC framework to trade energy with the grid, considering HVAC systems, photovoltaic, ESS and EV chargers. Although it expands the control strategy to include multiple DERs, its focus remains on cost minimization, without fully exploring the multi-objective potential of DER integration to enhance grid utility.

An essential aspect intertwined with EM optimization is the accurate modeling and control of thermal systems, particularly HVAC systems. Thermal modeling plays a critical role in predicting building energy consumption and designing effective control strategies for HVAC operations within microgrids. For example, Tindale [4] introduces a third-order lumped parameter simulation method called the "third constant model" (3TC). This model addresses the limitations of simpler second-order models, which can be overly responsive when applied to medium to heavy-weight buildings, by incorporating additional thermal dynamics to provide a more accurate representation of such structures. Incorporating advanced thermal modeling techniques, such as 3TC, improves the precision of energy consumption forecasts and improves the coordination of HVAC systems with other DERs, ultimately contributing to more efficient and flexible microgrid operations.

Real-time control deals with low-level management, such as voltage and frequency regulation. For example, [15] proposes an optimal cooperative voltage regulation controller designed for second-stage control. This type of real-time control focuses on ensuring grid

stability and addressing technical challenges at a granular level, rather than higher-level energy management.

1.3 Gaps in Existing Work

While the existing literature addresses individual aspects such as forecasting, DER modeling, or optimization, few studies provide a validated integrated framework that simultaneously coordinates various electrical and thermal DERs. Current studies often focus on one aspect, such as effective forecasting with stochastic considerations, detailed modeling of complex DERs like HVAC systems, or sophisticated optimization strategies for energy management, but generally do not demonstrate a unified, validated approach that addresses the co-optimization of electrical storage (ESS), electric vehicles (EVs), and complex HVAC systems within a single, adaptable control architecture.

This thesis begins to bridge that gap. First, it presents and validates a hierarchical receding horizon control (RHC) architecture for coordinating electrical DERs (ESS, EVs and PV generation), driven by an uncontrolled power forecasting model and incorporating dynamic demand-response signals as soft constraints. Second, recognizing the critical role of thermal loads, it develops two distinct and advanced HVAC control methodologies, one based on data and one based on physics, as foundational work specifically designed for future integration into such hierarchical frameworks. Validating these advanced HVAC methodologies in high-fidelity simulations demonstrates their individual potential and establishes the necessary foundation for their future integration into the RHC scheduling framework.

1.4 Contributions

This work presents significant advances in building energy management by developing and validating key components for coordinated control. As illustrated in Figure 1.1, these contributions collectively address the core problem of enhancing grid utilization. This overarching goal is pursued through three sub-objectives, each tackling distinct aspects of this challenge. The first sub-objective, local grid utilization, aims to enhance local load profiles by shaving peaks, smoothing demand, and optimizing local battery storage. The

second, demand response, enables the microgrid to dynamically respond to the wider grid's needs. These two aspects are primarily addressed by **Contribution 1**, which details a Receding Horizon Control (RHC) framework for scheduling battery energy storage systems (ESS) and electric vehicle (EV) chargers, incorporating photovoltaic (PV) generation. The third sub-objective, smart HVAC control, targets the significant energy load from heating, ventilation, and air conditioning systems. This is addressed by **Contribution 2**, focusing on a physics-based lumped-sum HVAC modeling approach, and **Contribution 3**, which develops a complementary black-box optimal control framework for HVAC systems. The specific contributions are detailed below:

Contribution 1: **A hierarchical, receding horizon control (RHC) system was developed and evaluated to coordinate electrical energy storage systems (ESS) and electric vehicle (EV) chargers, considering photovoltaic generation.** This system uses external load forecasting for load prediction and employs a novel sequential optimization approach that prioritizes operational objectives: minimizing peak power, minimizing power range, and maximizing battery state of charge (SOC). The framework incorporates demand response signals as dynamic soft bounds, and simulation results demonstrated significant peak power reduction.

Contribution 2: **A physics-based optimal control methodology for HVAC systems was developed using lumped-sum Resistor-Capacitor (RC) thermal modeling.** This included adapting existing models to a 5R5C structure incorporating water storage heating dynamics and creating a systematic graph-based formulation to generate state-space models of complex multizone systems. The inherent stability of these models was proven (Lemma 1), and optimal control strategies (formulated as Quadratic Programs) were applied, demonstrating effective temperature regulation and the ability to incorporate constraints such as peak power

minimization, although highlighting challenges in translating optimized power plans to simpler controllers such as those based on hysteresis.

Contribution 3: Complementing the physics-based approach, a black-box optimal control framework for HVAC set-point control was developed using linear regression (ARMAX). This data-driven model, identified using high-fidelity building thermal envelope and HVAC simulation platforms (like BOPTEST), maps temperature set-points and external conditions to energy consumption. Its use within an optimization framework demonstrated the potential to manage HVAC energy use according to various objectives, providing a pathway to integrate standard thermostat control into larger energy management schemes.

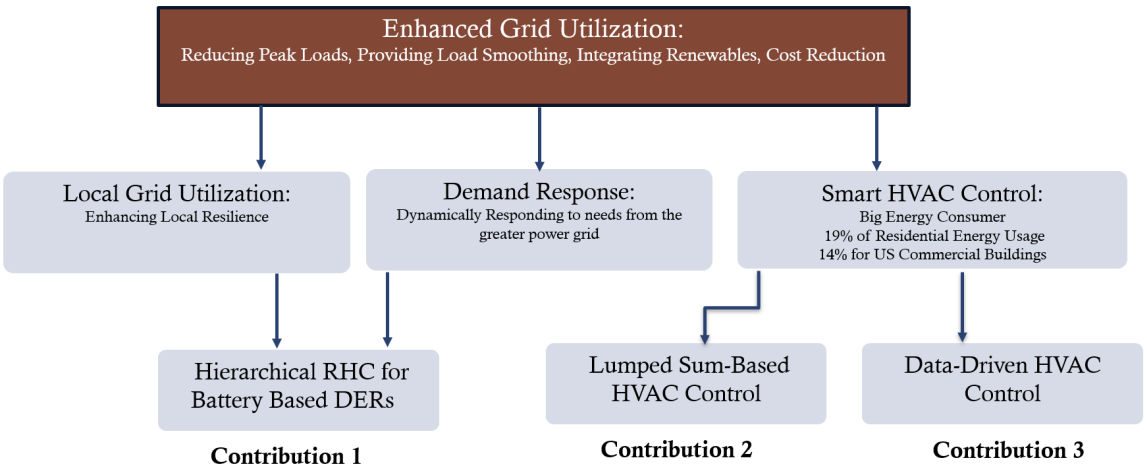

Fig. 1.1: Diagram of Contributions illustrating the overarching goal of enhanced grid utilization and the supporting sub-objectives addressed by each contribution.

The remainder of this document outlines the contributions of this work. Chapter 2 provides an overview of previous research and a detailed presentation of the contributions developed here. Chapter 3 presents the hierarchical control framework developed for ESS and EV chargers, detailing the sequential optimization objectives, the forecasting approach,

the integration of demand response, and the simulation results validating its effectiveness in reducing peak loads. Chapter 4 describes the physics-based HVAC control methodology using lumped-sum RC modeling, presenting the 5R5C model, the general multizone graph formulation, the stability analysis, and the results from optimal control simulations. Chapter 5 details the development and integration of the black-box regression-based HVAC control framework, including model identification using BOPTEST and the formulation for combined DER optimization. Finally, Chapter 6 summarizes these contributions and discusses potential directions for future work.

CHAPTER 2

BACKGROUND and OBJECTIVES

This chapter aims to thoroughly outline the objectives of the proposed work. It begins with Section 2.1, which provides a detailed overview of the chapter and provides a visualization of comparing the existing literature. Next, we explore relevant background information for microgrid planning in Section 2.2. Section 2.3 provides an overview of the process of modeling the dynamics of all DERS contained in this work, except for HVAC. Section 2.4 provides a detailed overview of all HVAC control modeling methods. Section 2.5 provides background and definitions for the concepts of grid utilization and demand response. Lastly, Section 2.6 offers an overview of the motivation and overarching goals of the combined research efforts.

2.1 Literature Review

This section provides a comprehensive review of the existing literature relevant to the challenges and objectives addressed in this thesis. It delves into the key areas that underpin optimal scheduling and control of distributed energy resources (DER) within microgrids. First, Section 2.2 examines the landscape of microgrid planning, focusing specifically on the crucial aspects of Forecasting and Energy Management (EM) Optimization, which are central to this work. Subsequently, Sections 2.3 and Section 2.4 explore the state-of-the-art in modeling various DERs, including energy storage systems (ESS), electric vehicle (EV) chargers, photovoltaic (PV) systems, and particularly the complexities associated with Heating, Ventilation, and Air Conditioning (HVAC) systems. Finally, Section 2.5 discusses pertinent concepts related to grid utilization and demand response strategies. This review serves to contextualize the contributions of the present research within the broader field and highlights the specific gaps that this work aims to address.

To provide a comparative overview of the features addressed in existing works within

these domains, Table 2.1 summarizes key characteristics of relevant studies, highlighting aspects such as the optimization techniques used, the DERs considered, the grid utility features and forecasting capabilities. This table will serve as a reference point throughout the subsequent review.

	EM Optimization									**Grid Utility**		**Forecasting**	
	MILP	MINLP	Meta-Heuristic Algorithm	Hierarchical Optimization	RHC/MPC	ESS Consideration	EV Consideration	HVAC Consideration	PV Consideration	Dynamic Pricing	Demand Response	Generation Forecasting	Demand Forecasting
[11]	✓	-	-	-	✓	-	-	-	✓	-	-	✓	✓
[12]	-	-	✓	-	-	✓	✓	-	✓	✓	✓	-	-
[13]	✓	-	-	-	✓	-	-	✓	✓	-	-	-	-
[16]	-	✓	-	✓	-	✓	✓	-	-	-	-	-	-
[17]	✓	-	-	-	-	-	-	-	-	-	-	-	-
[18]	✓	-	✓	-	-	✓	-	-	✓	✓	-	-	-
[19]	-	-	✓	-	✓	-	-	✓	-	-	-	-	-
[20]	✓	-	-	✓	-	✓	-	-	✓	✓	✓	-	-
[21]	-	-	-	✓	-	-	-	✓	-	✓	✓	-	-
[22]	-	-	-	-	-	-	-	-	✓	-	-	-	✓
[23]	-	-	-	-	-	-	-	-	-	-	-	✓	-
[24]	-	-	-	-	-	✓	-	-	✓	✓	-	✓	✓
[25]	-	-	-	✓	-	✓	✓	-	✓	✓	-	-	-
Proposed	✓	-	-	✓	✓	✓	✓	✓	✓	-	✓	✓	✓

Table 2.1: Overview of features in the literature on microgrid management.

2.2 Microgrid Planning

The field of microgrid planning is commonly segmented into three distinct categories: Forecasting, Energy Management (EM) Optimization, and Real-Time Control [10]. This structure provides context for the contributions of this work. Each category addresses different aspects required for the effective aggregation, planning, and control of DERs within a microgrid.

Forecasting involves predicting the behavior of uncontrolled loads over various time horizons: short-term (seconds to minutes), medium-term (one hour to 24 hours), and long-term (one day to one week). Accurate forecasting is essential for effective planning and scheduling within a microgrid, as it allows for better anticipation of energy needs and generation patterns. EM Optimization focuses on the strategic planning and coordination of DERs to ensure efficient energy supply while aiming to reduce costs. This involves optimizing how different DERs interact to maximize the benefits of the microgrid system. Real-Time Control deals with the precise regulation of microgrid components, addressing variables such as voltage, frequency, and power. The focus of this work is on the first two categories, so the remainder of this section will give a greater context and background to forecasting and EM optimization.

This work primarily focuses on Forecasting and EM Optimization, with Real-Time Control intentionally excluded from the scope. The remainder of this section will provide a comprehensive overview and context for these two categories, outlining their significance and application in microgrid management. By understanding these aspects in greater detail, the subsequent sections will establish the foundation for the proposed methodologies and their impact on optimizing microgrid performance.

2.2.1 Forecasting

Forecasting is a fundamental aspect of developing an effective microgrid energy management system. It enables the accurate prediction of variables such as weather conditions, energy generation, and load demand, which is crucial for optimizing DER utilization. Precise forecasting helps organize DERs effectively and is essential for managing objectives

such as reducing peak power usage, where precise predictions of demand and generation are critical.

Forecasting techniques commonly employ machine learning, deep learning, and ensemble methods. These techniques require defining an error metric or loss function to evaluate model performance. These loss functions are used to compare the output of the network with the target. In this case, the target would be some section of time that happens after the training data occur that the system can use to predict.

Much of the work in forecasting electrical loads is done without considering the other objectives in microgrid planning. For example, [22] reviews various ensemble methods for forecasting, focusing on short-term predictions and providing information on the effectiveness of different approaches for forecasting and control. In [23], a multilayer perceptron (MLP) is demonstrated to be useful for short-term forecasting of wind power generation.

When forecasting work considers additional objectives, the scope of these objectives is typically limited. For example, [11] employs long-short-term memory (LSTM) networks to forecast PV power output and load demand over a 24-hour horizon. Although this work incorporates some planning, which will be discussed in the next section, it remains limited in scope. Another example of this is [26]. This work also leverages an LSTM and uses an evolutionary algorithm known as MOGA to generate a Pareto front that represents trade-offs between conflicting goals, such as minimizing the cost of grid power purchase and the cost of battery degradation.

These works demonstrate the diverse applications of forecasting power consumption and demand, showcasing the potential of machine learning, ensemble learning, and deep learning methodologies in this domain. Using these advanced techniques, researchers have achieved accurate and reliable power forecasts, which is crucial for effective microgrid management. Moreover, these studies highlight the necessity of integrating sophisticated forecasting methods into more complex problem formulations, such as multi-objective optimization and the coordination of various Distributed Energy Resources (DERs).

2.2.2 Energy Management Optimization

In microgrids, the presence of multiple sources of power generation and demand requires effective EM Optimization to ensure efficient and reliable operation. EM Optimization involves several critical decisions: determining which DERs to control, modeling the DERs and the broader system accurately, designing control strategies for individual DERs, defining overarching optimization objectives, and selecting appropriate mathematical methods to achieve these objectives.

Several studies have addressed various aspects of EM Optimization in microgrids. For example, [11] provides a solution for controlling battery storage within a microgrid to minimize energy cost relative to a static pricing structure. To optimize this objective, the authors propose a Mixed Integer Linear Programming (MILP) formulation along with a day-ahead planning structure. They also incorporate a Receding Horizon Control approach to account for stochastic variations, enhancing the system's responsiveness to uncertainties.

Similarly, [17] develops a day-ahead optimal scheduler for DERs based on electrical and natural gas usage. This solution considers operating constraints of the alternating current (AC) to ensure network security and comply with reactive power requirements. By linearizing the power equations, the problem is formulated as a MILP while still accounting for AC power flow. However, this work does not include considerations for stochastic elements in the optimization process.

Some EM optimization techniques opt for more straightforward approaches [25, 27] to develop a rule-based energy management system (EMS) for an isolated microgrid, structured in a hierarchical control architecture with the EMS as a high-level controller. Unlike traditional approaches, this RB-EMS avoids complex optimizations and long prediction horizons; instead, it uses a set of expert if–then rules based on historical data and very short-term forecasts to determine the set points for distributed generators, energy storage, and controllable loads.

Other EM optimization techniques leverage the Fuzzy Logic Controller (FLC) to handle uncertainties in electricity demand and supply [24, 28]. The work presented in [24] presents

an energy management system for a solar-diesel-battery microgrid based on a Fuzzy Logic Controller (FLC) to handle the uncertainty of renewable supply and load demand. Fuzzy EMS uses short-term generation and load forecasts as inputs and a set of fuzzy rules to decide real-time power splits - aiming to maximize the use of PV power and minimize diesel generator operation while keeping battery SOC within safe limits.

In another study, [16] focuses on a small-scale microgrid, specifically a single-family home with an energy conversion unit of less than 15 kW capacity. The emphasis of this work is on heat-led control, which means that the focus of control is the thermal energy consumption of the building. The authors consider DERs such as HVAC systems, ESS, and EV charging. They formulate a Mixed Integer Non-Linear Programming (MINLP) problem to solve a single-objective optimization focused on minimizing the cost of electricity and natural gas.

Expanding on multi-objective optimization, [12] proposes a comprehensive model that considers user preferences, demand response, and real-time electricity pricing. This work includes a wider array of DERs, such as ESSs, wind turbines, photovoltaic systems, and electric vehicle chargers. To optimize these objectives simultaneously, the authors employ a Constrained Non-Dominated Sorting Genetic Algorithm (NSGA-III), which effectively handles multiple conflicting objectives and provides a set of optimal solutions.

In [13], the authors consider the control of a thermal microgrid, focusing on DERs such as photovoltaic systems, HVAC systems, and hot water tanks. They implement lower-level controllers, such as Proportional-Integral-Derivative (PID) and rule-based controllers, for individual DERs. These controllers are integrated into a higher-level Model Predictive Control (MPC) framework to coordinate all the DERs within the microgrid. The work presents two different MPC-based approaches: one optimization-based and the other optimization-free. Although the optimization-based approach yields better performance in terms of meeting the control objectives, it requires significantly more computational time compared to the optimization-free method.

Although several studies have advanced EM optimization using MILP-based formulations and multi-objective genetic algorithms, alternative approaches also offer promising benefits. For example, Lu et al. [18] compare a traditional MPC framework with an optimization-free method to manage photovoltaic and battery systems, highlighting the trade-off between computational burden and control performance. Similarly, Hassan et al. [21] introduce a hierarchical framework that integrates a Markov process to discretize building energy consumption and uses a Markov Decision Process for thermostat control, thereby embedding stochastic considerations directly in optimizing the grid utility and SOC of batteries. Furthermore, Sandgani et al. [19] employ a particle swarm optimization technique to coordinate diverse renewable sources such as photovoltaic, wind and fuel cells, effectively addressing nonlinearities and varying operational scales. These approaches not only complement the MILP/MINLP and NSGA-III methods described previously, but also suggest directions for reducing computational complexity and enhancing system robustness in real-world microgrid management.

Together, these studies highlight the diverse methodologies and considerations in microgrid microenvironment optimization. They underscore the importance of selecting appropriate control strategies and optimization techniques tailored to the specific configurations and objectives of the microgrid. By addressing various DERs, modeling approaches, and optimization methods, these works contribute to the advancement of efficient and resilient microgrid energy management.

2.3 Distributed Energy Resource Modeling

Modeling Distributed Energy Resources (DERs) is essential to optimize grid utility and ensure efficient energy management within microgrids. This section provides a comprehensive background on modeling the DERs considered in this work, including ESS, PV systems and EV chargers. HVAC modeling is left for a seperate section due to the inherent complexity involved in modeling thermal dynamics. Accurate HVAC modeling is crucial because it significantly affects overall energy consumption, occupant comfort, and the integration of renewable energy sources within the microgrid. The complexity arises from the need

to accurately capture the interactions between various thermal components, environmental factors, and energy inputs.

2.3.1 ESS and EV Chargers

When modeling and controlling ESS and EV chargers, the fundamental objective is to influence the state of charge (SOC) of the battery by applying power (P). The primary distinction between these two DERs lies in their operational capabilities: EV chargers are typically unidirectional, meaning that they can only charge the battery and are not designed to supply power back to the grid, unlike certain ESS, which offer bidirectional power flow. In this work, the charging behavior of the battery is assumed to be proportional within certain limits. This assumption is valid because lithium-ion battery charging typically follows a Constant Current Constant Voltage (CCCV) curve, which exhibits a nearly linear increase in SOC during the constant current phase [29]. Specifically, the change in SOC (ΔS) is directly proportional to the power input (P) applied over a time interval (Δt), as described by Equation 3.26:

$$\Delta S = k_c P \Delta t \tag{2.1}$$

Where k_c is the proportionality constant representing the efficiency of the charging process. Figure 2.1 illustrates this proportional charging behavior, with the SOC depicted in green. The figure demonstrates that within a certain range, the charging profile maintains a proportional relationship between the power input and the increase in SOC. For this work, restrictions on the SOC are made so that it can only be charged while remaining in the SOC where proportionality is assumed.

2.3.2 Photovoltaic Systems (Solar Panels)

PV power generation converts solar radiation into electrical energy through the photoelectric effect [30]. Photovoltaic systems are recognized as the safest, most reliable, and pollution-free among renewable energy resources. However, their dependence on solar radiation restricts their operation to daytime hours and makes them susceptible to fluctuations

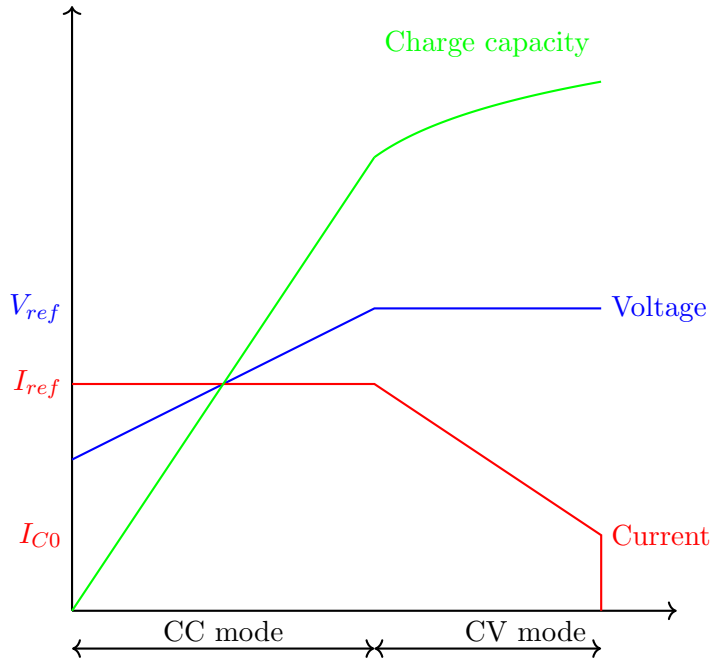

Fig. 2.1: Charging profile of Lithium Ion (Charge capacity shown in green)

as a result of weather conditions. In the context of DERs scheduling, the microgrid cannot directly control the power output of PV systems. Instead, the microgrid can only decide whether to send the generated power back to the grid or utilize it locally, necessitating an accurate prediction of the power output rather than control. There are two primary approaches for this: modeling the power output based on solar radiation or using historical data to forecast future power generation. [31] presents an analytical model for predicting PV power output, defined by the equation:

$$P_{PV}(t) = \frac{G_t(t)}{G_{\text{ref}}} \times P_{PV-\text{STC}} \times \eta_{PV} \times [1 - \beta_T(T_c - T_{C-\text{STC}})] \tag{2.2}$$

where $P_{PV}(t)$ denotes the PV power output at time t, $G_t(t)$ represents the solar radiation at time t, G_{ref} is the reference solar irradiance set at 1000 W/m^2, $P_{PV-\text{STC}}$ is the rated power of the PV panel under Standard Test Conditions (STC), η_{PV} is the generation efficiency, β_T is the temperature coefficient, T_c is the cell temperature, and $T_{C-\text{STC}}$ is the cell reference

temperature. The cell temperature T_c is calculated using the equation:

$$T_C = T_{\text{amb}} + (NOCT - 20) \times \frac{G_t(t)}{800} \tag{2.3}$$

where T_{amb} is the ambient temperature and $NOCT$ is the Normal Operating Cell Temperature.

The downside of this method is that it still must rely on forecasting instead of directly forecasting power generation solar radiation to be forecasted. This means that the guarantees provided by the first-principle modeling still rely on stochastic methods. It follows then that directly predicting power generation is equally reliable and requires fewer intermediate steps. [32] investigates several deep learning methods to forecast future power generation, including autoencoders, LSTM, and CNNs.

2.4 HVAC Systems

Effective control and optimization of building HVAC systems depend on accurate thermal modeling. There are three primary approaches: white-box, gray-box, and black-box modeling. White-box models are derived from fundamental physical laws (first principles). In contrast, black-box models learn input-output relationships directly from data, treating the system's internal workings as unknown. Gray-box models blend these techniques, using physics to establish a model structure whose specific parameters are then estimated from the data. Each of these modeling approaches will be elaborated upon in the subsequent sections.

2.4.1 White-box modeling

White box modeling is based on fundamental physical principles such as the conservation of mass, energy, and momentum [33]. This bottom-up method involves detailed modeling of the thermal envelope of the building, including occupants, schedules, HVAC systems, and various parameters. Although these models are derived from first principles, they often simplify higher-order terms that have minimal impact on the system's response to reduce

computational complexity. These methods are typically implemented using comprehensive simulation software packages such as EnergyPlus [34], ESP-r [35], and Modelica [36].

Modeling the exact flow of heat energy throughout a building is not necessary to develop control systems to manage the temperature in a building. A typical thermostat does not consider the temperature throughout the room; it typically considers regulating the temperature at a small number of points, often just one. Thus, the model of heat flow throughout the building can be simplified by modeling the flow between points of interest. The basic premise behind lumped-sum modeling is that the building is decomposed into various 'lumps', the 'sum' of which constitutes the entire building. These "lumps" are modeled as nodes in a thermal schematic model, e.g., Fig. 2.2, which are connected to a thermal capacitor. A node may correspond to air in part of a room, the whole room, the surface of a wall, a portion of the inside of a wall, etc. The adjacent nodes are connected via thermal resistance, which models the conduction, convection, and thermal radiation mechanisms between lumps. The combination of thermal capacitors and resistors in an electric circuit-like architecture is used to model the flow of heat between lumps/nodes.

The concept of lumped sum modeling is not unique to thermal systems. Direct parallels are found in mechanical translation, mechanical rotation, electrical systems, and fluid systems [37,38]. In the same way that the voltage difference represents the potential energy of a circuit, the temperature difference, measured in Kelvin, is the potential energy of a thermal system. The electrical current represents the flow of electrons through the circuit; the heat flow rate, measured in joules per second, represents the transfer of heat energy through the system. The electrical resistance describes the opposition to the flow of electrons in a circuit; the opposition in a material or object to the flow of heat is referred to as thermal resistance, measured in squared meters-Kelvin-seconds/Joule. Finally, the thermal capacitance of a lump, measured in Joules and Kelvin, is directly related to the capacitor found in a circuit. Note that there is no equivalence to an inductor within a thermal system.

The phrase "direct parallel" is used as the mathematical equations for a circuit are equivalent to those of a lumped-sum thermal model. Consider zones 1 and 2 associated

with temperatures T_1 and T_2. Let Q_{21} denote the flow rate from zone 2 to zone 1, and let R_{21} denote the resistance between the two zones[1]. The flow can be determined using Ohms law as,

$$Q_{21} = \frac{1}{R_{21}}(T_2 - T_1). \tag{2.4}$$

In the same fashion as in a circuit, thermal capacitance provides a dynamic relationship for heat flow. Given a capacitor with capacitance C_1 connected to the node associated with the temperature of zone 1, T_1, the heat flow rate into the material associated with zone 1 is modeled as

$$Q_1 = C_1 \frac{d}{dt} T_1 = C_1 \dot{T}_1, \tag{2.5}$$

where the dot notation denotes the time derivative.

Kirkoff's current law states that the flow into a node must be equal to the flow out of a node. The same applies to the heat flow rate in thermal systems. For example, Fig. 2.2 shows three flows into and a single flow out of zone 1. Kirkoff's current law provides a relationship that can be used to determine the dynamic equation for T_1 in terms of the temperatures of the surrounding components as

$$C_1 \dot{T}_1 = \frac{1}{R_{01}}(T_0 - T_1) + \frac{1}{R_{21}}(T_2 - T_1) + \frac{1}{R_{31}}(T_3 - T_1). \tag{2.6}$$

Or, written more succinctly,

$$\dot{T}_1 = \frac{1}{C_1}\left(\sum_{j=2}^{3} \frac{1}{R_{j1}}(T_j - T_1) + \frac{1}{R_{01}}(T_0 - T_1)\right) \tag{2.7}$$

Extending this principle across a network of interconnected nodes, each representing a thermal component (like an air zone or wall layer) with associated capacitance and linked by thermal resistances, yields a system of first-order differential equations. This resulting state-space model captures the essential thermal dynamics based on physical principles, providing a robust foundation for developing physics-based HVAC control strategies.

[1]Note that this work only considers symmetric thermal circuits, so $Q_{12} = -Q_{21}$ and $R_{12} = R_{21}$

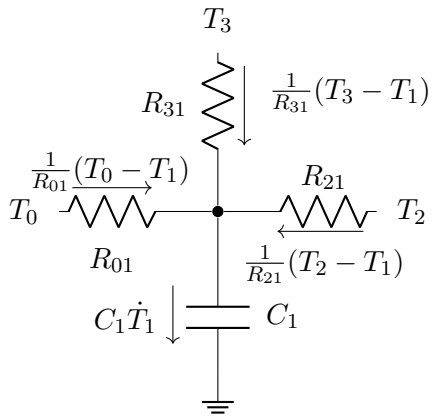

Fig. 2.2: Simple circuit illustrating Kirchhoff's law. Arrows indicate power flow through the circuit.

2.4.2 Black-Box Modeling

Black-Box Modeling is a data-driven approach that does not rely on prior knowledge of the system's internal mechanisms. Instead, it treats the system as a "black box," modeling its behavior solely on the basis of input-output data. Typically, these modeling methods employ machine learning algorithms to capture complex system dynamics. For example, [39] presents a model predictive controller that leverages the Auto-Regressive Integrated Moving Average with Exogenous Inputs (ARIMAX) model to capture a linear approximation of the dynamics of an HVAC system. This work utilizes building examples from the Modelica Buildings library to fit a linear model, enabling effective control of the system through model predictive control. Additionally, [40] explores three separate black-box methods for modeling a building's thermal response: Gaussian Mixture Regression Models, Gaussian Process Regression Models, and Deep Neural Networks (DNNs). By applying these models, the study successfully models the energy use of buildings, demonstrating the efficacy of black-box approaches in capturing complex thermal behaviors.

Black-box methods offer significant advantages when dealing with highly complex, non-linear systems that are difficult to model from first principles. They can adapt to patterns in the data without explicit physical modeling, making them flexible tools for system identification. However, these methods come with certain downsides. They are highly dependent on the quantity and quality of available data; insufficient or poor quality data can lead

to inaccurate models. In addition, complex machine learning algorithms may require substantial computational resources and time to train, which can be a limitation in practical applications.

2.4.3 Grey-box modeling

Grey-box modeling is a combination of the white-box and black-box methods. In gray-box approaches, the model structure is determined on the basis of physical principles, but the individual parameters are fitted to the data to model the system's response. In thermal systems, this typically involves using the lumped RC model discussed in the white-box modeling section. However, instead of determining the values of the parameters through building materials, regression techniques are used to fit the model parameters to the observed data, given the structure of the predetermined RC model.

A least squares parameter estimation method to estimate the parameters of such a system is presented in [2]. In this work, the model is fixed to the 3R2C model presented in [41]. The regression is performed on data collected from a real building and the study details the challenges and solutions associated with selecting appropriate parameters for this model. Furthermore, [42] introduces a 2R2C model to fit individual rooms. Instead of using a least squares method to estimate the building's parameters, this work employs a genetic algorithm. The use of a genetic algorithm allows for a more global search of the parameter space, which can be advantageous in complex models where the parameter estimation problem may be non-linear or have multiple local minima.

2.5 Grid Utilization and Demand Response

When considering interactions with the power grid, one might assume that reducing overall power consumption is an optimal loading strategy. However, this approach overlooks the operational and economic challenges that power companies face, particularly because minimizing power consumption at the end of the planning horizon can cause problems in subsequent periods [43]. The grid infrastructure is typically designed to meet peak demand, and significant increases in electrical loads can strain equipment, increase maintenance costs,

and require expensive investments in additional capacity. In contrast, minimizing power consumption during off-peak periods can lead to under-utilization of the network, resulting in inefficiencies both operationally and economically [44]. Consequently, power companies prioritize managing electrical loads to prevent over- and under-loading to maintain system reliability. Managing this regularized power profile is a strategy known as grid utilization [45]. With grid utilization, power companies can achieve a more balanced load profile, reduce the need for peaking generators, and enable the grid to operate closer to its optimal capacity, ultimately benefiting both the utility and its customers.

In addition, these grid management objectives also offer substantial benefits to consumers. For large consumers, peak power constitutes the most expensive component in power bills, so lowering peak loads directly translates into significant cost savings [46]. In addition, ensuring robust power usage during off-peak periods allows consumers to maximize their energy value by fully utilizing the available grid capacity. In this way, demand response strategies and balanced load management not only enhance grid stability and reliability, but also give consumers the most bang for their buck by aligning operational efficiency with economic benefits.

To improve grid utilization, modern power grids have also adopted demand response strategies to address fluctuations in power supply and demand. The response to demand takes advantage of the inherent flexibility in electricity consumption to provide essential grid services, improving the stability of the grid by integrating renewable energy sources [47]. Although often associated with the adoption of renewable energy, response to demand serves a broader purpose, improving the resilience of the grid and optimizing energy distribution in response to real-time conditions.

Demand response is not a new concept and has been implemented in various forms over the years. One of the oldest and most common implementations is Time-of-Use (TOU) pricing, where higher costs are applied during peak usage periods [47]. TOU programs typically target larger customers, as they require the installation of specialized metering infrastructure [48]. Beyond TOU pricing, some utilities have adopted other demand response

programs, such as critical peak pricing (CPP) [47], an event-based tariff that imposes higher prices during utility-defined peak demand events. This pricing structure has been implemented by California's Independent System Operator (ISO) to reduce loads during summer afternoons [49]. In addition, demand response can be facilitated through dynamic pricing structures that operate on shorter time scales. The continuous response to demand, introduced as early as 1988 by [50], was termed Spot Pricing of electricity. Spot pricing involves varying price structures based on real-time demand and supply conditions, providing flexible load management, and improving grid stability.

Demand response programs can vary widely in their notification times, which generally fall into three categories [51]. First, ultra-short notification programs, measured in seconds, are designed to respond to emergency fluctuations in the grid [52]. Second, moderate notification programs provide advance notice ranging from thirty minutes to several hours, balancing rapid response with practical operational adjustments [53]. Finally, long-notification programs issue signals days in advance, allowing planning around major preplanned events and granting the grid greater operational flexibility [51]. For the purposes of this work, it is assumed that these instructions are typically communicated to the microgrid within a moderate timeframe.

The proposed framework employs a pair of demand response signals as "soft bounds." The power company issues these signals to the microgrid operator, who retains the flexibility to adjust consumption accordingly. The first signal encourages increased power consumption during periods of low demand, whereas the second signal promotes reduced power consumption during periods of high demand.

Soft-bound implementations are found in the literature. A study introduced a soft limit at the household level for EV charging by setting a critical power threshold that, if exceeded, incurs a penalty fee and, if not, offers a reward [54]. This mechanism effectively encourages residential users to manage their peak loads in a manner consistent with grid requirements without imposing strict mandates. In another study, [55], a Demand Turn-Up strategy was proposed in which the electricity price is set to zero during periods of renewable

energy surplus, causing consumers to increase their consumption, such as charging thermal storage or heating water, and therefore influencing all related constraints. Building upon these local optimization goals, demand response mechanisms provide an additional layer of control to enhance greater grid stability, particularly during dynamic conditions.

In summary, this review highlights significant progress in microgrid forecasting, optimization, and DER modeling. However, critical gaps remain, particularly in validated frameworks that integrate advanced forecasting with multiobjective hierarchical control that coordinates diverse electrical DERs alongside complex thermal loads, such as HVAC systems. Existing optimization strategies often prioritize economic objectives over operational grid stability metrics or require complex post-processing, while HVAC modeling approaches face trade-offs between accuracy, data requirements, and suitability for control integration. This analysis establishes the clear need for the integrated hierarchical RHC framework and the dual HVAC control methodologies developed and evaluated in the subsequent chapters of this thesis.

2.6 Research Motivation

The preceding review of the literature highlights both significant advances and persistent challenges in the management of DERs within microgrids. While sophisticated techniques exist for forecasting, optimization, and modeling individual components, a key gap remains in validated, integrated frameworks capable of coordinating diverse electrical resources such as ESS and EV chargers alongside complex thermal loads, particularly Heating, Ventilation, and Air Conditioning (HVAC) systems, under dynamic grid conditions. Existing optimization approaches often struggle to effectively balance multiple operational priorities or require complex post-processing, while accurately modeling and controlling HVAC systems presents substantial hurdles related to data requirements, model fidelity, and practical implementation.

This thesis is motivated by the need for more holistic and operationally focused microgrid management solutions, particularly for large-scale multifamily developments like the Citizens West project in Salt Lake City, which features a rich mix of controllable resources

(hot water, HVAC, ESS, PV, EV chargers). Without a unified system, optimizing these assets individually risks increased costs and suboptimal grid interaction. Therefore, this research was undertaken to develop and explore key components of such a comprehensive system.

Specifically, this work focused on two main thrusts. The first was the development and validation of a hierarchical RHC framework designed to coordinate electrical DERs (ESS, EVs, considering PV) by explicitly prioritizing operational goals such as peak shaving and load smoothing, while incorporating dynamic demand response signals. The second thrust involved investigating advanced control methodologies for HVAC systems, recognizing their significant energy impact and potential flexibility. Two distinct approaches were developed: a physics-based method using lumped-sum RC models amenable to optimal control and a data-driven, regression-based method focused on controlling standard temperature setpoints.

However, while the RHC framework for electrical DERs was implemented and tested via simulation, and the HVAC methodologies were developed and evaluated independently, the scope of this thesis did not extend to the full integration and co-optimization of these HVAC controllers within the validated RHC loop. Thus, the research presented provides a robust framework for electrical DER coordination (detailed in Chapter 3) and foundational advancements in HVAC modeling and control (detailed in Chapters 4 and 5) as essential steps towards a truly unified system, leaving the final integration as a critical area for future work. The overarching goal remains to provide solutions that enable buildings such as Citizens West to operate more efficiently, benefiting both owners and the larger electrical grid.

CHAPTER 3

Hierarchical Battery Control

3.1 Introduction

The increasing penetration of renewable energy sources presents both immense opportunities and significant challenges for modern power grids, requiring more sophisticated control strategies to manage variability and ensure stability. Microgrids, as localized electrical systems that integrate various distributed energy resources (DER), such as solar PV, wind turbines, energy storage (ESS), and electric vehicle (EV) chargers, offer a promising solution. These systems improve local resilience and enable customized energy management, acting as flexible units that can dynamically interact with the larger grid. However, effectively coordinating the complex interplay between numerous DERs, fluctuating renewables, variable loads, and grid interactions to meet multiple, often conflicting, objectives remains a critical research challenge.

Hierarchical control structures have been explored to manage such complexity. For example, some frameworks utilize Markov Decision Processes (MDPs) to handle uncertainty and optimize across different energy vectors (electricity, heat, gas), focusing on learning optimal policies for aggregated consumers, as seen in [21]. Although powerful for capturing stochastic transitions, MDP-based approaches can require significant effort in modeling state spaces and reward functions, and the resulting policies might not always align directly with explicit real-time operational priorities.

Another common approach involves multi-objective optimization techniques, often employing evolutionary algorithms such as MOGA to generate a Pareto front that represents trade-offs between conflicting goals, such as minimizing the cost of generating the power from the grid and the cost of battery degradation [26]. These methods excel at exploring the solution space and identifying optimal compromises between the chosen objectives, typically

economic ones. However, selecting a single operating point from the resulting Pareto front often requires a subsequent, potentially complex, or subjective, decision-making process (for example, fuzzy logic). Furthermore, the computational effort to generate a full Pareto front can be substantial, potentially limiting its application in control loops that require frequent updates.

This chapter proposes a different paradigm: a sequential and hierarchical optimization framework designed to directly embed operational priorities crucial for grid stability and resilience. This work contends that, for many operational scenarios, certain objectives take precedence. The hierarchy of this work prioritizes minimizing peak power consumption (critical for reducing grid strain and demand charges), followed by minimizing the power range (to smooth the load profile presented to the grid), then maximizing the state-of-charge (SOC) of essential building batteries (for local resilience), and finally maximizing the overall system SOC (for flexibility). This specific order reflects common utility and operator concerns, aiming for a stable and predictable grid interaction first, before optimizing storage levels. While this sequential approach deliberately forgoes exploring the full Pareto trade-off space inherent in multi-objective methods, it yields a single, unambiguous operating schedule directly aligned with these pre-defined operational priorities, simplifying decision-making in dynamic contexts. This work acknowledges the solution's sensitivity to this chosen order, which reflects a specific operational philosophy, but could be adapted for different requirements.

The works [21] and [26] serve as illustrative examples that represent different philosophies to tackle hierarchical control challenges in microgrids. The former [21] delves into stochastic modeling, focusing on learning optimal long-term policies under uncertainty. The latter [26] employs multi-objective evolutionary algorithms to explicitly map the Pareto frontier, effectively exploring the trade-offs primarily between economic objectives, though requiring a subsequent decision step to select an operating point from the generated set.

Between these specific approaches lies a broad spectrum of research that focuses on microgrid control and optimization. This includes various implementations of Model Pre-

dictive Control (MPC) or Receding Horizon control (RHC) often using linear or quadratic programming for faster computation [11, 13], day-ahead weighted-sum optimization techniques that aggregate multiple objectives into a single function [17], rule-based hierarchical systems offering simplicity but potentially sacrificing optimality [25], and fuzzy logic controllers for handling imprecise objectives or complex decision-making [24].

The novelty of this work's contribution lies not in the concept of sequential optimization itself but in its specific, integrated application within a comprehensive microgrid management system. This work uniquely combines: (1) this particular operational hierarchy tailored for grid stability; (2) advanced load forecasting using Extreme Gradient Boosting (XGBoost); (3) a receding horizon control (RHC) structure for adaptability; and (4) a distinct formulation of demand response signals acting as dynamic "soft bounds" integrated directly into the hierarchical optimization constraints. This integrated system offers a potentially robust and computationally efficient alternative that is aimed squarely at achieving prioritized operational performance in microgrids with significant penetration of DER.

The remainder of this chapter is organized as follows. First, Table 3.1 provides the problem terminology. Then, Section 3.2 details the hierarchical planning structure, including the overall control loop, dynamic constraints, and describes the objectives for the local utilization of the grid and the implementation of demand response strategies. Finally, Sections 3.3 and 3.4 present the experimental setup, simulation results, and concluding remarks.

3.2 Hierarchical Planning

This section details the proposed hierarchical control framework, outlining the overall control loop, variable definitions, and the dynamic constraints governing the system. Then it introduces two types of hierarchical objectives: local grid utilization and demand response.

3.2.1 Overarching Control Loop

The focus of this work is to develop a control loop that can achieve a series of hierarchical objectives that can respond to dynamic signals from the power company and changes

Table 3.1: Problem Terminology

Problem Terminology	
N_t	Number of time steps
N_{dt}	Number of time steps for demand response
N_b	Number of batteries
δ	Time Step (Hr)
k	A given time step to denote a time window
Battery Terminology	
S_k	SOC at time step (kWh) $\in \mathbb{R}^{N_b}$
S_{min}	The minimum allowed SOC for a battery (kWh) $\in \mathbb{R}^{N_b}$
S_{max}	The maximum allowed SOC for a battery (kWh) $\in \mathbb{R}^{N_b}$
$U_{b,k}$	Applied power at time step k (kW) for battery b $\in \mathbb{R}^{N_b}$
$U_{b,\max}$	Maximum allowable discharge from battery b (kW) $\in \mathbb{R}^{N_b}$
$U_{\min}$	Maximum allowable charge to the grid from battery b(kW) $\in \mathbb{R}^{N_b}$
$P_{b,k}$	Power utilization for battery at time step k (kW) $\in \mathbb{R}^{N_b}$
Power Terminology	
$P_{n,k}$	Forecasted net external load/generation at time step k (kW) $\in \mathbb{R}$
P_k	Total power utilization for the Microgrid in step k (kW) $\in \mathbb{R}$
k_c	Efficiency factor for the charging/discharging
p_{min}	The minimum power for utilized (kW) $\forall k \in \mathbb{R}$
p_{max}	The maximum power for utilized (kW) $\forall k \in \mathbb{R}$
O_{pos}	Incentivized Demand response vector $(kW) \in \mathbb{R}^{N_t}$
O_{neg}	Demand minimization vector $(kW) \in \mathbb{R}^{N_t}$
P_{eff}	Augmented power vector (kW) $P + O \in \mathbb{R}^{N_t}$
P_o	Normalized power vector for demand response $(kW) \in \mathbb{R}^{N_t}$
$\mathbb{D}$	Binary indicator vector
P_{const}	Preprocessed power vector (kW) $\mathbb{D}$ for demand minimization

in the system. To address this, a receding horizon control (RHC) loop is proposed. RHC can be implemented to account for the inherent uncertainty and variability in energy generation and demand by ingestion of feedback from the system. This method introduces a dynamic iterative approach in which the optimization problem is solved at regular intervals, typically every few minutes to every few hours, using updated forecasts that reflect the most recent data. Receding-horizon control adds robustness by continuously adjusting the strategy based on real-time measurements and updated predictions. This allows the system to respond to stochastic elements such as weather variations, unexpected demand spikes, or renewable energy fluctuations.

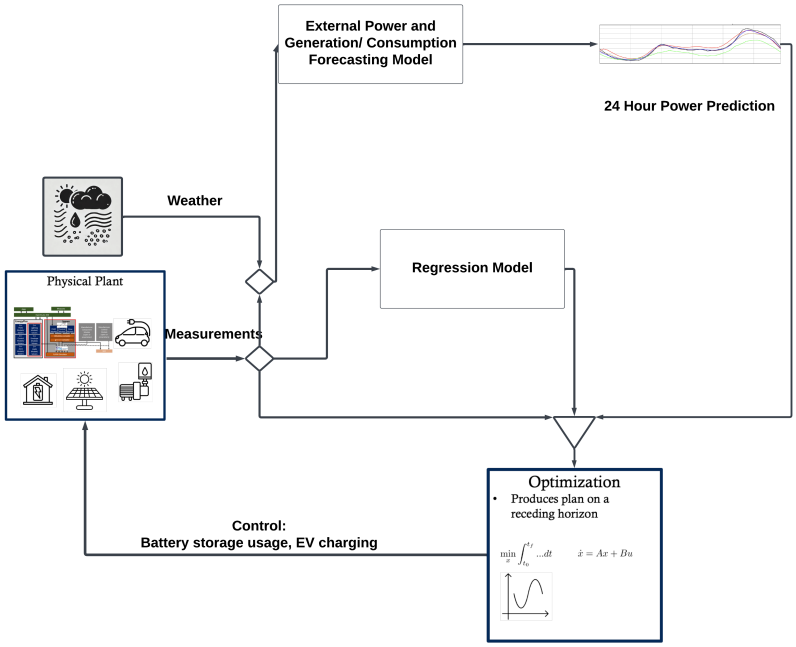

Fig. 3.1: Receding Horizon Control Loop Diagram

Figure 3.1 provides a visual representation of the control loop logic used in the receding horizon planner, detailing the interactions between the measurement, modeling, and optimization components.

3.2.2 Variable and Solution Vector Definition

Key parameters and constants must be defined to develop the mathematical formulation. These key optimization parameters include applied battery power, U, SOC, S, and power usage, P. Each of these parameters are a matrix or vector that associates every characteristic of the problem with a discrete time step. This problem is discretized over a 24-hour horizon. Each time step, $\delta \in \mathbb{R}$, corresponds to a time interval. The time step is equal to the total time horizon divided by the number of time steps.

The parameter U is a matrix of the applied power to each battery for all time steps N_t, and for all batteries N_b, it has dimensions of $N_b \times N_t$. When U is indexed by k, it is a vector that corresponds to the total power applied to each battery at a single time step with dimensions N_b. The parameter S is a matrix for the SOC for all batteries and all time

steps plus an additional time to allow for a final SOC. Similarly to U, when S is indexed by k, it is a vector that corresponds to the SOC of each battery for a single time step with dimensions $\mathbb{R}^{N_b}$. S has the dimensions of $N_b \times N_t + 1$. The final major parameter, P, is a vector of the total power use for each time step and has the dimensions of N_t.

Two additional variables allow the definition of maximum and minimum power usage for the entire time horizon and are denoted as p_{max} and p_{min}, respectively. These parameters are arranged within a solution vector, represented as y. Since y is the solution vector, the parameter matrices (U, S) must be vectorized (reshaped into a one-dimensional form). This vectorization, denoted using the operator $vec()$, stacks the elements of a matrix (e.g., column-by-column) into a single column vector, ensuring they are sequentially stored within y.

$$y = \begin{bmatrix} \text{vec}(U)^T & \text{vec}(S)^T & P^T & p_{min} & p_{max} \end{bmatrix}^T \tag{3.1}$$

3.2.3 Dynamic Constraints

The following dynamic constraints formalize the evolution of the SOC and the total power consumption.

State of Charge Propagation: For each battery b, the SOC is updated according to its power input:

$$S_{b,k+1} = S_{b,k} + \delta \, k_c \, U_{b,k}, \quad k = 0, \cdots, N_t + 1. \tag{3.2}$$

This equation relates the SOC at time k to the SOC at time $k + 1$, where the change is determined by the control input $U_{b,k}$, the discretization of time δ, and the charge efficiency factor k_c.

Total Power Consumption: The total power consumption at time k is defined by subtracting the aggregated battery inputs from the overall power consumption, thus accounting for the forecast external load. This relationship holds that the external power signal $P_{n,k}$ is equal to the total power consumption P_k and the sum of the total power applied to all batteries considered $U_{b,k}$ for each time step k.

$$P_{n,k} = P_k - \sum_{b=1}^{N_b} U_{b,k}, \quad k = 0, \cdots, N_t + 1. \tag{3.3}$$

Equation (3.3) shows this equality relationship that must exist for each time step.

Bound Constraints In addition to the propagation constraints, upper and lower bounds are imposed on both the SOC and the battery power input:

$$S_{min} \leq S_{b,k} \leq S_{max} \quad \forall k \tag{3.4}$$

$$U_{min} \leq U_{b,k} \leq U_{max} \quad \forall k \tag{3.5}$$

These dynamic constraints establish the operational framework for the microgrid and form the basis for the subsequent optimization of the utilization of the local grid.

3.2.4 Local Grid Utilization

The local utilization of the grid improves the performance of the microgrid by driving the stability and efficiency of the key outcomes. In terms of stability, the approach is designed to smooth out power consumption by minimizing peaks and increasing troughs. This creates a more regular and predictable energy profile that reduces stress on the system. Efficiency is achieved through reduced operational costs; By intelligently managing energy flow, the microgrid minimizes waste and the need for costly peak power operations, ensuring that energy is utilized optimally. The local grid utilization objective consists of four sub-objectives, each ordered in descending importance: minimizing peak power, minimizing power range, maximizing the SOC of the building battery, and maximizing the overall battery SOC.

Minimizing peak power is prioritized because it reduces stress on the grid infrastructure, prevents overload conditions, and lowers demand charges, thus improving the overall reliability of the grid. Minimizing power range smooths fluctuations in energy consumption, aiding in grid stability by providing a more predictable power profile to the larger grid. As stated in Maximizing the SOC of the building battery ensures that local storage is used

effectively, providing resilience against power outages and allowing for improved load shifting to reduce reliance on grid power. Finally, maximizing the overall SOC of the battery supports broader energy availability, allowing greater flexibility in responding to demand fluctuations and integrating renewable energy sources more effectively.

To address all of these objectives, a process known as lexicographic optimization is leveraged. The process of creating a lexicographic method is to order a set of objective functions in importance order. Each objective in order is then solved with the following algorithm. This lexigraphic method is outlined in Equation 3.6.

$$\begin{aligned} &\min F_i(x) \\ &\text{given that } F_j(x) \leq F_j(x_j^*),\ j = 1, 2, \ldots, i-1;\ i > 1 \\ &i = 1, 2, \ldots, m, \end{aligned} \tag{3.6}$$

Here, i represents the order of the optimization functions, and $F_j(x)$ is the optimal solution of the objective function j^{th} of the iteration j^{th}. Each subsequent optimization problem is allowed to find new solutions. But, each iteration adds a new constraint that any new solutions cannot degrade previous objective functions.

Minimize Peak Power

The primary objective is to minimize peak power consumption to reduce grid strain during peak periods. Let $\mathbb{1}$ denote a column vector where each element is one, with dimensions appropriate for element-wise comparison with the vector it operates on (in this case, $N_t \times 1$ to align with the power vector P). The scalar slack variable p_{max} is introduced to represent the maximum power value on the planning horizon.

The variable p_{max} is constrained such that it must be greater than or equal to the power consumed in every time step:

$$\mathbb{1}\, p_{max} \geq P \tag{3.7}$$

Equation (3.7) uses the vector $\mathbb{1}$ to ensure the single value p_{max} forms an upper bound for all elements within the power vector P. This slack variable p_{max} is then directly minimized in the first objective function, subject to the system's dynamic and bound constraints.

$$
\begin{aligned}
\min_{y} \quad & p_{max} \\
\text{s.t.} \quad & \mathbb{1}\, p_{max} \geq P, \\
& S_{b,k+1} = S_{b,k} + \delta\, k_c\, U_{b,k}, \quad k = 0, \cdots, N_t + 1, b = 0, \cdots, N_b \\
& P_{n,k} = P_k - \sum_{b=1}^{N_b} U_{b,k}, \quad k = 0, \cdots, N_t + 1, \\
& U_{b,\min} \leq U_{b,k} \leq U_{b,\max}, \quad \forall b, \forall k \\
& S_{b,min} \leq S_{b,k} \leq S_{b,max}, \quad \forall b, \forall k
\end{aligned}
\tag{3.8}
$$

For the sake of brevity, the constraints in (3.8) are represented using three vertical dots in the following problems, Equations (3.9) - (3.12).

Minimize Power Range

A secondary objective is to minimize the difference between maximum and minimum power usage over the defined time horizon. This encourages efficient utilization of the available power resources. This can be done with a structure similar to that of the previous problem. The slack variable p_{min} can be introduced and constrained element-wise to be less than or equal to the power vector. To pass the previous solutions maximum to this problem the variable $p_{max,prev}$ is set as an equality constraint to this problem's p_{max}

$$\begin{aligned} \max_{y} \quad & p_{min} \\ \text{s.t.} \quad & \mathbb{1} p_{max} \geq P \\ & p_{max} = p_{max,prev} \\ & \mathbb{1} p_{min} \leq P \\ & \vdots \end{aligned} \tag{3.9}$$

Maximize Building Battery SOC

A tertiary objective focuses on maximizing the final SOC of the building's storage batteries. This ensures that building energy storage is at the highest SOC for unplanned events beyond the planning horizon. To ensure that the current solution meets or exceeds the previous minimum, $p_{min,\,\mathrm{prev}}$ is enforced as an equality constraint in p_{min}.

$$\begin{aligned} \max_{y} \quad & S_{1,N_t+1} \\ \text{s.t.} \quad & \mathbb{1} p_{max} \geq P \\ & p_{max} = p_{max,prev} \\ & \mathbb{1} p_{min} \leq P \\ & p_{min} = p_{min,prev} \\ & \vdots \end{aligned} \tag{3.10}$$

Maximize Overall Battery SOC

Finally, the overarching objective is to maximize the SOC for every battery at each time interval, ensuring that both ESS and EV chargers maintain high charge levels to enhance system robustness against unexpected disturbances. This objective is achieved by aggregating the SOC values across all batteries during the planning period. Let S be the matrix of SOC values, where each column corresponds to an individual battery. The overall SOC is obtained by first summing the SOC over time for each battery and then summing

these values across all batteries. This operation is mathematically represented as

$$\sum_{b=1}^{N_b}(\sum_{k=1}^{N_t+1} S_{b,k}) \tag{3.11}$$

To ensure continuity with the previous solution, the final SOC of the building storage battery, $S_{1,\,N+1,\text{prev}}$, is imposed as an equality constraint on the final SOC of the current, $S_{1,\,N+1}$.

$$\begin{aligned}
\max_{y} \quad & \sum_{b=1}^{N_b}(\sum_{k=1}^{N_t+1} S_{b,k}) \\
\text{s.t. } & \mathbb{1}p_{max} \geq P \\
& p_{max} = p_{max,\,prev} \\
& \mathbb{1}p_{min} \leq P \\
& p_{min} = p_{min,\,prev} \\
& S_{1,N_t} = S_{1,\,N_t+1,\,prev} \\
& \vdots
\end{aligned} \tag{3.12}$$

3.2.5 Demand Response

Although local grid utilization strategies focus on optimizing energy usage during regular operating conditions through objectives such as peak power reduction and efficient storage management, they do not fully address grid behavior during dynamic or emergency conditions. To further enhance grid stability and responsiveness, demand response mechanisms are integrated into the framework. These mechanisms enable the microgrid to adapt its consumption patterns in real time, either by incentivizing increased usage when the larger grid is underutilized or by enforcing consumption curtailment during periods of high demand. The following section details these demand response strategies and explains how they complement the local grid utilization objectives.

To optimize the overall utility of the external power grid, the microgrid accepts signals

that indicate specific periods during which it encourages the microgrid to draw power from or supply power to the grid. As discussed in 2.5, there can be three types of communication periods for demand response programs. For this work, it is assumed that these instructions are typically communicated to the microgrid in a moderate time frame.

To effectively support the overarching objectives of demand response, two distinct demand response signals are introduced: one aimed at incentivizing increased power consumption and another designed to encourage power curtailment. The first signal, termed an incentivized demand response, denotes periods when the utility encourages increased power usage above a specified threshold to stabilize the grid during periods of low demand. The second signal, identified as demand minimization, specifies intervals in which power consumption must be constrained to a predetermined maximum to mitigate grid strain or avoid peak demand charges. Two approaches are proposed for managing the demand minimization signal: the first ideal scenario fully satisfies the consumption limits defined by the utility, while the second scenario provides a fallback solution used when operational constraints make the primary response to demand infeasible. Both signals play an integral role in shaping the operational strategy of the microgrid and are embedded in the optimization framework detailed in subsequent sections.

Incentivized Demand Response

To account for periods when the power company wants to increase utilization, the concept of incentivized demand response is introduced. This signal is represented by the vector $O_{pos} \in \mathbb{R}^{N_t}$, where each element, indexed by time k, indicates the level of "free" power (measured in kilowatts) designated by the power company at that time. Here, "free" power means that any power consumption below the specified level is not charged on the microgrid's power bill. If no signal is provided at a particular time, the corresponding element is set to 0.

$$O_{pos} \in \mathbb{R}^{N_t} \quad \begin{cases} O_{pos,k} > 0 & \text{if incentivize d demand exists} \\ 0 & \text{o.w.} \end{cases}$$

With the incentivized demand response defined, the signal O_{pos} must be integrated with the power consumption data to form new objective functions. To achieve this, a new vector P_{eff} is introduced that represents the amount of energy that the utility company will incur costs. In other words, P_{eff} accounts for the effective power consumption after considering the "free" power provided by O_{pos}. To compute P_{eff}, a preprocessing step rescales the original power vector, P, by subtracting O_{pos}. This operation yields the effective power cost seen by the utility:

$$P_{\text{eff}} = P - O_{pos} \quad \in \mathbb{R}^{N_t}. \tag{3.13}$$

With this effective power vector, the slack variable p_{max} can be constrained to be greater than or equal to P_{eff} rather than P.

$$p_{max}\mathbb{1} \geq P_{\text{eff}} \tag{3.14}$$

With this constraint, the objective function for this demand response can then be defined as the following

$$\begin{aligned} \min_{y} \quad & p_{max} \\ \text{s.t.} \quad & P_{\text{eff}} = P - O_{pos} \\ & p_{max}\mathbb{1} \geq P_{\text{eff}} \\ & \vdots \end{aligned}. \tag{3.15}$$

Demand Minimization

To define demand minimization, the signal that represents this process must be defined:

$$O_{neg} \in \mathbb{R}^{N_t} \quad \begin{cases} O_{neg,k} > 0, & \text{if a demand response limit exists,} \\ 0, & \text{otherwise.} \end{cases} . \tag{3.16}$$

Here, O_{neg} is analogous to O_{pos}; at each index k, it specifies the power level below which the microgrid is required to operate during a demand response event. Two approaches are defined for demand minimization: the primary approach assumes that the microgrid has sufficient energy reserves to handle the utility's signal, while the backup approach applies when the microgrid lacks the necessary reserves, but still aims to reduce energy consumption as much as possible.

Both of these methods require preprocessing of this O_{neg} vector. To do this preprocessing, the binary indicator vector $\mathbb{D} \in \mathbb{R}^{N_t}$ is introduced

$$\mathbb{D} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \quad \text{where } \mathbb{D}_k = \begin{cases} 1, & \text{if } O_{neg,k} > 0, \\ 0, & \text{otherwise.} \end{cases} . \tag{3.17}$$

This vector allows us to extract the power values at the times when demand minimization is required.

Constrained Demand Minimization: Assuming that the local grid can handle the curtailment signal defined by the utility, power usage is constrained during periods when the

demand response limit is active. A diagonal matrix is formed from $\mathbb{D}$:

$$\mathbb{D}_{mat} = \text{diag}(\mathbb{D}) \tag{3.18}$$

This matrix $\mathbb{D}_{mat}$ can be used to filter the power vector P to create the vector P_{const}. This new power vector will contain only power values during the relevant times for demand minimization."

$$P_{const} = \mathbb{D}_{mat} \cdot P \tag{3.19}$$

Subsequently, the following element-wise constraint is imposed:

$$P_{const} \leq O_{neg}, \tag{3.20}$$

ensuring that power consumption during these periods does not exceed the specified demand limit. This constraint is then incorporated into the set of dominating constraints defined in Section 3.2.4. To illustrate, the first level of these constraints is redefined as follows:

$$\begin{aligned}
\min_{y} \quad & p_{max} \\
\text{s.t.} \quad & \mathbb{1}\, p_{max} \geq P_{const}, \\
& S_{k+1} = S_k + \delta\, k_c\, U_{b,k}, \quad k = 0, \cdots, N_t + 1, \\
& P_{n,k} = P_k - \sum_{k=1}^{N_b} k_c\, U_{b,k}, \quad k = 0, \cdots, N_t + 1, \\
& P_{const} = \mathbb{D}_{mat} \cdot P, \\
& P_{const} \leq O_{neg}, \\
& U_{b,\min} \leq U_{b,0} \leq U_{b,\max}, \\
& U_{b,\min} \leq U_{b,N_t} \leq U_{b,\max}.
\end{aligned} \tag{3.21}$$

Backup Demand Minimization: The backup demand minimization method also directly targets the specific periods during which power consumption must be reduced to meet

the designated threshold. When the constrained demand response becomes infeasible, this backup minimization will be executed. This means that everything defined in Section 3.2.5 is ignored, and instead this second structure is executed.

This backup process performs the optimization differently. Instead of redefining the top-level objective function as in Section 3.2.5, it introduces a new objective function that will take the highest priority in the hierarchical structure. This means that the new optimization problem takes precedence over all previously defined problems (see Section 3.2.4). In effect, the solution to this optimization will serve as an upper bound on power consumption during demand response events, thereby influencing all related constraints.

This approach, similar to constrained demand minimization, first identifies all time indices where a demand response limit is required with the vector $\mathbb{D}$. For each identified period, excess power consumption is calculated over the set limit, and this surplus becomes the key metric for analysis. This normalization is achieved by constructing a vector, P_o, defined as the original power consumption with the offset O_{neg} subtracted, thereby quantifying the normalized excess power.

The adjusted power consumption vector, P_o, is defined as:

$$P_o = P - O_{neg} \quad \in \mathbb{R}^{N_t}. \tag{3.22}$$

To optimize power usage during demand minimization, the variable $p_{\max,\mathrm{d}}$, which captures the maximum deviation between the planned power usage and the limit of curtailment, is introduced. It is constrained to be at least as large as the power consumption P_o at every time step k when a demand response limit is active:

$$\mathbb{1}p_{\max,\mathrm{d}} \geq P_o \quad \forall k \text{ with a demand response limit} \tag{3.23}$$

To enforce this restriction only during the response periods to demand, the defined binary indicator vector $\mathbb{D} \in \mathbb{R}^{N_t}$ can be used. Using this indicator vector, the constraint on

peak power during demand response times is expressed as:

$$\mathbb{D} \odot p_{\max,\,\mathrm{d}} \geq P_o. \tag{3.24}$$

$$\begin{aligned}
&\min_{y} \quad p_{max,\,d} \\
&\text{s.t. } S_{k+1} = S_k + \delta\, k_c\, U_{b,k}, \quad k = 0, \cdots, N_t + 1, \\
&\qquad P_{n,k} = P_k - \sum_{k=1}^{N_b} k_c\, U_{b,k}, \quad k = 0, \cdots, N_t + 1, \\
&\qquad P_o = P - O_{neg} \\
&\qquad \mathbb{D} \odot p_{max,\,d} \geq P_o \\
&\qquad P_{const} = \mathbb{D}_{mat}\, P, \\
&\qquad P_{const} \leq O_{neg}, \\
&\qquad U_{b,\min} \leq U_{b,0} \leq U_{b,\max}, \\
&\qquad U_{b,\min} \leq U_{b,N_t} \leq U_{b,\max}
\end{aligned} \tag{3.25}$$

3.3 Experimental Setup

To comprehensively present the results, this section details the experimental setup that underpins this study. First, the data sets used to generate realistic power consumption are described, ensuring that the experiments are grounded in real-world scenarios. These datasets provide the essential data specifically power usage and power consumption.

The experimental framework is organized into two primary segments: day-ahead testing and receding-horizon testing. In the planned day-ahead scenario, planning is performed independently of forecasting, which isolates the impact of pre-scheduled strategies. In contrast, the receding-horizon approach integrates real-time forecasting, thereby enabling the control strategy to adapt dynamically to evolving conditions. This dual framework not only highlights the benefits of proactive planning, but also demonstrates the system's ability to

respond to uncertainties in a more agile manner.

3.3.1 Datasets

Four data sets were used to develop different components of the experimental setup. These data sets are carefully curated to generate realistic power consumption and generation profiles, providing a solid basis for evaluating our energy management strategies.

The first data set used in this study is [56]. This data set contains data on hourly power consumption (in kWh) from a two-story house in Houston, Texas, that spans from June 1, 2016 to August 2020. The next data set used is [57]. Data were collected over 34 days from two solar power plants in India. The third data set used is [58]. Originally from the UK's smart meter rollout initiative, this dataset provides a refactored sample of electrical consumption data for 5,567 London households, collected between November 2011 and February 2014. The final data set was collected from the OpenMeteo API [59]. This api was used to augment the weather data from [58] with the solar irradiance data.

3.3.2 Day-ahead Testing

To develop and tune the objective functions, the day-ahead planning approach uses historical average forecasting to create a static external profile. This profile is constructed by combining two signals: an averaged power consumption profile and an averaged power generation profile.

Power Consumption Profile

The power consumption profile is derived from historical data from single-family homes during the summer months and is scaled to reflect the peak loads typically observed in multi-family residences, as obtained from the dataset [56].

Figure 3.2 illustrates the static average power consumption profile used in the day-ahead tests.

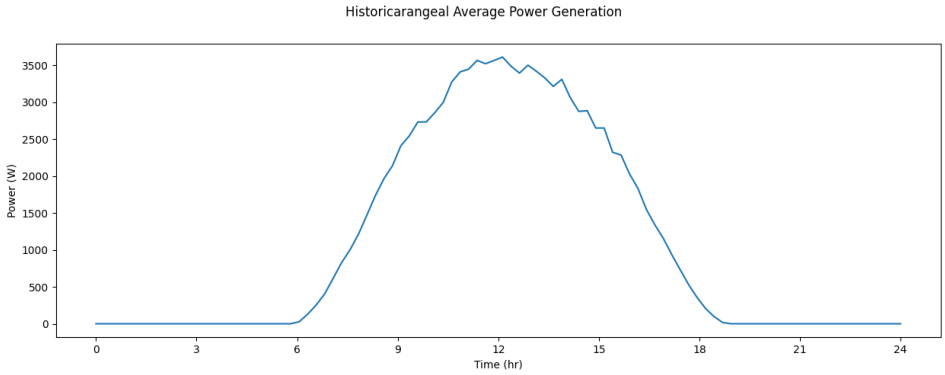

Fig. 3.2: Static Average Power Consumption Profile

Power Generation Profile

The power generation profile is based on averaged data from several solar farms in India [57], collected during the summer months under high solar irradiance conditions.

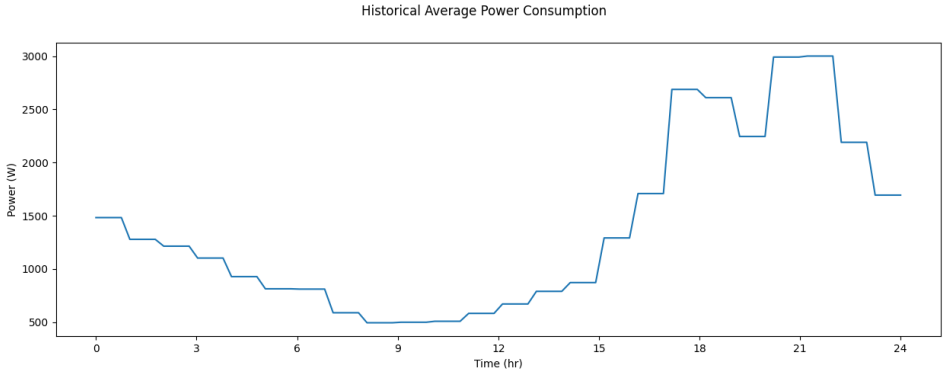

Fig. 3.3: Static Solar Generation Profile

Day-ahead Planning Specifics

As shown in Figure 3.3, the static solar generation profile is a critical element of our day-ahead tests. The net external load is computed by subtracting the solar generation profile from the power consumption profile, resulting in an external load that peaks at 301 kW. The system is equipped with a battery wall rated at 300 kWh, with both charging and discharging capped at 100 kW. In addition, demand response signals are varied to show the

flexibility and capabilities of the planning architecture. Finally, 10 EVs are connected and charged at the start of the planning horizon, and all vehicles are disconnected 30 minutes before its conclusion.

3.3.3 Receding Horizon Testing

Receding horizon testing offers a dynamic framework to evaluate the proposed energy management system under conditions that mimic real-time operation, explicitly addressing the challenge of uncertainty inherent in forecasts. The primary mechanism for handling uncertainty in this framework is the iterative nature of RHC: The system frequently re-forecasts and re-optimizes its plan based on the latest available measurements and predictions. However, it is crucial to note that the optimization problem solved within each RHC step (as defined in Sections 3.2.4 and 3.2.5) is deterministic. It relies on point forecasts for load and generation without incorporating stochastic or robust optimization techniques to explicitly hedge against uncertainty within the optimization formulation itself.

This setup necessitates a distinction between the data used for planning and the data representing the "ground truth" against which the system's performance is implicitly evaluated during simulation. The planner operates using potentially inaccurate forecasts, while the simulation environment progresses based on (simulated) actual conditions. The following subsections detail the forecasting methods that provide input to the planner and other dynamic aspects of the RHC setup.

Power Consumption Forecasting

The XGBoost model was chosen to perform to perform forecasting. XGBoost is an advanced ensemble learning algorithm that implements the gradient boosting framework [60]. In gradient boosting, a series of decision trees is iteratively built, with each successive tree trained to correct the errors of its predecessors. At every iteration, the algorithm minimizes a loss function by taking a step in the direction of the gradient, effectively reducing the prediction error. XGBoost further enhances this process by incorporating regularization techniques to prevent overfitting, optimizing both the complexity of the model and the

speed of training. Its ability to handle missing data, support parallel processing, and efficiently manage large-scale data sets makes it particularly well suited for the complex and variable nature of time-series data in energy systems.

The input to the model is constructed by sliding a window over the time series data. For each instance, a sequence of 49 half-hour steps is used as input and the target is defined as the subsequent 48 half-hour steps of average energy consumption. The feature vector for each input sequence is formed by concatenating the following variables:

- Average energy consumption recorded over a specific time interval.
- Binary indicator denoting whether the day is a public holiday.
- Measurement of atmospheric visibility (for example, in kilometers or miles).
- Ambient temperature recorded at the time of observation.
- Temperature of the dew point, indicating the level at which the air becomes saturated.
- Atmospheric pressure measurement (typically in hPa or millibars).
- Wind speed measurement indicating the rate of air movement.
- Categorical description of the type of precipitation (e.g., rain, snow, or none).
- Relative humidity, reflecting the moisture content in the air.
- Encoded representation of the time of day (hour and minute).
- The month during which the measurement was taken.
- The day of the month corresponding to the observation.
- The year in which the data was recorded.
- Binary indicator specifying whether the day is a weekday.

Table 3.2: XGBoost Model Parameters

Parameter	**Value**	**Description**
n_estimators	1000	Number of boosting rounds
learning_rate	0.01	Step size shrinkage
early_stopping_rounds	50	Rounds for early stopping

The target variable consists of the next 48 half-hour measurements of average energy, capturing the short-term dynamics in power consumption. All variables were normalized between 0 and 1 before training.

The XGBoost model is configured as summarized in Table 3.2.

Power Generation Forecasting

This forecasting model is only dedicated to predicting the power consumption of the microgrid. For power generation, solar irradiance data from [59] is utilized. This API allows for the specification of the panel tilt and azimuth; in these experiments, the tilt is set to 45% and the azimuth to 0° (South). The obtained solar irradiance value is then input into an analytical model that converts it into a generated power value. [31] presents an analytical model for predicting PV power output, defined by the equation:

$$P_{PV}(t) = \frac{G_k}{G_{\text{ref}}} \times P_{PV-\text{STC}} \times \eta_{PV} \times \left[1 - \beta_T \left(T_c - T_{C-\text{STC}}\right)\right], \tag{3.26}$$

where $P_{PV}(t)$ denotes the PV power output at time t, G_k represents the solar radiation at time t, G_{ref} is the reference solar irradiance (1000 W/m^2), $P_{PV-\text{STC}}$ is the rated power of the PV panel under Standard Test Conditions (STC), η_{PV} is the generation efficiency, β_T is the temperature coefficient, T_c is the cell temperature, and $T_{C-\text{STC}}$ is the reference cell temperature. The cell temperature, T_c, is calculated using the equation:

$$T_c = T_{\text{amb}} + (NOCT - 20) \times \frac{G_t(t)}{800}, \tag{3.27}$$

where T_{amb} is the ambient temperature and $NOCT$ is the Normal Operating Cell Temperature.

Table 3.3: Parameter Values for PV Power Output Equations

Parameter	Symbol	Value	Units
Reference Solar Irradiance	G_{ref}	1000.0	W/m^2
Rated PV Power under STC	$P_{PV-\text{STC}}$	340	W
PV Generation Efficiency	η_{PV}	0.21	–
Temperature Coefficient	β_T	0.004	1/°C
Reference Cell Temperature	$T_{C-\text{STC}}$	25.0	°C
Normal Operating Cell Temperature	NOCT	45.0	°C
Number of Panels	n_{panels}	70	–

Table 3.3 outlines the parameters used in Equations (3.26) and (3.27). Since these are historical irradiance values, they are assumed to be accurate. To simulate the forecasting of these irradiance values, Gaussian noise is applied to these values for planning purposes.

Receding Horizon Specifics

To better simulate a true microgrid environment beyond just using forecasts, several dynamic modifications have been introduced compared to the day-ahead setup. In the day-ahead approach, demand response signals are static and generated only once. In contrast, the receding-horizon approach integrates dynamically generated demand signals for time intervals between 4 and 6 hours into the future planning horizon at each RHC step.

Additionally, the receding-horizon experiment features more dynamic charging windows for EV chargers connected to the microgrid. The start and end times for the EV connection are randomly generated, and whenever an EV is detached, a new EV is introduced into the planning horizon with updated connection schedules.

Furthermore, to simulate real-world battery dynamics more accurately, including potential model mismatch or unmodeled effects during operation, stochastic noise is introduced into the SOC propagation during the simulation execution phase. This is distinct from the idealized, deterministic propagation assumed by the planner (Eq. 3.2). Specifically, after

the optimal control action $U_{b,k}$ is determined by the planner for the current interval and applied in the simulation, the subsequent simulated SOC update incorporates a small random perturbation that represents process noise. This tests the RHC's ability to correct for deviations between planned and actual battery states.

As mentioned earlier, while the RHC framework adapts to forecast errors and process noise over successive time steps through feedback and re-planning, the optimization problem solved within each time step remains deterministic. This work does not employ stochastic or robust optimization techniques to explicitly manage uncertainty within the optimization formulation. Incorporating such methods to potentially improve robustness against significant short-term forecast errors or system variability is acknowledged as a limitation and remains an important direction for future work.

3.4 Results

This section presents the experimental results of the proposed hierarchical optimization framework for microgrid energy management. First, the performance of the forecasting model is examined through quantitative metrics. Next, simulation results are discussed for day-ahead local grid utilization planning, illustrating the capability of the framework to reduce peak power consumption and optimize battery usage under static conditions. Subsequently, the results for the day-ahead demand response scenarios are presented. Finally, the performance of the Receding Horizon Control (RHC) implementation over a 24-hour simulation is analyzed, showcasing its adaptability to forecast errors and dynamic conditions, and comparing its outcome to the initial day-ahead plan. Together, these results validate the effectiveness of our approach and identify directions for future refinement.

3.4.1 Forecasting Training Results

Three loss functions were used to assess forecast precision: the L1 loss (Mean Absolute Error), which penalizes errors uniformly; the L2 loss (Mean Squared Error), which disproportionately penalizes large errors; and the explained variance, which evaluates how well the predicted trends match the actual data. Table A.1 reports the loss values for the trained

XGBoost model used in the RHC simulations.

Table 3.4: Loss Results for XGBoost Forecasting Model

Loss Type	**XGBoost**
L1 Loss	3.268E-2
L2 Loss	1.765E-3
Explained Variance	0.9654

The low L1 loss of approximately 0.033 indicates that, on average, the absolute error in the normalized predictions is minimal. Similarly, the very low L2 loss suggests that large errors are infrequent. An explained variance of 0.9654 reveals that the model accounts for approximately 96.5% of the variability in the observed load data. Together, these metrics suggest that the XGBoost model achieves strong forecasting performance, captures the underlying trends in power consumption with high precision, providing a reliable basis for the RHC planner.

Figures 3.4 and 3.5 provide visual assessments of the performance of the forecasting system at different points during the simulation (Iteration 0 and Iteration 6, respectively). In both figures, the top panel compares the predicted net external load (consumption minus forecasted generation, orange line) against the true net load (blue line), while the bottom panel compares the PV generation forecast derived from noisy irradiance data (orange line) against the generation calculated from true irradiance (blue line). These plots confirm the reasonable accuracy of the forecasting methods while visualizing the forecast errors that the RHC must manage.

3.4.2 Results of Day-Ahead Local Grid Utilization Planning

The hierarchical objectives for the local grid utilization were tested using a static daily profile derived from historical average data (Figures 3.2 and 3.3), resulting in a maximum net external load of 301 kW before optimization. The system included a primary 300 kWh ESS (100 kW charge/discharge limits) and 10 simulated EVs. Figure 3.6 illustrates the outcome after each hierarchical step. Each pair of panels shows the resulting grid power profile (top)

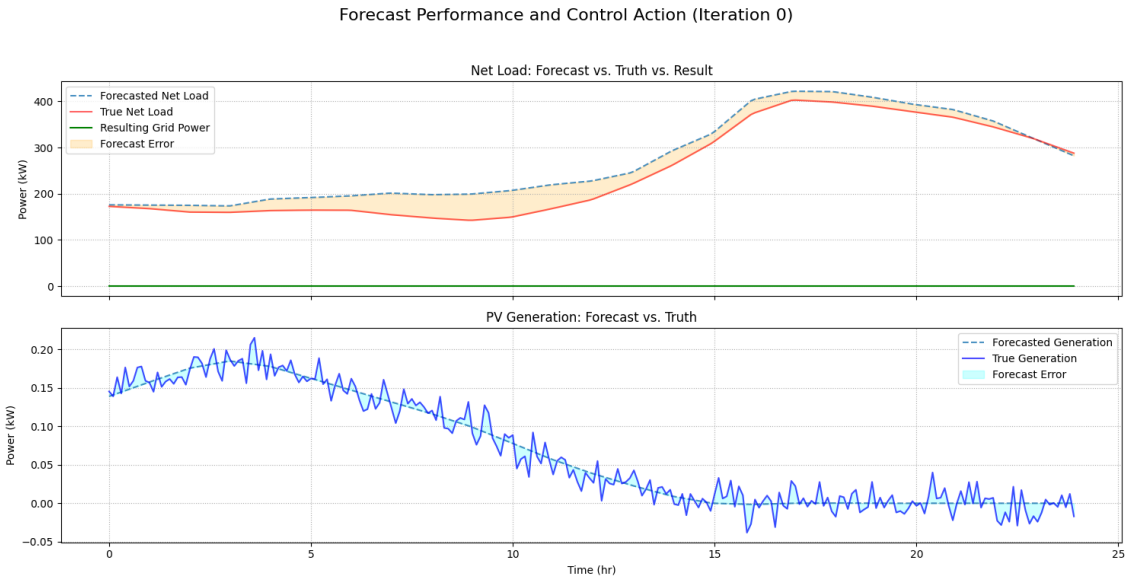

Fig. 3.4: Forecasting Performance at RHC Iteration 0. (Top) Net Load Forecast vs. Truth. (Bottom) PV Generation Forecast vs. Truth.

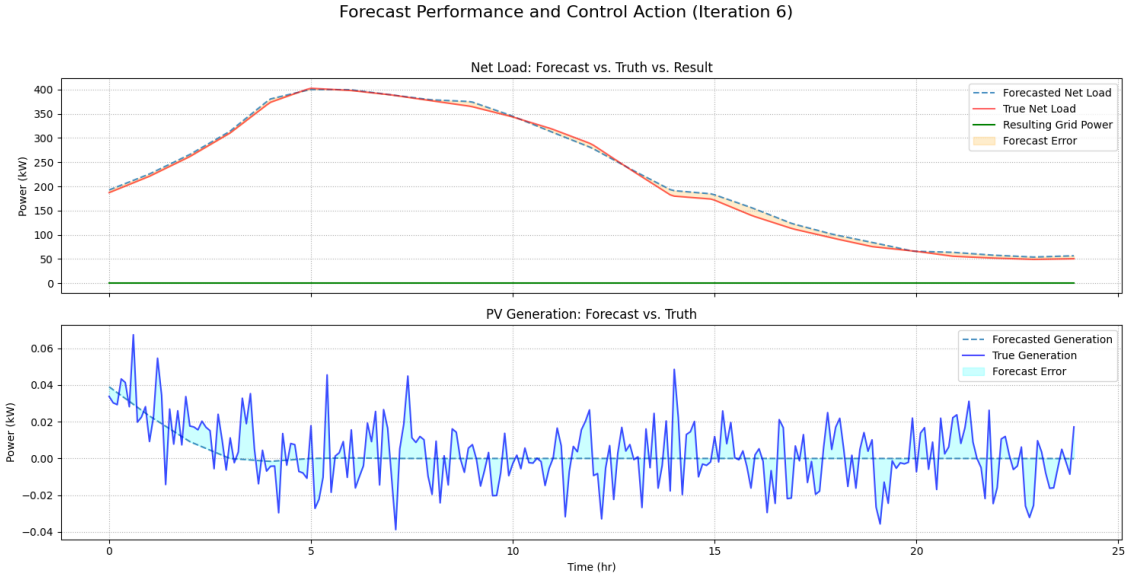

Fig. 3.5: Forecasting Performance at RHC Iteration 6.

and the corresponding battery SOC trajectories (bottom). Initially (top-left), minimizing peak power (p_{max}) significantly reduces the maximum grid draw to approximately 150 kW. Subsequently (bottom-left), maximizing the minimum power (p_{min}) further smooths the

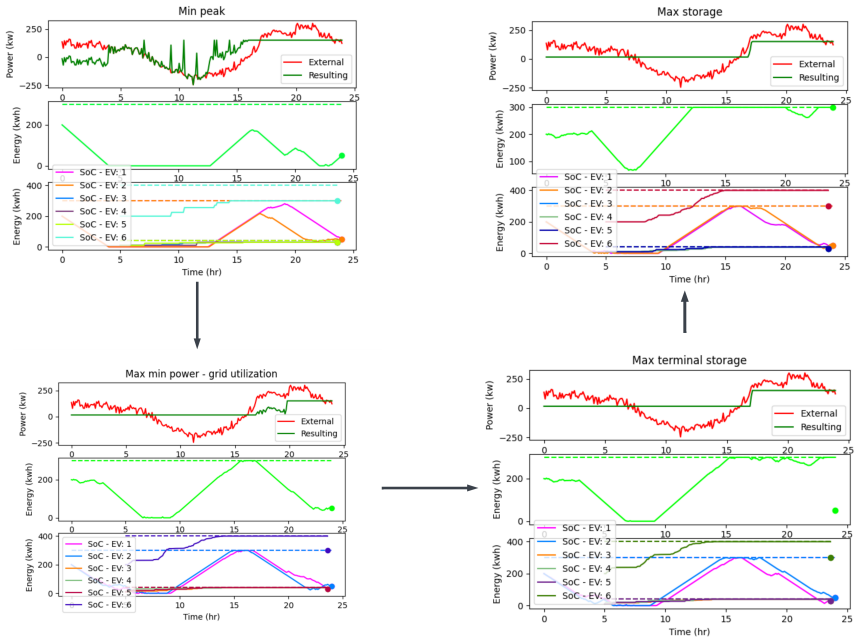

Fig. 3.6: Day-Ahead Local Grid Utilization Results. Each quadrant shows Grid Power (top plot) and Battery SOCs (bottom plot) after sequential optimization: Minimize Peak Power (Top Left), Maximize Minimum Power (Bottom Left), Maximize Terminal SOC Bat 0 (Bottom Right), and Maximize Overall SOC (Top Right).

profile. The third step (bottom-right), maximizing the terminal SOC of the main ESS (S_{1,N_t+1}), adjusts operation to meet this resilience goal. Finally (top-right), maximizing the overall system SOC ($\sum\sum S_{b,k}$) fine-tunes EV charging within the remaining flexibility. These results demonstrate the hierarchy successfully achieves prioritized operational goals in a day-ahead context.

3.4.3 Demand Response Day-Ahead Results

Constrained Demand Response

This scenario demonstrates the framework's ability to handle demand response signals within the day-ahead planning context, using the same system setup as Section 5.4-B but with added DR signals: an incentive (O_{pos}) of 100 kW during hour 1 and a curtailment limit (O_{neg}) of 60 kW during hour 4. Figure 3.7 shows the results after each hierarchical

step, with subplot layouts mirroring Figure 3.6. The optimization successfully incorporates both signals. The incentive allows higher grid draw in hour 1 without penalty, while the curtailment limit forces grid power below 60 kW during hour 4. The hierarchy proceeds similarly to the local utilization case but respects these additional DR constraints, showing the framework's capability to integrate DR signals into the prioritized planning.

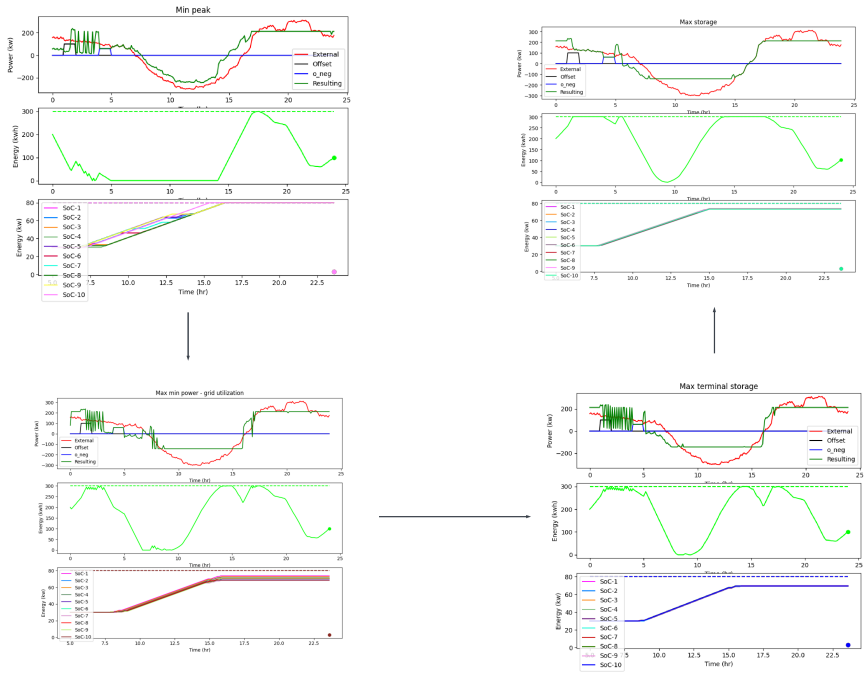

Fig. 3.7: Day-Ahead Demand Response Results (Constraint Method). Subplots show outcomes after each hierarchical objective (layout identical to Figure 3.6), incorporating O_{pos} and O_{neg} signals.

3.4.4 Receding Horizon Control (RHC) Simulation Results

This subsection presents the outcome of the RHC strategy applied over a 24-hour simulation period, which included dynamic forecasting, variable EV availability, and simulated process noise on the battery's state of charge (SOC).

A key insight from the RHC simulation was the behavior of the strict hierarchical optimization under dynamic conditions. In every iteration, successfully solving the first three

high-priority objectives (minimizing peak power, minimizing power range, and maximizing the building battery's final SOC) left no feasible solution space for the fourth and final objective (Maximize Overall SOC).

This outcome triggered the activation of the pre-defined backup solver, demonstrating the framework's built-in robustness. The system is designed to prioritize the most critical operational goals, and when they fully constrain the problem, it seamlessly transitions to a backup strategy. The effectiveness of this approach is confirmed by the fact that this backup solver successfully met all 111 active curtailment signals without a single violation, as shown later in Figure 3.12.

24-Hour Aggregated RHC Results

Figure 3.8 aggregates the RHC results for the entire 24-hour simulation. The top panel shows the initial day-ahead plan against the actual external load and the final grid power achieved by the RHC. The RHC effectively adapted to changing conditions, maintaining a stable power profile despite deviations from the initial forecast. The middle and bottom panels illustrate the dynamic nature of the demand response signals, with the RHC receiving and reacting to different incentive and curtailment signals than those anticipated in the day-ahead plan.

Figure 3.9 provides a closer look at the system's internal response. The top panel shows the SOC trajectories for all batteries, including the main ESS and all connected EVs, while the bottom panel displays the power charge/discharge actions taken by the controller.

RHC Performance Analysis

To better assess the peak-shaving performance of the RHC, Figure 11 compares the peak power of the unplanned external load, the day-ahead plan, and the final RHC-controlled outcome.

To better assess the peak-shaving performance of the RHC, Figure 3.10 compares the peak power of the unplanned external load, the day-ahead plan, and the final RHC-controlled outcome.

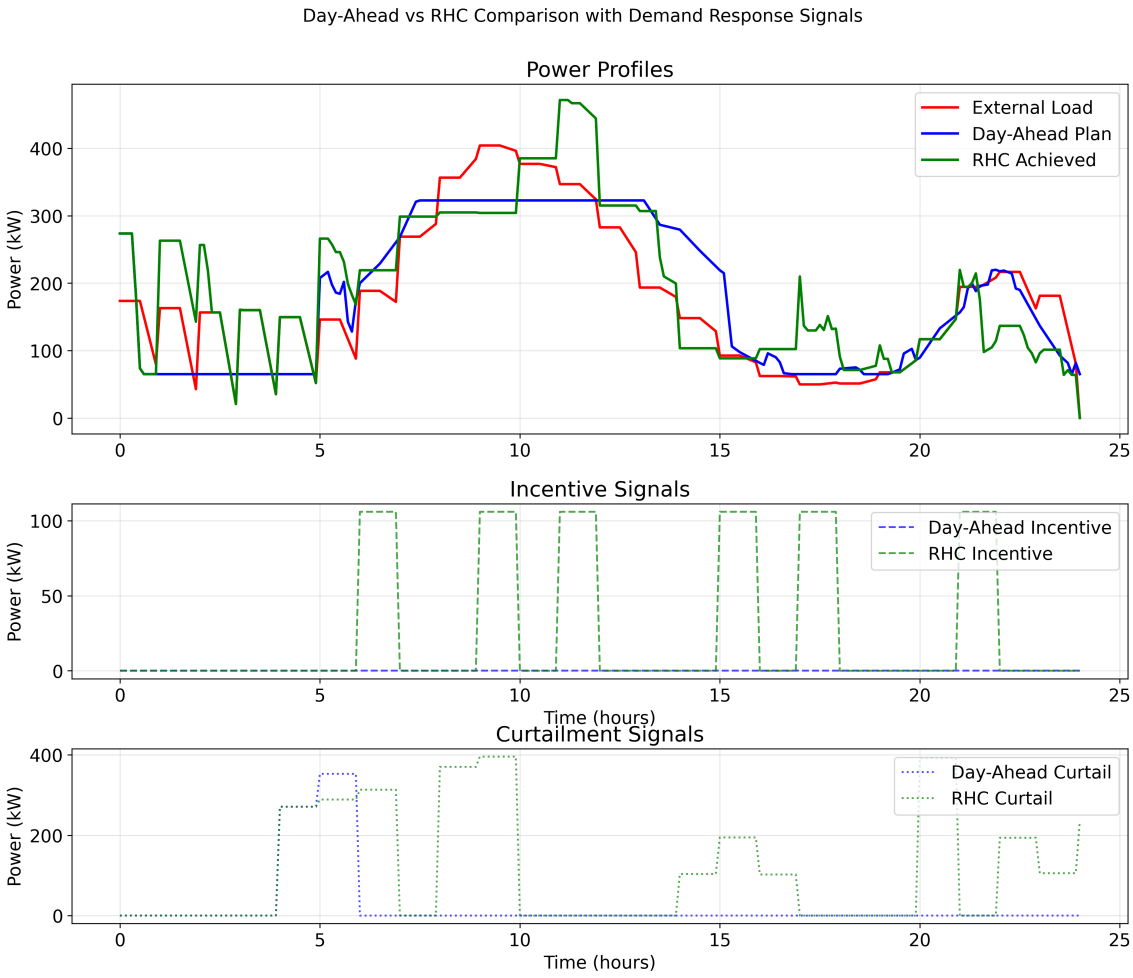

Fig. 3.8: Comparision of RHC Power plan vs day ahead planner. (Top) Power Profiles: Initial Forecast, Resulting Grid Power, Day Ahead Plan. (Middle) Incentive Signals: Day Ahead Plan (blue dashed line), and the RHC (green dashed line), (Bottom) Curtailment Signals: Day Ahead Plan (blue dashed line), and the RHC (green dashed line)

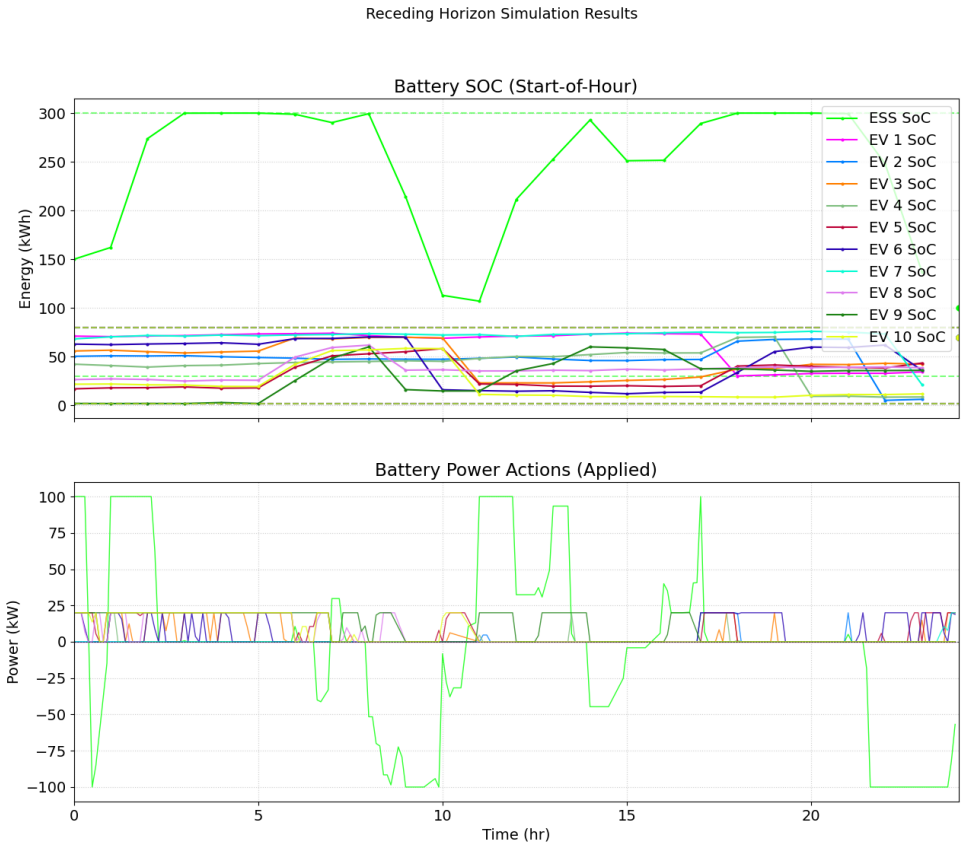

Fig. 3.9: Aggregated System Response Over 24-Hour RHC Simulation. (Top) Battery SOCs (including noise). (Bottom) Applied DER Power Actions.

Interestingly, the day-ahead plan resulted in a lower peak than the RHC implementation. To understand this, Figure 3.11 shows the RHC's plan at hour 13, which was the time of the peak.

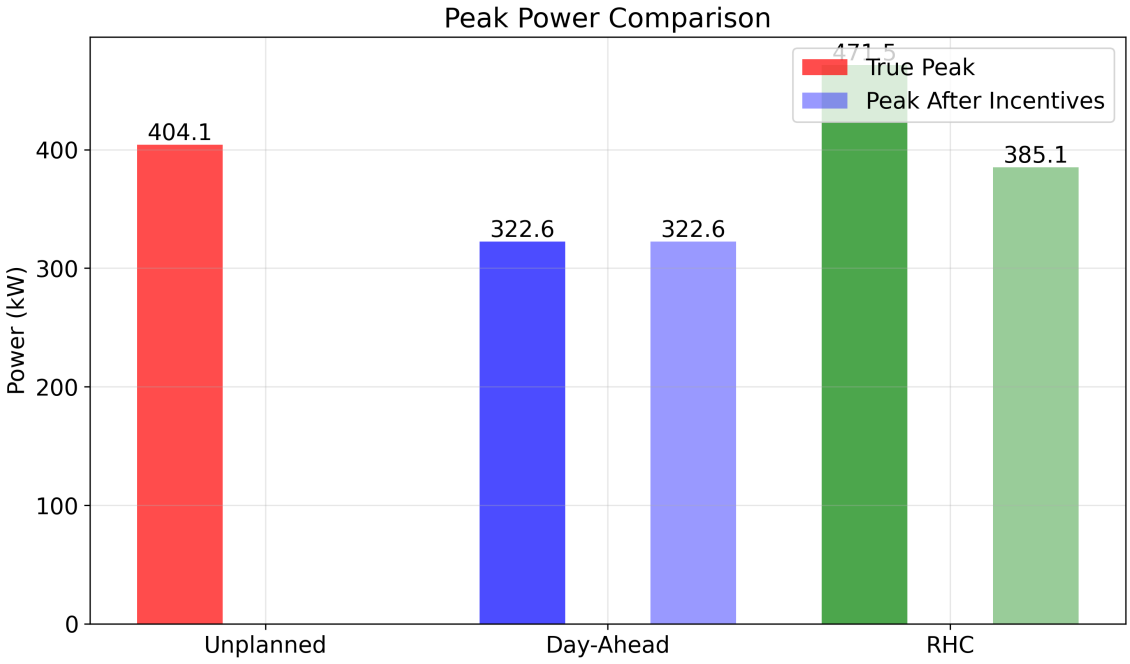

Fig. 3.10: This chart compares the unplanned peak power with the planned peak power from the day-ahead schedule and the actual peak power achieved with the RHC controller. It also distinguishes between the true peak and the peak after accounting for any "free power" from incentive signals.

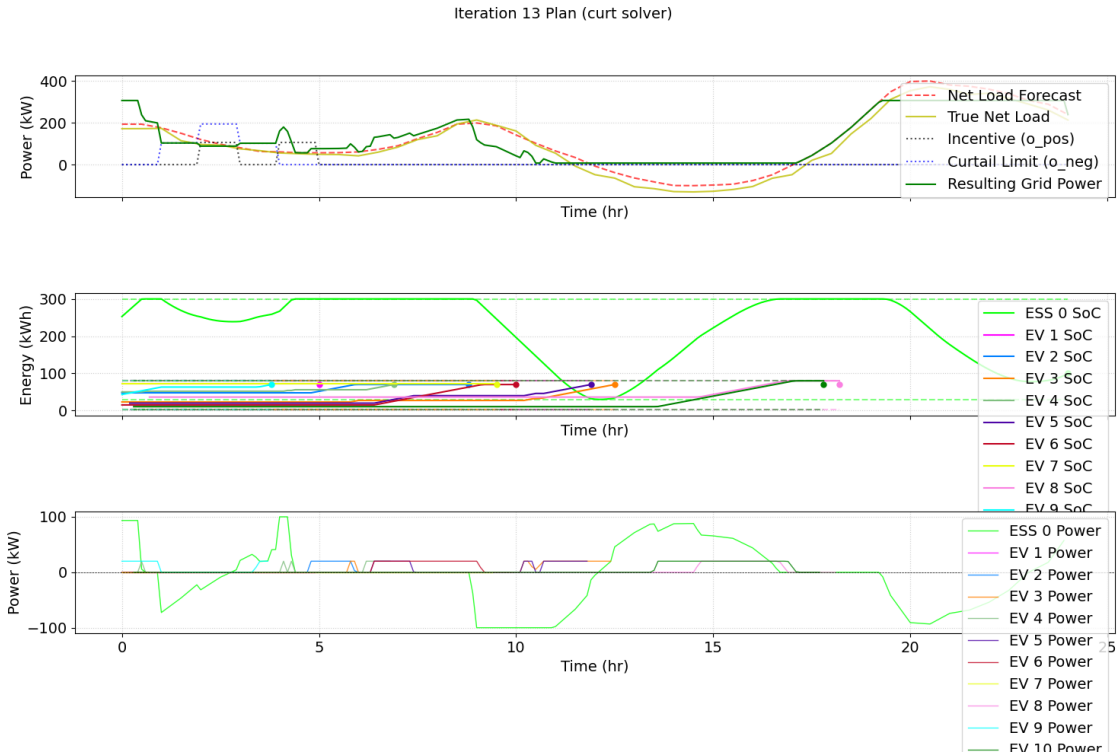

Fig. 3.11: This figure shows the RHC's plan at hour 13, including the net load forecast, the resulting grid power, and the status of all batteries and EVs.

The planner at hour 13 was attempting to minimize a higher peak forecasted for later in its 24-hour horizon. This forward-looking action, while optimal for the entire window, led to a higher peak at that moment than the day-ahead plan scheduled. The controller prioritized mitigating the larger future peak, making the immediate peak a secondary, non-limiting factor in its executed plan.

Despite the difficulties of the primary solver, the backup solver proved to be highly effective. Figure 3.12 demonstrates the RHC's success in meeting all curtailment signals.

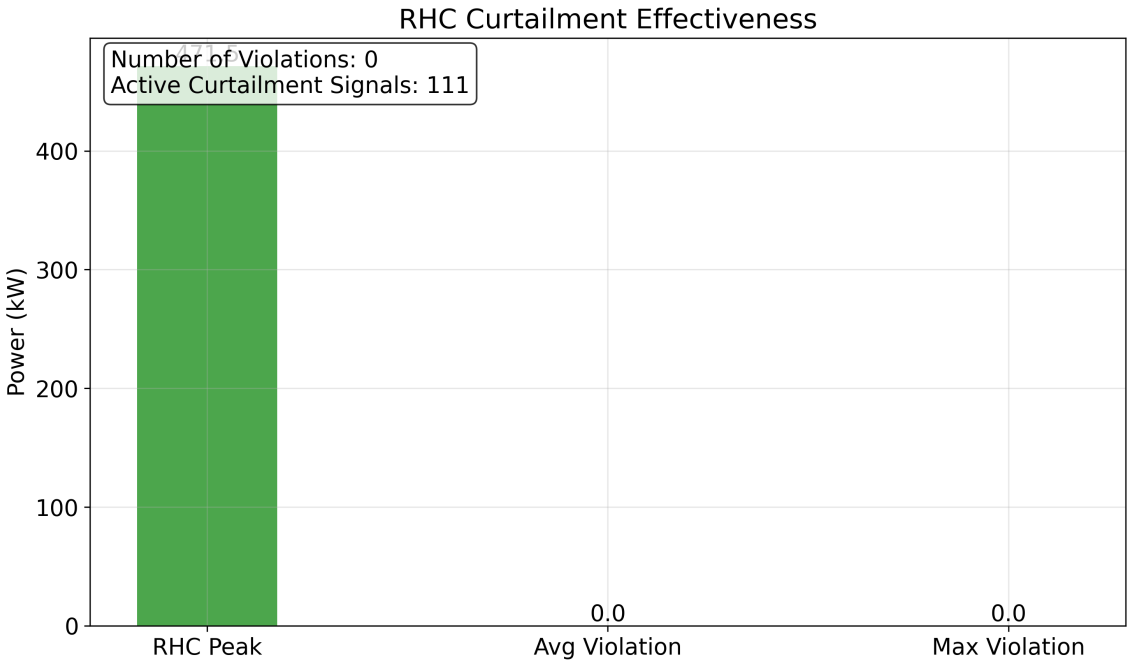

Fig. 3.12: This chart shows the RHC's performance in response to curtailment signals. It highlights that there were zero violations despite 111 active curtailment signals during the simulation.

As shown in Figure 3.12, the RHC successfully met all 111 curtailment signals without a single violation, confirming the robustness of the backup strategy.

3.4.5 Conclusion

This work presented a comprehensive hierarchical optimization framework for microgrid energy management that integrates various DERs, including ESS, EV chargers, and photovoltaic systems, into a unified control strategy. The framework leverages both day-ahead planning and receding-horizon control, combined with advanced XGBoost forecasting, to balance peak load reduction, grid utilization, and demand response objectives. By incorporating dynamic constraints and a layered objective structure, the framework ensures that higher-priority goals, such as minimizing peak power consumption and power range, are addressed while sequentially optimizing additional objectives like maximizing battery SOC.

Experimental results confirmed the effectiveness of the framework. The XGBoost forecasting module demonstrated strong performance with low error metrics and high explained variance, providing a reliable input for the planner. The day-ahead simulations illustrated a significant reduction in peak power (e.g., from 301 kW to 150 kW) and the successful incorporation of demand response signals. The 24-hour RHC simulation further highlighted the adaptability of the framework, managing DERs effectively despite forecast inaccuracies and dynamic conditions.

Furthermore, the 24-hour RHC simulation highlighted the practical implications of a strict hierarchical approach. The simulation revealed that fully optimizing for the highest-priority objectives consistently constrained the problem to the point where the final, least-critical objective became infeasible. However, the framework's design anticipated this outcome. The consistent and effective activation of the backup solver demonstrated the system's resilience, ensuring that all critical grid-level objectives, such as demand curtailment, were successfully met without violation.

This finding underscores a fundamental trade-off in real-time control between maintaining a strict priority order and the need for operational flexibility. The results suggest that for practical implementation, a robust hierarchical controller benefits from including fallback strategies like the one presented.

This outcome warrants further investigation into constraint relaxation techniques or alternative hierarchical formulations for the primary solver. Despite the observed solver behavior, the dual experimental approach illustrates the framework's potential in addressing both predictable day-ahead scenarios and the dynamic, uncertain conditions encountered in real-time operation.

CHAPTER 4

Lumped Sum Graph Creation

4.1 Introduction

Efficient thermal management in buildings is crucial for both energy conservation and occupant comfort. However, designing effective control systems or accurately predicting thermal behavior requires navigating the complexities of heat transfer, a field intertwining thermodynamics and fluid mechanics. Detailed models involving partial differential equations (PDEs) can capture these dynamics precisely, but are often computationally prohibitive for real-time control applications or large-scale simulations.

To address this challenge, this work employs the lumped-sum modeling approach. The principal advantage of the lumped-sum method lies in its ability to represent complex thermal systems using a set of ordinary differential equations (ODEs) in time, significantly simplifying analysis and control design. This simplification is achieved by replacing the distributed nature described by partial differential equations with ODEs representing the thermal behavior at discrete points within the system.

This approach necessitates certain simplifying assumptions. Firstly, it assumes that thermal properties such as resistance and capacitance are concentrated, or 'lumped', at specific points rather than being distributed throughout the space. Secondly, to achieve linear models suitable for many control techniques, non-linear heat transfer mechanisms like radiation are often approximated using their dominant linear terms.

The state of the art in lump sum building thermal modeling encompasses a variety of Resistor-Capacitor (RC) network structures, ranging from simple first-order models (1R1C) to more complex configurations (e.g. 3R2C, 5R4C, 6R4C) that incorporate detailed dynamics of walls, internal mass, and interactions with heating/cooling systems [1, 3]. The choice of model complexity typically involves a trade-off between fidelity and computational

tractability [3]. For multizone buildings, modeling interactions between zones becomes essential, leading to larger-scale interconnected RC networks. Graph-based methods have been proposed to systematically construct these multizone models, leveraging the network structure inherent in the building layout [6]. The control strategies applied to these models range from classical feedback and PID control to advanced methods such as model predictive control (MPC), which explicitly uses the dynamic model to optimize control actions over a future horizon [1, 61].

This chapter focuses on the development, analysis, and control of lumped-sum RC models for the construction of thermal zones. Key contributions are presented throughout this work. An adaptation and extension of an existing 5R4C model [1] to a 5R5C representation is introduced, explicitly including the dynamics and energy input associated with a heating element of water storage, providing a more integrated view of energy consumption. Furthermore, a systematic, graph-based formulation for constructing state-space models of complex multizone thermal systems is developed. This formulation generalizes previous approaches [6] by offering greater flexibility in handling connections to various external node types and allowing inputs or disturbances to affect linear combinations of state variables. A significant theoretical contribution is a rigorous analysis of the stability properties of the aggregate multizone system matrix derived from this graph-based RC network. Specifically, Lemma 1 provides a novel proof that this matrix is Hurwitz, guaranteeing inherent stability and detectability based on the physical structure of the thermal network, leveraging Gershgorin and Taussky's theorems [62]. This result underpins the reliability of control designs based on these models. Finally, the work details the application and comparison of state-space feedback control (including integral action for disturbance rejection) and optimal control strategies formulated as Quadratic Programs (solved using OSQP [63]) for single-zone and multizone examples, with the optimal control framework incorporating practical considerations such as peak power minimization.

This chapter is structured as follows. First Table 4.1 Section 4.2 introduces the concept of thermal zones and the fundamentals of lumped-sum RC modeling, illustrating various

configurations of the RC model. Next, Section 4.3 details a representative single-zone 5R5C model, adapting an existing 5R4C model to include water storage dynamics, and deriving its state-space representation. Section 4.4 then discusses the design of discrete-time state-space feedback control for the single-zone model, including equilibrium calculation and integral control for disturbance rejection, while Section 4.5 presents the formulation of optimal control for the single-zone system as a quadratic program, including how to wrap state-space integral control around the optimal solution. Following this, Section 4.6 develops a systematic graph-based approach for modeling multizone thermal systems, defining graph matrices (adjacency, degree, Laplacian) and deriving the aggregate system dynamics; this section also includes a proof (Lemma 1) of the Hurwitz stability of the aggregate system matrix. Section 4.7 provides a detailed example of a multizone (two-floor, eight-zone) system, discussing its model configuration and the formulation of an optimal control problem that incorporates peak power considerations. Subsequently, Section 4.8.2 (labeled in the provided text as "Results" and containing subsections "Verification of Controllability and Observability", "Single-zone Control Observations", and "Multi-zone Results") presents simulation results and analysis for both single-zone and multizone control strategies, comparing different control approaches and verifying model properties. Finally, Section 4.9 provides the closing remarks and conclusions for the chapter.

Table 4.1: Chapter 4 Problem Terminology

Lumped-Sum RC Modeling Terminology	
R	Thermal Resistance (e.g., R_{ij} between nodes i, j)
C	Thermal Capacitance (e.g., C_i at node i)
T	Temperature at a node (e.g., T_i, T_a, T_r, T_c, T_w)
$\dot{Q}$	Heat flow rate (input/disturbance, e.g., Q_{fcu}, Q_{ws}, Q_S)
x	State vector (temperatures of nodes with capacitance)

Continued on next page

Table 4.1: Chapter 5 Problem Terminology – Continued

Lumped-Sum RC Modeling Terminology (Continued)	
u	Control input vector (applied heat flow rates)
d	Disturbance vector (external temperatures, other heat flows)
A, B, E	Continuous-time state-space matrices ($\dot{x} = Ax + Bu + Ed$)
$\bar{A}, \bar{B}, \bar{E}$	Discrete-time state-space matrices ($x_{k+1} = \bar{A}x_k + \bar{B}u_k + \bar{E}d_k$)
Feedback Control Terminology (Sec 4.4)	
x^d	Desired equilibrium state (temperature)
x^e, u^e	Equilibrium state and control for nominal disturbance d^e
$z_k = x_k - x^e$	State error vector at time k
$\hat{u}_k = u_k - u^e$	Control error vector at time k
Q, R	Weighting matrices for LQR cost function
K	State feedback gain matrix ($\hat{u}_k = -Kz_k$)
σ	Integral state for disturbance rejection
δ	Augmented state vector including integral state ($[z; \sigma]$)
A_{aug}, B_{aug}	Augmented state-space matrices for integral control
K_{aug}	Augmented state feedback gain matrix ($\hat{u}_k = -K_{aug}\delta_k$)
$\mathcal{C}$	Controllability Matrix
$\mathcal{O}$	Observability Matrix
Optimal Control Terminology (Sec 4.5)	
x_k^d, u_k^d	Desired state/input trajectory at time k
x_k^e, u_k^e	State/input error vectors ($x_k - x_k^d, u_k - u_k^d$)
P	Terminal state weighting matrix
x_{min}, x_{max}	State bounds
u_{min}, u_{max}	Input bounds

Continued on next page

Table 4.1: Chapter 5 Problem Terminology – Continued

Lumped-Sum RC Modeling Terminology (Continued)	
y	Augmented optimization vector ($[x_1..x_N; u_0..u_{N-1}]$)
H, c	Quadratic Program (QP) cost matrices ($0.5y^T Hy + c^T y$)
A_{aug}, l, u	QP constraint matrices/vectors ($l \leq A_{aug} y \leq u$)
p_k	Power consumed at time k (sum of inputs)
p_{peak}	Peak power over optimization horizon
Multi-zone Graph Formulation Terminology (Sec 4.6)	
v_i	Node i in the thermal graph
x_w, x_z	State vectors for wall and zone nodes, respectively
d_e	Vector of external node temperatures (disturbances)
d_w, d_z	Disturbance vectors acting directly on wall/zone nodes
$\mathcal{N}_i$	Neighborhood set (connected nodes) for node i
A_G	Adjacency matrix (conductances $1/R_{ij}$) of the full graph
D_G	Degree matrix (sum of conductances from node i) of the full graph
$L_G = D_G - A_G$	Laplacian matrix of the full graph
$L_{ww}, L_{wz}, ...$	Blocks of the Laplacian matrix based on node type
C_w, C_z	Diagonal capacitance matrices for wall/zone nodes
$A_{ww}, A_{wz}, ...$	Blocks of the aggregate state matrix A
$B'_u, B_{we}, ...$	Blocks of aggregate input/disturbance matrices B, E

4.2 Lumped Sum Modeling for Thermal Zones

A building often contains multiple distinct areas where temperature is controlled, typically by a dedicated thermostat. Such an area, which might encompass one or more rooms, is referred to as a *thermal zone*. Analyzing the complex heat transfer within these zones often requires simplified approaches suitable for simulation and control design. The *lumped-sum modeling* method provides such an approach, representing the thermal behavior using

an analogy with electrical circuits, commonly known as the Resistor-Capacitor (RC) approach. In this framework, thermal resistances (representing pathways for heat transfer like walls or windows) are modeled as resistors (R), and thermal capacitances (representing energy storage in building mass like walls, air, or furniture) are modeled as capacitors (C). Models are often designated by the number of components used; for instance, a 3R2C model employs three resistors and two capacitors.

The selection of an appropriate RC model involves balancing accuracy with complexity. While more complex models with additional R and C elements can capture thermal dynamics more precisely, they also demand greater computational resources. Simpler models may suffice for scenarios with slow temperature variations, whereas more complex models might be necessary to accurately represent systems with faster dynamics or intricate heat transfer paths [3]. Therefore, the choice depends on the specific building characteristics, the required fidelity, and the goals of the analysis or control design.

Various RC configurations are used in practice, often increasing in complexity to represent more detailed physical phenomena (see Fig. 4.2). Simpler configurations like 1R1C or 2R1C typically model the direct thermal connection between the zone's air mass (T_r) and external temperatures (T_a), often neglecting the dynamics within walls or internal structures.

More complex models, such as 3R2C structures, incorporate additional thermal effects. However, it is crucial to understand the specific schematic associated with a designation like "3R2C," as different arrangements exist [2–6]. For example:

- One common 3R2C configuration (Fig. 4.1, left, similar to [2–4]) models heat transfer between the outside (T_a) and the internal zone (T_r) through two primary paths: one representing heat flow through an external wall (modeled with resistance and capacitance) and another potentially representing heat flow through a window (modeled with resistance only).

- Another distinct 3R2C schematic (Fig. 4.1, right, similar to [5,6]) focuses on modeling the thermal dynamics *within* a wall itself, dividing its thermal mass into two sections

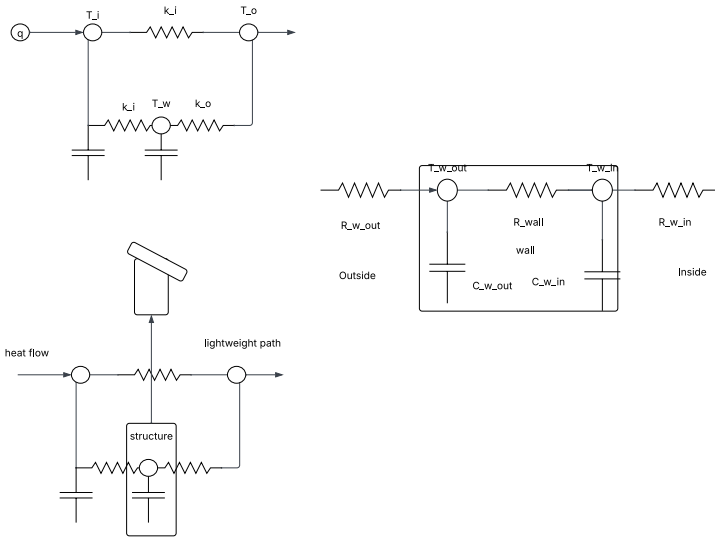

Fig. 4.1: This figure shows two different configurations referred to as 3R2C models. The left image shows a schematic similar to models used in [2–4], while the right shows a schematic similar to models used in [5, 6].

(lumps) connected by resistors.

- Yet another 3R2C arrangement (Fig. 4.2-C, from [3]) adds a capacitor (C_c) to represent the thermal mass of the internal construction (e.g., furniture, internal walls), linking it to the main zone temperature (T_r). This type can be more suitable than 1R1C/2R1C models when internal mass significantly affects thermal dynamics, particularly during more rapid temperature changes [3].

Even more elaborate models, like the 6R4C structure shown in Fig. 4.2, add further elements to capture effects such as detailed multi-layer wall dynamics or thermal interactions between the primary zone and adjacent zones or building structures. While offering potentially higher precision, the improvement over a well-chosen simpler model (like a suitable 3R2C) may sometimes be marginal, as observed in [3].

Understanding these different model structures and their associated trade-offs is essential for selecting an appropriate representation.

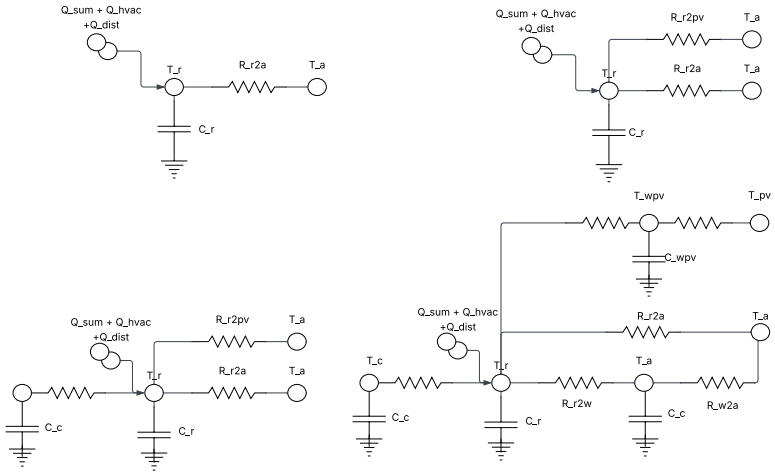

Fig. 4.2: Figure showing various thermal RC models. Adapted from [3] with the addition of the highlight of the 3R2C model (C). T_a corresponds to an external temperature, T_r to the zone of interest, T_c to the internal construction (walls, furniture, etc.) of the building, and T_w to the external wall. Capacitors have the same subscripts, resistor subscripts denote connections between temperature nodes, and $\dot{Q}$ denotes additional heat flow (e.g., solar, HVAC, occupants).

4.3 A Representative Single Zone 5R5C Model

This section develops a representative model of a single-zone system with both reactive and capacitive components. A brief overview of the 5R4C model is given in Section 4.3.1. As the capacitive heating element was designed to be open loop, a representative capacitive component is added in Section 4.3.2 to provide a better sense of energy usage during the control design. The resulting model is a 5R5C model and a state-space representation is provided in Section 4.3.3 and discretized in Section 4.3.4.

4.3.1 5R4C Model from [1]

Fig. 4.3 shows a representation of the 5R4C model developed in [1]. The model includes several major components that are representative of a single-temperature zone. There are six sources of input:

1. T_{ws}: The temperature of the water storage used to heat the floor slabs
2. T_e: The temperature external to the building
3. P_{fcu}: The heat flow from a fan control unit (FCU)
4. P_{ahu}: The heat flow from an air handling unit (AHU)
5. P_S: The heat flow caused by solar radiation
6. P_g: The heat flow caused by the occupants of the building

Both temperature inputs, T_{ws} and T_e, are assumed to be external signals that are not controlled. The equations describing each of these inputs are shown in equations (4.1) - (4.4) and the explanations are left to [1]. Four inputs, P_{fcu}, P_S, P_g, and P_{ahu} are modeled to directly affect the internal temperature, T_i. The other two have indirect affects.

$$P_g = \eta_{occ} + P_{occ} + \sum_i \eta_{dev,i} \times P_{dev,i} \tag{4.1}$$

$$P_{fcu} = \dot{m}_{fcu} \cdot C_p(T_{fcu} - T_i) \tag{4.2}$$

$$P_{ahu} = (1 - \text{eff}) \times \dot{m}_{acu} \cdot C_p(T_e - T_i) \tag{4.3}$$

$$P_s = \left[1 - (1 - \alpha) \times \frac{S_{\text{glazing}}}{S_{\text{glazing}} + S_{\text{walls}}}\right] \times Q_s \tag{4.4}$$

The temperature of the water supply, T_{ws}, directly affects the temperature of the floor slab, T_{fs}, through a modeled conductive resistance R_{fs}. The floor slab has a heat capacitance given by C_{fs}. R_{fs} is again used to denote the resistance to conductive heat between the center of the slab and the surface of the slab, modeled as having temperature T_S. The slab surface then has a convective resistive element (or heat exchange), R_{he1}, between T_S and T_i.

The second indirect input is the external temperature, T_e. T_e has a direct effect on the temperature of the external wall, denoted T_w, through the conductive and convective

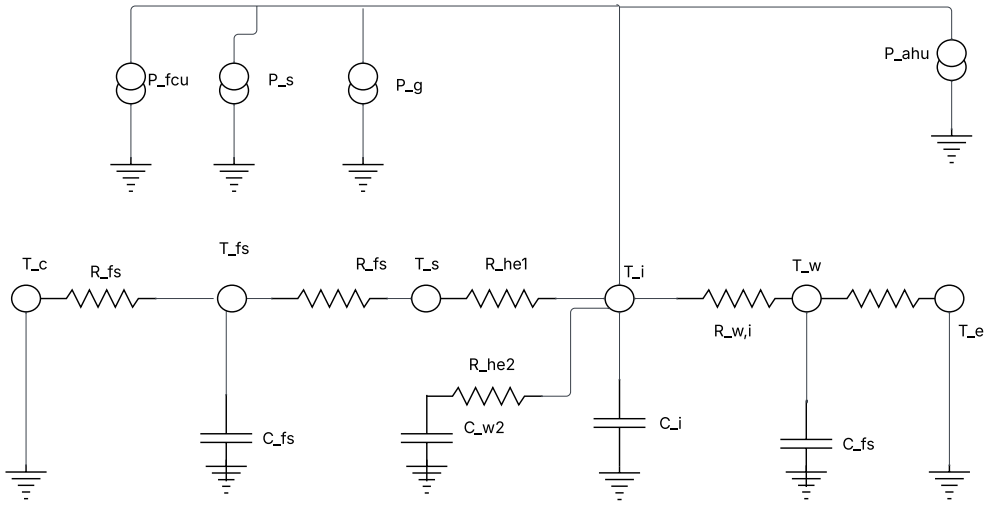

Fig. 4.3: Figure 5 from [1] depicting the 5R4C Model

resistance modeled by $R_{w,e}$. The external wall is assumed to have a heat capacitance of C_w. The external wall temperature then affects the internal room temperature through a heat resistance denoted as $R_{w,i}$ that effectively models both conductive and convective heat transfer.

The internal room has a direct heat capacitance denoted as C_i. There is also an indirect heat capacitance C_{w2} that models the heat capacitance of the internal walls and really any other large item (furniture, etc.). The temperature of the interior wall and the interior room is related through a resistive element denoted as R_{he2}.

In general, there are five internal temperature nodes lumped together:

- T_{fs}: Center of the floor slab
- T_S: The surface of the floor slab (The only node not having an associated heat capacitance)
- T_e: Exterior wall temperature

- T_{w2}: Interior wall temperature
- T_i: Room temperature in the interior (that is, the temperature of interest).

4.3.2 5R5C Model Description

This section takes the inputs to the model developed in [1] and presented in Section 4.3.1 and modifies them. First, the assumption that the water storage temperature is an uncontrollable external signal is relaxed. Second, the inputs that directly affect interior temperature are consolidated.

Having the assumption that the water storage temperature is an uncontrollable external variable may be consistent with HVAC design. However, this does not model well the energy usage for maintaining the temperature constant. As energy usage is fundamental to this work, a model is introduced to relate energy expenditures to the temperature of water storage. T_{ws} is considered a temperature node connected to the lumped sum model as in Fig. 4.3, with two additional connections. The first connection is to an air flow source, Q_{ws}, which has units of Joules per second. The heating of the water storage is assumed to have an efficiency factor that is given by $\kappa_{ws} \in [0\ 1]$. The second connection is to a capacitive element with capacitance C_{ws}.

Water has a specific heat capacity of $4.186 J/gK$ with a rule of thumb of approximately 0.12 liters per meter of pipe (60 liters per 500 meters of pipe). The pipes are typically installed approximately 6 to 8 inches apart. A single liter of water has 1000 grams of water. Assuming a total of 180 liters of water, we have $C_{ws} = 753,480 J/K$.

The second major change in the model is to consolidate each of the external direct inputs into a single input, denoted Q_{fcu}. The resulting model is shown in Fig. 4.5. It is again assumed that the FCU has a known efficiency factor, denoted as $\kappa_{fcu} \in [0\ 1]$.

4.3.3 5R5C State Space Model

Using Kirkoff's current law (applied to heat flow for thermal systems), which states that the input current and output current must be equivalent, a model for the temperature

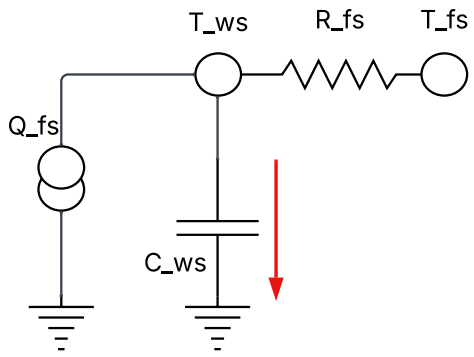

Fig. 4.4: Model for water storage energy usage

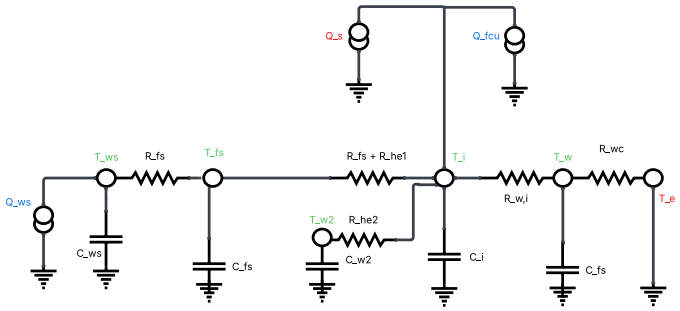

Fig. 4.5: 5R5C Model for water storage energy usage

at each node can be stated. Note that the heat flow through a capacitive element can be stated as $C \cdot \dot{T}$, where C is the heat capacitance and $\dot{T}$ is the time derivative of the temperature.

The sign convention $C \cdot \dot{T} > 0$ is used to denote an outward flow from the node into the capacitive element, as shown in Fig. 4.4. The inward flow to node 1 from node 2 with resistive element R is denoted as

$$\frac{1}{R}(T_2 - T_1).$$

Using these relations / conventions and the node connections depicted in Fig. 4.5, the time evolution for the node at each temperature can be written as

$$\begin{aligned}
\dot{T}_{fs}C_{fs} =& \frac{1}{R_{fs}}\left(T_{ws} - T_{fs}\right) + \frac{1}{R_{fi}}\left(T_i - T_{fs}\right) \\
\dot{T}_{w2}C_{w2} =& \frac{1}{R_{he2}}\left(T_i - T_{w2}\right) \\
\dot{T}_i C_i =& \frac{1}{R_{fi}}\left(T_{fs} - T_i\right) + \frac{1}{R_{he2}}\left(T_{w2} - T_i\right) \\
& \frac{1}{R_{w,i}}\left(T_w - T_i\right) + \kappa_{fcu}Q_{fcu} + Q_S \\
\dot{T}_w C_w =& \frac{1}{R_{w_i}}\left(T_i - T_w\right) + \frac{1}{R_{w_e}}\left(T_e - T_w\right) \\
\dot{T}_{ws}C_{ws} =& \frac{1}{R_{fs}}\left(T_{fs} - T_{ws}\right) + \kappa_{ws}Q_{ws}
\end{aligned} \tag{4.5}$$

A standard state space representation of this model can be made by assigning the states, inputs, and disturbances as

$$x = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} T_{fs} \\ T_{w2} \\ T_i \\ T_w \\ T_{ws} \end{bmatrix}, \; u = \begin{bmatrix} Q_{fcu} \\ Q_{ws} \end{bmatrix}, \; d = \begin{bmatrix} T_e \\ Q_S \end{bmatrix}, \tag{4.6}$$

and rewriting the equations as follows:

$$\begin{aligned}
\dot{x}_0 =& \frac{1}{C_{fs}}\Big[-(\frac{1}{Rfs}+\frac{1}{R_{fi}})x_0+\frac{1}{R_{fi}}T_i+\frac{1}{R_{fs}}T_{ws}\Big]\\
\dot{x}_1 =& \frac{1}{C_w 2}\Big[-\frac{1}{R_{he2}}x_1+\frac{1}{R_{he2}}x_2\Big]\\
\dot{x}_2 =& \frac{1}{C_i}\Big[\frac{1}{R_{fi}}x_0+\frac{1}{R_{he2}}x_1-\\
& (\frac{1}{R_{fi}}+\frac{1}{R_{wi}}+\frac{1}{R_{he2}})x_2+\frac{1}{R_{wi}}x_3+\kappa_{fcu}u_0+d_1\Big]\\
\dot{x}_3 =& \frac{1}{C_w}\Big[\frac{1}{R_{w,i}}x_2-(\frac{1}{R_{w,i}}+\frac{1}{R_{w,e}})x_3+\frac{1}{R_{w,e}}d_0\Big]\\
\dot{x}_4 =& \frac{1}{C_{ws}}\Big[\frac{1}{R_{fs}}x_0-\frac{1}{R_{fs}}x_4+\kappa_{ws}u_1\Big]
\end{aligned} \tag{4.7}$$

Written in vector form, the state equations become

$$\dot{x} = Ax + Bu + Ed, \tag{4.8}$$

where

$$A = \begin{bmatrix}
a_{00} & 0 & \frac{1}{R_{fi}C_{fs}} & 0 & \frac{1}{R_{fs}C_{fs}}\\
0 & \frac{-1}{R_{he2}C_{w2}} & \frac{1}{R_{he2}C_{w2}} & 0 & 0\\
\frac{1}{R_{fi}C_i} & \frac{1}{R_{he2}C_i} & a_{22} & \frac{1}{R_{wi}C_i} & 0\\
0 & 0 & \frac{1}{R_{w,i}C_w} & a_{33} & 0\\
\frac{1}{R_{fs}C_{ws}} & 0 & 0 & 0 & \frac{-1}{R_{fs}C_{ws}}
\end{bmatrix}, \tag{4.9}$$

$$\begin{aligned}
a_{00} &= \frac{-1}{C_{fs}}(\frac{1}{R_{fs}}+\frac{1}{R_{fi}}),\\
a_{22} &= \frac{-1}{C_i}(\frac{1}{R_{fi}}+\frac{1}{R_{wi}}),\\
a_{33} &= \frac{-1}{C_w}(\frac{1}{R_{w,i}}+\frac{1}{R_{w,e}}),
\end{aligned}$$

$$B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \frac{\kappa_{fcu}}{C_i} & 0 \\ 0 & 0 \\ 0 & \frac{\kappa_{ws}}{C_{ws}} \end{bmatrix}, \tag{4.10}$$

$$E = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & \frac{1}{C_i} \\ \frac{1}{R_{w,e}C_w} & 0 \\ 0 & 0 \end{bmatrix}. \tag{4.11}$$

4.3.4 Forming a Discrete-time State-Space Model

An exact discretization of a continuous-time system can be performed under the assumption that the inputs are constant over the discretization period [64]. This is not only possible for an HVAC system, but is quite desirable as rapid changes in inputs can be detrimental to the system itself [65]. The disturbances (outside temperature and solar radiation energy), on the other hand, will be continuously varying. However, the temperature and radiation vary slowly and could be assumed constant at discretization levels of several minutes. The approach taken here is to assume that both the inputs and the disturbances are constant and that the techniques discussed in Section 6.4 of [64] are employed.

In Section 6.4 of [64], a continuous time system of the form $\dot{x} = Ax + Bu$ is disretized as $x_{k+1} = \bar{A}x_k + \bar{B}u_k$ over the time interval T, where

$$\bar{A} = e^{AT},\ \bar{B} = \int_0^T e^{A(T-\tau)}d\tau B, \tag{4.12}$$

and e^A represents the matrix exponential.

Note that (4.8) includes an additional term for the disturbance. This can be accounted for by combining the control and disturbances into a single vector to match the form

$$\dot{x} = Ax + B'\nu, \text{ where } \nu = \begin{bmatrix} u \\ d \end{bmatrix} \text{ and } B' = [B, E], \tag{4.13}$$

where $[B, E]$ is a column-wise concatenation of the matrices. This immediately allows the use of (4.12) to be used for obtaining $\bar{A}$ and $\bar{B}'$ where

$$\bar{B}' = \int_0^T e^{A(T-\tau)} d\tau B'. \tag{4.14}$$

The discretized input and disturbance matrices could be obtained from $\bar{B}'$ simply by noting that $\bar{B}' = [\bar{B}, \bar{E}]$. Thus, $\bar{B}$ is obtained from the first two columns of $\bar{B}'$ and $\bar{E}$ is obtained from the final two. The resulting discretized dynamics for the system can be expressed as

$$x_{k+1} = \bar{A}x_k + \bar{B}u_k + \bar{E}d_k. \tag{4.15}$$

4.4 Single-zone Feedback Control

This section develops a discrete-time feedback control law that can be used to maintain the single-temperature zone at the desired temperature. The controller is developed by first determining an equilibrium state and control in Section 4.4.1 to maintain the zone at the desired constant temperature. This equilibrium is used as a basis for a disturbance-free state-space feedback control based on the Discrete-time Algebraic Riccati Equation (DARE) in Section 4.4.2. As disturbances slowly vary, Section 4.4.3 introduces an integral term to compensate for disturbances.

4.4.1 Calculating an Equilibrium State and Control

The overall objective of the controller is to maintain the internal temperature, T_i, at a desired temperature given, T_{id} (or, in state-space form, x_{d2}). There are no other specifications for the states or inputs. Various possibilities could exist, but the lowest

energy solution is desired. Thus, an optimization problem is developed to find the states and inputs that keep the system in equilibrium.

For a nominal disturbance value, denoted as d_e, the definition of an equilibrium state and control, x_e and u_e, respectively, is given as

$$\dot{x}_e = 0 = Ax_e + Bu_e + Ed_e. \tag{4.16}$$

A quadratic cost is used to solve for the values of x_e and u_e using the following optimization problem:

$$\begin{aligned} &\min_{x_e, u_e} u_e^T Q u_e \\ &\text{s.t. } Ax_e + Bu_e = -Ed_e \\ &\qquad x_{e2} = x_{d2} \end{aligned} \tag{4.17}$$

Note that a continuous-time equilibrium value will also be a discrete-time equilibrium using the exact discretization, i.e.,

$$x_e = \bar{A}x_e + \bar{B}u_e + \bar{E}d_e \tag{4.18}$$

Also note that, generally, there is no guarantee for an equilibrium solution to exist. However, an equilibrium exists for the problem at hand.

To justify this, consider the steady–state equation of our system

$$\dot{x} = Ax + Bu + Ed.$$

At equilibrium ($\dot{x} = 0$) we have

$$Ax_e + Bu_e = -Ed_e.$$

The existence of a solution (x_e, u_e) follows from the following considerations:

1. **Structure of A:** The matrix A is obtained by applying Kirchhoff's law to the RC thermal network. It takes a (scaled) Laplacian form, reflecting the energy balance via resistive interconnections, where each diagonal term (scaled by the inverse capacitance) is the sum of conductances leaving the node. Fixing a reference (for example, $x_{e2} = x_{d2}$) to remove invariance to uniform temperature changes, A becomes invertible in the relevant subspace.

2. **Control via B:** The matrix B maps the available control actions (such as heating and cooling inputs) into the state space. Physical design ensures that these controls affect all modes critical to maintaining the desired temperatures.

3. **Disturbances via E:** The matrix E captures the effect of constant external disturbances (e.g., ambient temperature or solar radiation). With the disturbances d_e assumed constant at equilibrium, $-Ed_e$ is a fixed vector.

4. **Existence of a Solution:** With the reference state fixed, the equilibrium equation

$$Ax_e + Bu_e = -Ed_e$$

   is a linear system in the unknowns (x_e, u_e). The structure of the RC model ensures that the right-hand side, $-Ed_e$, is in the range of the augmented matrix $\begin{bmatrix} A & B \end{bmatrix}$, thus guaranteeing that there is a solution.

Thus, for any constant disturbance d_e, there exist unique states x_e and controls u_e satisfying the equilibrium condition.

4.4.2 Feedback Control Without Disturbances

Given the equilibrium state and control, a controller is designed to keep the state at equilibrium, assuming that the state does not differ from the nominal disturbance (i.e. the predicted external temperature and solar irradecense).

First, consider the change of state $z_k = x_k - x_e$ and let the control be defined as $u_k = \hat{u}_k + u_e$. The discrete-time dynamics can be written as,

$$\begin{aligned} z_{k+1} &= x_{k+1} - x_e \\ &= \bar{A}x_k + \bar{B}(\hat{u}_k + u_e) + \bar{E}d_e - x_e \\ &= \bar{A}(z_k + x_e) + \bar{B}\hat{u}_k + \bar{B}u_e + \bar{E}d_e - x_e \\ &= \bar{A}z_k + \bar{B}\hat{u}_k + \underbrace{\bar{A}x_e + \bar{B}u_e + \bar{E}d_e}_{x_e \text{ (see (4.18))}} - x_e \\ &= \bar{A}z_k + \bar{B}\hat{u}_k \end{aligned} \tag{4.19}$$

The numerical examples in Section 4.8.2 will show that the system $(\bar{A}, \bar{B})$ is completely controllable. This allows the use of the discrete-time Algebraic Riccati Equation (DARE) to be used to determine an optimal controller [66]. The controller takes the form $u_k = -Kz_k$ and optimizes the cost,

$$\begin{aligned} &\min_{u_k} J = \sum_{k=0}^{\infty} \left(z_k^T Q z_k + u_k^T R u_k\right) \\ &\text{s.t. } z_{k+1} = \bar{A}z_k + \bar{B}u_k \end{aligned} \tag{4.20}$$

where $Q = C^T C \succeq 0$ and $R \succ 0$. A solution exists if $(\bar{A}, \bar{B})$ is completely controllable and $(\bar{A}, C)$ is completely observable.

Note that the only state of concern is z_2, corresponding to the error between T_2 and T_{d2}. Choosing $C = \begin{bmatrix} 0 & 0 & \sqrt{q_2} & 0 & 0 \end{bmatrix}$ where $q_2 > 0$ is shown in Section 4.8.2 to satisfy the observability condition. The control weighting matrix is defined as

$$R = \begin{bmatrix} r_0 & 0 \\ 0 & r_1 \end{bmatrix}, \text{ where } r_i > 0. \tag{4.21}$$

Together, the weighting matrices Q and R allow the balance between the deviations of T_{d2} and the energy used by each of the inputs.

Allowing $\hat{u}_k$ to be defined using the resulting value of K, the aggregate control law becomes

$$u_k = u_e - K(x_k - x_e). \tag{4.22}$$

4.4.3 State-space Integral Control for Disturbance Rejection

The external temperature and solar radiation slowly vary throughout any given day. Thus, a control element is needed to compensate for the deviations from the nominal. Due to the fact that the variations are slowly varying, a single integrator is added to the system to counteract the near-constant disturbances.

The integrator is modeled as

$$\dot{\sigma} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \end{bmatrix} z = Gz, \tag{4.23}$$

where $z = x - x_e$ in continuous time. Note that the developments for a change of state in the previous section could have been performed in continuous time before the exact discretization was performed. This results in

$$\dot{z} = Az + B\hat{u}. \tag{4.24}$$

Forming an augmented state, $\delta = [z;\ \sigma]$, to include the integral, the augmented dynamics become

$$\dot{\delta} = \begin{bmatrix} A & 0 \\ G & 0 \end{bmatrix} \delta + \begin{bmatrix} B \\ 0 \end{bmatrix} \hat{u} = A_{aug}\delta + B_{aug}\hat{u}, \tag{4.25}$$

where each 0 is a matrix of the appropriate dimension. Using the discretization process in (4.12), the discretized matrices $\bar{A}_{aug}$ and $\bar{B}_{aug}$ can be found, resulting in the augmented discrete dynamics given by

$$\delta_{k+1} = \bar{A}_{aug}\delta + \bar{B}_{aug}\hat{u}_k. \tag{4.26}$$

A feedback control law, $\hat{u}_k = -K_{aug}\delta_k$ can be found using the DARE approach with the modification that

$$C_{aug} = \begin{bmatrix} 0 & 0 & \sqrt{q_2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sqrt{q_\sigma} \end{bmatrix}, q_2 > 0, q_\sigma > 0. \tag{4.27}$$

As σ constitutes the integral of the error, the resulting control,

$$u_k = u_e - K_{aug}\delta_k, \tag{4.28}$$

is a proportional-integral controller.

4.5 Single-zone Optimal Control

This section discusses the formulation of control as an optimization problem. This is done in a simultaneous formulation [61], which includes the state as a slack variable. Section 4.5.1 presents the problem to be solved, while Sections 4.5.2 through 4.5.3 transform it into a formulation that is suitable for use with an optimization toolbox. Section 4.5.4 provides some details on the representation of the optimization problem without the state slack variables. Section 4.5.5 then discusses wrapping the state-space integral control developed in Section 4.4.3 around the optimal trajectory developed here.

4.5.1 Simultaneous Problem Formulation

A recursive form of the quadratic problem to be solved is presented. Given the desired state and input at time k represented as $x_{d,k}$ and $u_{d,k}$, respectively, with the errors given by

$x_{e,k} = x_k - x_{d,k}$ and $u_{e,k} = u_k - u_{d,k}$, the quadratic programming problem takes the form,

$$\begin{aligned}
&\min_{x_k,u_k} \left[\frac{1}{2} u_{e,0}^T R u_{e,0} + \sum_{k=1}^{N-1} \left(\frac{1}{2} u_{e,k}^T R u_{e,k} + \frac{1}{2} x_{e,k}^T Q x_{e,k} \right) \right. \\
&\qquad \left. + \frac{1}{2} x_{e,N}^T P x_{e,N} \right] \\
&\text{s.t. } x_{x+1} = \bar{A} x_k + \bar{B} u_k + \bar{E} d_k \qquad , \\
&\quad x_{min} \leq x_k \leq x_{max} \quad k = 1, ..., N-1 \\
&\quad x_{min,N} \leq x_N \leq x_{max,N} \\
&\quad u_{min} \leq u_k \leq u_{max} \quad k = 0, ..., N-1
\end{aligned} \tag{4.29}$$

where Q, R and P are weighting matrices, $\leq$ is interpreted element-wise, and the variables of optimization consist of $x_1, x_2, ..., x_N$ and $u_0, u_1, ..., u_{N-1}$. The bounds on x_N effectively allow for a tighter condition on the terminal state for the sake of convergence.

4.5.2 Cost as Augmented Quadratic

While the recursive formulation in (4.29) is convenient for the sake of formulation, the problem must be massaged to fit within the framework of an optimization library. In the sequel, the OSQP [63] library is used as a target. It optimizes a problem of the form

$$\begin{aligned}
&\min_y \frac{1}{2} y^T H y + c^T y \\
&\text{s.t. } l \leq A_{aug} y \leq u
\end{aligned} \quad . \tag{4.30}$$

This section focuses on transforming the objective function in (4.29) to an objective of the form $\frac{1}{2} y^T H y + c^T y$ that will produce the same optimal solution (note that any constant terms are removed). This is accomplished by first expanding the recursive formulation into a standard quadratic form with the constants eliminated. The updated form is then readily amenable to the augmented formulation.

The error terms in (4.29) are first expanded to

$$\begin{aligned}
&\frac{1}{2}\left(u_0^T R u_0 - 2u_{d,0}^T R u_0 + u_{d,0}^T R u_{d,0}\right) + \\
&\frac{1}{2}\sum_{k=1}^{N-1}\Big(x_k^T Q x_k - 2x_{d,k}^T Q x_k + x_{d,k}^T Q x_{d,k} + \\
&u_k^T R u_k - 2u_{d,k}^T R u_k + u_{d,k}^T R u_{d,k}\Big) + \\
&\frac{1}{2}\left(x_N^T P x_N - 2x_{d,N}^T P x_N + x_{d,N}^T P x_{d,N}\right)
\end{aligned} \tag{4.31}$$

Note that the final term on each line is a constant, thus the value produced by minimizing (4.31) would be the same as that of minimizing

$$\begin{aligned}
&\frac{1}{2}\left(u_0^T R u_0 - 2u_{d,0}^T R u_0\right) + \\
&\frac{1}{2}\sum_{k=1}^{N-1}\Big(x_k^T Q x_k - 2x_{d,k}^T Q x_k + \\
&u_k^T R u_k - 2u_{d,k}^T R u_k\Big) + \\
&\frac{1}{2}\left(x_N^T P x_N - 2x_{d,N}^T P x_N\right)
\end{aligned} \tag{4.32}$$

Each instance of u_k and x_k in (4.32) is now in a quadratic form ($z^T M z$) or in a linear form ($c^T z$).

Allow the augmented vector of optimization variables be defined as

$$y = [x_1; x_2; ...; x_N; u_0; u_1; ...; u_{N-1}], \tag{4.33}$$

where the Matlab-style semicolon is used to denote a new row. The quadratic terms in (4.32) can be represented as $\frac{1}{2}y^T H y$, where H is defined as the block diagonal

$$\begin{aligned} H &= blkdiag([\underbrace{Q,...,Q}_{N-1},P,\underbrace{R,...,R}_{N}]) \\ &= \begin{bmatrix} Q & 0 & ... & ... & ... & ... & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & Q & \ddots & & & \vdots \\ \vdots & & \ddots & P & \ddots & & \vdots \\ \vdots & & & \ddots & R & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & ... & ... & ... & ... & 0 & R \end{bmatrix}. \end{aligned} \tag{4.34}$$

The linear term can be written as

$$\begin{aligned} c^T = -[&x_{d,1}^T Q, ..., x_{d,N-1}^T Q, x_{d,N}^T P, \\ &u_{d,0}^T R, ..., u_{d,N-1}^T R] \end{aligned} \tag{4.35}$$

4.5.3 Augmented Constraint

As shown in (4.29), the augmented constraints must be of the form $l \leq A_{aug} y \leq b$. The three elements of this augmented constraint will consist of the dynamic constraint, the limits on the state, and the limits on the input.

To handle the dynamic constraint, it is helpful to expand the recursive dynamic equation and write the terms with the optimization variable on the left and constants on the right, as follows

$$\begin{aligned} x_1 &= Ax_0 + Bu_0 + Ed_0 \rightarrow Bu_0 - x_1 = -Ax_0 - Ed_0 \\ x_2 &= Ax_1 + Bu_1 + Ed_1 \rightarrow Ax_1 + Bu_1 - x_2 = -Ed_1 \\ x_3 &= Ax_2 + Bu_2 + Ed_2 \rightarrow Ax_2 + Bu_2 - x_3 = -Ed_2 \\ &\vdots \end{aligned} \tag{4.36}$$

Given the augmented optimization vector in (4.33), the recursive constraint relationship can be written in matrix form as $A_{dyn}y = b_{dyn}$, where

$$A_{dyn} = \begin{bmatrix} -I & 0 & \dots & \dots & 0 & \bar{B} & 0 & \dots & \dots & 0 \\ \bar{A} & -I & 0 & \dots & \vdots & 0 & \bar{B} & & & \vdots \\ 0 & \ddots & \ddots & & \vdots & \vdots & & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & 0 & \vdots & & & \ddots & 0 \\ 0 & \dots & 0 & \bar{A} & -I & 0 & \dots & \dots & 0 & \bar{B} \end{bmatrix} \tag{4.37}$$

$$b_{dyn} = \begin{bmatrix} -\bar{A}x_0 - \bar{E}d_0 \\ -\bar{E}d_1 \\ \vdots \\ -\bar{E}d_{N-1} \end{bmatrix} \tag{4.38}$$

This equality constraint can quickly be converted to an inequality constraint by noting that $A_{dyn}y = b_{dyn}$ is the same as $l_{dyn} \leq A_{dyn}y \leq u_{dyn}$, where $l_{dyn} = -b_{dyn}$ and $u_{dyn} = b_{dyn}$.

The upper and lower limits can be converted to the augmented inequality form by allowing $A_{lim} = I$ and writing the constraint as

$$l_{lim} \leq A_{lim}y \leq u_{lim} \tag{4.39}$$

where

$$\begin{aligned} l_{lim} &= [\underbrace{x_{min}; ...; x_{min}}_{N-1}; x_{min,N}; \underbrace{u_{min}; ...; u_{min}}_{N}] \\ u_{lim} &= [\underbrace{x_{max}; ...; x_{max}}_{N-1}; x_{max,N}; \underbrace{u_{max}; ...; u_{max}}_{N}]. \end{aligned} \tag{4.40}$$

The complete constraint is created by stacking each of the limit vectors and A matrices vertically.

4.5.4 A Note on the Sequential Formulation

The sequential approach to solving the problem is to optimize over u_k and then use the dynamic equations to calculate the state. Each state is actually a linear combination of the inputs. Although this approach requires far fewer optimization variables, the upper and lower limit constraints are no longer sparse, and neither is the quadratic cost matrix. Due to the OSQP solver taking advantage of sparsity, the sequential approach was found to take much longer to solve. Thus, it is not presented here.

4.5.5 Wrapping the State-space Integral Control Around the Optimal Solution

The optimal solution produced by solving (4.30) produces the best solution, given the objective, assuming that the disturbances are perfectly known. Although an approximate knowledge of the external temperature and amount of solar radiation to be received throughout the day can be fairly accurate, deviations from the forecast will certainly be experienced.

Various approaches could be taken to account for the deviations such as improved modeling and estimation techniques, but no approach will be perfect. These deviations are slow. Thus, the state-space integral control developed in Section 4.4.3 can be a very reasonable solution to adapt to the changes from the expected optimal.

To utilize the feedback control, a single minor, but significant, change is made. Instead of using the desired equilibrium state and control as part of the feed-forward, the optimal trajectory is used. Denote u_k^* and x_k^* as the resulting control and state at discrete time k produced by solving (4.30). A state-space integral control is defined by adapting (4.22) to be written as

$$u_k = u_k^* - K(x_k - x_k^*), \tag{4.41}$$

where K is defined as in Section 4.4.3.

4.6 A Formulaic Approach to Modeling a Multi-zone System

Recall that Section 4.3.3 applied Kirkoff's current law to formulate dynamic equations for each lumped sum node (temperature) to which a capacitive element is directly con-

nected. In this section, a methodical procedure is developed to use Kirkoff's current law with an associated graph-based representation of the lumped sum model. The nodes of the graph correspond to connection points in the lumped sum schematic model that include a capacitive element. Edges connect nodes and always have an associated resistive element. The procedure for graph creation is developed first, followed by a generic form for the individual dynamics of each dynamic element in the lumped sum model. The matrices associated with the resulting graph are defined, and the aggregate dynamics of the lumped sum model is presented. Next, a proof for the stabalizability of the graph is presented. The section ends with two examples using the formulation.

Note that the developments in the Section are largely taken from [6]. A brief remark is made in Section 4.6.4 to denote the differences.

4.6.1 Multi-zone Graph Creation

The graph is created from the lumped sum model by defining the nodes and connecting edges. The nodes and edges are defined methodically to enable the modeling of the aggregate system in Section 4.6.4. In particular, the nodes are defined to enable a well-defined structure for the aggregate dynamics. As mentioned above, the edges form purely resistive paths between the nodes through the lumped sum schematic.

Three types of nodes are defined: wall nodes, zone nodes, and external nodes. Each node has a capacitive element associated with it, except the external nodes, which are assumed to have infinite capacitance. The wall nodes correspond to fixed elements that separate the controlled temperature zones. In reality, they do not have to be walls at all. In the example given below, one "wall" actually corresponds to the floor slab. A better description for the wall nodes would be uncontrolled nodes as they represent lumped sum points that do not have a control input but do have a connected capacitive element. The "zone" nodes represent distinct thermal zones that are being controlled. Although the "single zone" model in Section 4.3.2 has a single zone of interest, it is modeled herein as having two zones: the water storage temperature zone and the internal temperature zone (that is, the zone of interest). External nodes are nodes that are not controlled and have

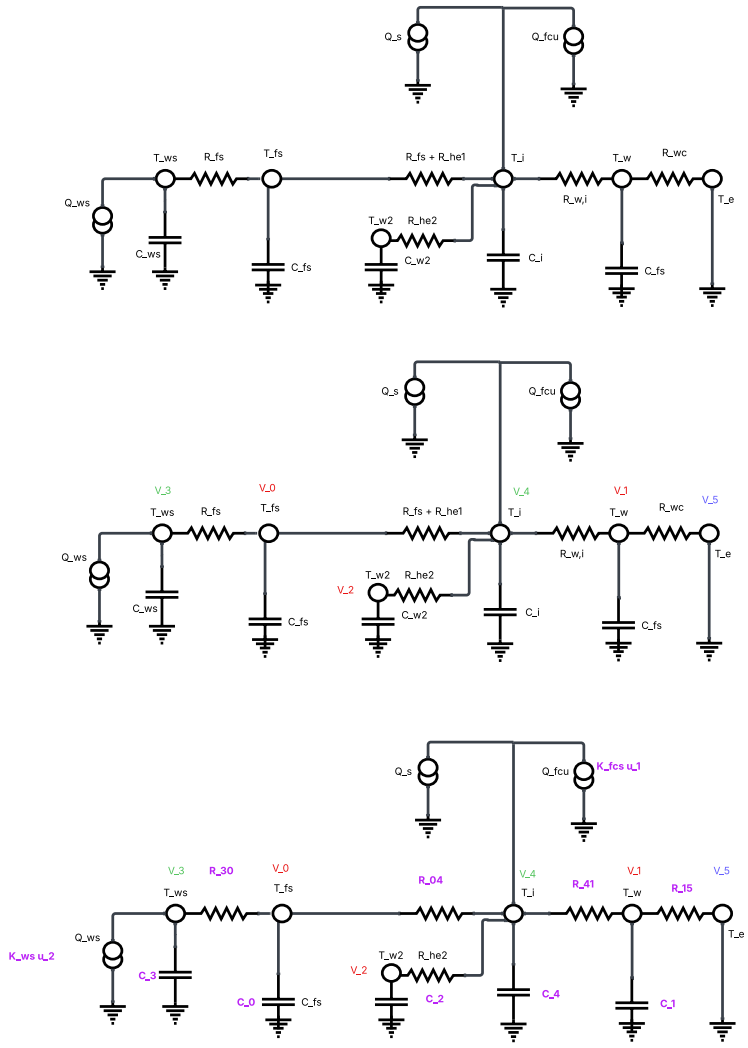

Fig. 4.6: 5R5C Model relabelled using the generic procedure. The top figure duplicates the schematic in Fig. 4.5. The middle figure shows the labeling of vertices (red – walls, green – zones, black – external). The bottom figure relabels the capacitive and resistive elements to according to the node numbering scheme.

no capacitive element. There is a single external node in the 5R5C model that corresponds to the outside temperature of ambient air. Recall that it acts on the system akin to a disturbance. The 5R5C model in Section 4.3.1 would be modeled with two external nodes: one for the ambient outside temperature and one for the water storage temperature.

To achieve a desirable block structure for dynamics in Section 4.6.4, the nodes are numbered in a specific order. The wall nodes are all numbered first, followed by the zone nodes and then by the external nodes. Within each group of nodes, the order is not important. An example is shown in Fig. 4.6, which shows the labeling of the three wall nodes, two zone nodes, and one external node from the 5R5C model in Section 4.3.3. Fig. 4.6 shows a relabeling of the capacitive elements to correspond to the node labeling. C_j corresponds to the capacitance connected to node v_j.

The numbers of wall, zone, and external nodes are represented as n_w, n_z, and n_e, respectively. Each of the nodes forms an element of the node set, V, which is a set of integers. Note that $|V| = n_w + n_z + n_e$. Each wall and zone node will correspond to a dynamic state in the system. A vector of wall states is represented as $x_w \in \mathbb{R}^{n_w}$ and zone states are represented as $x_z \in \mathbb{R}^{n_w}$, with the vector of all states being the stacked vector,

$$x = \begin{bmatrix} x_w \\ x_z \end{bmatrix}. \tag{4.42}$$

The external nodes are considered disturbances and are represented by the vector $d_e \in \mathbb{R}^{n_e}$. The subscript i denotes the i^{th} element of the particular vector (i.e., x_{w_i} is the i^{th} element of x_w).

Edges are associated with pure resistive pathways through the graph between two nodes. For example, the original diagram in Fig. 4.3 shows two resistive values, R_{fs} and R_{he1}, between the nodes corresponding to T_{fs} and T_i. However, T_s does not have a capacitive element connected to it and therefore would not form a node in the graph. T_{fs} and T_{he1} are combined into a single resistance in Figs. 4.5 and 4.6 and would correspond to a single edge.

Associated with each edge is a resistance. Fig. 4.6 shows a relabeling of the resistive elements to correspond to the labeling of the node and edge. $R_{ij} = R_{ji}$ corresponds to the resistance between nodes v_i and v_j. Note that the thermal graph is inherently bidirectional (i.e. heat transfer can occur in both directions), and the resistance is assumed to be symmetric.

The edge set, $E \subset [V]^2$, helps define a neighborhood set for each node. The neighborhood set for node i is denoted as $\mathcal{N}_i$ and corresponds to the set of all edges that begin at node i. Formally, the neighborhood set is defined as

$$\mathcal{N}_i = \{j | \{i, j\} \in E\}. \tag{4.43}$$

4.6.2 Generic Form for Individual Node Dynamics

The labels defined in Section 4.6.1 allow for a generic description of the system dynamics. Recall that Kirkoff's current law allows for the development of dynamic equations for different nodes in the graph. Assuming that the capacitive heat flow is out of the node, the remaining heat flows can be defined as coming into the node.

If heat flow from neighboring nodes were the only incoming flow, the dynamics for node i could be written as

$$\begin{aligned} \dot{x}_i C_i &= \sum_{j \in \mathcal{N}_i} \frac{1}{R_{ij}} (x_j - x_i) \\ \dot{x}_i &= \frac{1}{C_i} \sum_{j \in \mathcal{N}_i} \frac{1}{R_{ij}} (x_j - x_i) \end{aligned} \tag{4.44}$$

However, incoming heat flows come in three forms. The first is the flow of heat from a neighboring node through the resistive element that connects the nodes, as shown in (4.44). The second comes from the control input that acts on the node. The third comes from disturbances on the node.

The defining difference between a wall node and a zone node is the absence of a control input to the wall node, so the wall-node dynamics are defined first. The vector of disturbances for the wall is denoted as d_w and has dimension n_{dw}. They affect x_w through the

matrix $B_w \in \mathbb{R}^{n_w \times n_{wd}}$, where B_{w_i} denotes the i^{th} row of B_w. The dynamics of x_{w_i} can therefore be written as

$$\begin{aligned} \dot{x}_{w_i} C_i &= \sum_{j \in \mathcal{N}_i} \frac{1}{R_{ij}} (x_j - x_i) + B_{w_i} d_w \\ \dot{x}_{w_i} &= \frac{1}{C_i} \Big[\sum_{j \in \mathcal{N}_i} \frac{1}{R_{ij}} (x_j - x_i) + B_{w_i} d_w \Big] \end{aligned} \tag{4.45}$$

Similarly, disturbances $d_z \in \mathbb{R}^{n_{dz}}$ and control input $u \in \mathbb{R}^{n_u}$ affect the zone states, which are contained in the subvector x_z. To properly incorporate these effects within the aggregate state vector

$$x = \begin{bmatrix} x_w \\ x_z \end{bmatrix},$$

a mapping σ_i is introduced. This mapping identifies the index in the full state vector x corresponding to the i^{th} element of x_z. In other words,

$$x_{z_i} = x_{\sigma_i}.$$

Using this convention, the dynamics of the i^{th} zone node can be written as

$$\begin{aligned} \dot{x}_{z_i} \, C_{\sigma_i} &= \sum_{j \in \mathcal{N}_{\sigma_i}} \frac{1}{R_{\sigma_i j}} \Big(x_j - x_{z_i} \Big) + B_{z_i} \, d_z + B_{u_i} \, u, \\ \dot{x}_{z_i} &= \frac{1}{C_{\sigma_i}} \left[\sum_{j \in \mathcal{N}_{\sigma_i}} \frac{1}{R_{\sigma_i j}} \, (x_j - x_{z_i}) + B_{z_i} \, d_z + B_{u_i} \, u \right]. \end{aligned} \tag{4.46}$$

Note that in Equation (4.42) the aggregate state vector x is defined to include only the dynamic states, x_w and x_z. For notational convenience in Equations (4.45) and (4.46), it is assumed that the disturbance vector d_e has been appended to the state vector, even though d_e is not part of the dynamic state components.

4.6.3 Graph Matrices

In Section 4.6.2, the formulation of the equations for a single node was significantly

simplified through the enumeration of the nodes, the relabeling of the resistive and capacitive elements, and the definition of a neighborhood set. To extend this formulation from individual node dynamics to a full system representation, three matrices are introduced: the adjacency matrix, the degree matrix, and the Laplacian matrix.

The adjacency matrix captures the pairwise interactions between all nodes in the system. Each row corresponds to a node, and each column corresponds to the node connected to it. Typically, the adjacency matrix is square and symmetric. In this formulation, the full interaction graph (which includes both dynamic nodes and external nodes) is considered. The complete adjacency matrix is defined as

$$A_G \in \mathbb{R}^{(n_w+n_z+n_e)\times(n_w+n_z+n_e)},$$

where the entry in the i^{th} row and j^{th} column represents the conductance between nodes v_i and v_j. Formally, denoting a_{ij} as the (i,j) entry, the adjacency matrix is given by

$$A_G = [a_{ij}], \quad \text{where } a_{ij} = \begin{cases} \frac{1}{R_{ij}}, & \text{if } j \in \mathcal{N}_i, \\ 0, & \text{otherwise.} \end{cases} \tag{4.47}$$

The degree matrix is a diagonal matrix that records the total conductance of each node. It is defined by

$$D_G = [d_{ij}], \quad \text{where } d_{ij} = \begin{cases} \sum_{k\in\mathcal{N}_i} \frac{1}{R_{ik}}, & \text{if } i = j, \\ 0, & \text{otherwise.} \end{cases} \tag{4.48}$$

The full graph Laplacian, which is used directly to model aggregate dynamics, is defined as the difference between the degree and the adjacency matrices:

$$L_G = D_G - A_G, \tag{4.49}$$

with

$$L_G \in \mathbb{R}^{(n_w+n_z+n_e)\times(n_w+n_z+n_e)}.$$

For clarity, L_G is divided into block components corresponding to the different types of nodes (wall, zone, and external). This partitioning can be expressed as:

$$L_G = \begin{bmatrix} L_{ww} & L_{wz} & L_{we} \\ L_{zw} & L_{zz} & L_{ze} \\ L_{ew} & L_{ez} & L_{ee} \end{bmatrix}, \tag{4.50}$$

where, assuming zero-indexing or proper ordering as defined in Section 4.6.1, the blocks are characterized as follows:

- $L_{ww} \in \mathbb{R}^{n_w \times n_w}$**:** The square submatrix corresponding to the interactions among the wall nodes.
- $L_{wz} \in \mathbb{R}^{n_w \times n_z}$**:** The submatrix corresponding to the interactions between the nodes on the wall (rows 1 to n_w) and the nodes on the zone (columns $n_w + 1$ to $n_w + n_z$).
- $L_{we} \in \mathbb{R}^{n_w \times n_e}$**:** The submatrix corresponding to the interactions between the nodes on the wall (rows 1 to n_w) and the external nodes (columns n_w+n_z+1 to $n_w+n_z+n_e$).
- $L_{zw} \in \mathbb{R}^{n_z \times n_w}$**:** The submatrix corresponding to the interactions between zone nodes (rows $n_w + 1$ to $n_w + n_z$) and wall nodes (columns 1 to n_w).
- $L_{zz} \in \mathbb{R}^{n_z \times n_z}$**:** The square submatrix corresponding to the interactions among the zone nodes.
- $L_{ze} \in \mathbb{R}^{n_z \times n_e}$**:** The submatrix corresponding to the interactions between zone nodes (rows $n_w + 1$ to $n_w + n_z$) and external nodes (columns $n_w + n_z + 1$ to $n_w + n_z + n_e$).
- $L_{ew} \in \mathbb{R}^{n_e \times n_w}$**:** The submatrix corresponding to the interactions between the external nodes (rows $n_w + n_z + 1$ to $n_w + n_z + n_e$) and the wall nodes (columns 1 to n_w).

- $L_{ez} \in \mathbb{R}^{n_e \times n_z}$**:** The submatrix corresponding to the interactions between the external nodes (rows $n_w + n_z + 1$ to $n_w + n_z + n_e$) and the zone nodes (columns $n_w + 1$ to $n_w + n_z$).

- $L_{ee} \in \mathbb{R}^{n_e \times n_e}$**:** The square submatrix representing the interactions among external nodes.

In subsequent analysis, the full matrices are truncated to focus only on the dynamic system (i.e., the wall and zone nodes), while the external nodes are treated as disturbances.

4.6.4 Aggregate Dynamics for Multi-zone System

Given the definition of the Laplacian graph for the multizone thermal system, the aggregate dynamics of the system can be stated. The dynamics takes the form

$$\dot{x} = Ax + Bu + Ed,$$

just as before. This permits the descretization and control schemes previously outlined to be followed quite closely. The particular form of the matrices is now given:

$$\begin{bmatrix} \dot{x}_w \\ \dot{x}_z \end{bmatrix} = \begin{bmatrix} A_{ww} & A_{wz} \\ A_{zw} & A_{zz} \end{bmatrix} \begin{bmatrix} x_w \\ x_z \end{bmatrix} + \begin{bmatrix} 0 \\ B'_u \end{bmatrix} u + \begin{bmatrix} B_{we} & B_{wd} & 0 \\ B_{ze} & 0 & B_{zd} \end{bmatrix} \begin{bmatrix} d_e \\ d_w \\ d_z \end{bmatrix} \tag{4.51}$$

where

$$
\begin{aligned}
A_{ww} &= -C_w^{-1} L_{ww} \\
A_{wz} &= -C_w^{-1} L_{wz} \\
A_{zw} &= -C_z^{-1} L_{zw} \\
A_{zz} &= -C_z^{-1} L_{zz} \\
B_{we} &= -C_w^{-1} L_{we} \\
B_{wd} &= -C_w^{-1} B_w \\
B_{ze} &= -C_z^{-1} L_{ze} \\
B_{zd} &= -C_z^{-1} B_z \\
B'_u &= -C_z^{-1} B_u
\end{aligned}
\tag{4.52}
$$

and $C_w \in \mathbb{R}^{n_w \times n_w}$ and $C_z \in \mathbb{R}^{n_z \times n_z}$ are a diagonal matrices with the capacitance of each wall node and zone node, respectively, in the order they are defined by the node enumeration of Section 4.6.1.

Remark 1. In addition to minor changes in notation, the differences from the above formulation and those in [6] are a few minor generalizations. In [6], a separation is defined between two external nodes – ambient air and ground. In this work, any number of external nodes could be defined. In [6], only the walls can be connected to external nodes. In this work, both walls and zones could be connected to external zones. In [6], the inputs and disturbances effect exactly one state each. In this work, a linear combination of inputs and disturbances can affect the various states.

4.6.5 Stabalizability and Detectability of Aggregate Dynamics

In control theory, ensuring that a system is both stabilizable and detectable is critical to designing controllers that guarantee robust and reliable performance. Stabilizability means that even if not all state variables can be directly controlled, every unstable mode of the system can be influenced by the available inputs so that the overall closed-loop behavior can

be driven to a stable state. Detectability, on the other hand, is the dual of stabilizability; it guarantees that any state that is not directly measurable or reconstructable from the outputs is at least asymptotically stable. Together, these two properties ensure that the unstable behaviors of the system can be corrected and that any hidden dynamics will naturally decay over time.

For the multizone aggregate thermal dynamics, formulated in Section 4.6.4, these properties are essential. They are prerequisites for applying advanced optimal control techniques (such as those based on the Discrete-time Algebraic Riccati Equation) and for designing reliable state observers. Without stabilizability, there could exist unstable dynamics that no controller can mitigate, and without detectability, unmeasured states might lead to undetected instabilities.

To rigorously analyze these properties in the context of the thermal system, it is crucial to bound the spectrum of the aggregate state matrix. A classical tool for this purpose is the Gershgorin disk Theorem, which locates all eigenvalues of a matrix using only its diagonal and off-diagonal entries. This theorem is particularly valuable for our RC-based thermal models, whose sparsity patterns allow us to derive tight eigenvalue bounds.

Theorem 1 (Gershgorin Disk Theorem [62])**.** Let $A = [a_{ij}]$ be a complex $n \times n$ matrix. For $i = 1, \ldots, n$, let R_i be the sum of the absolute values of the non-diagonal entries in the i-th row:

$$R_i = \sum_{j \neq i} |a_{ij}|$$

Let $D(a_{ii}, R_i)$ be the closed disk centered at a_{ii} with radius R_i:

$$D(a_{ii}, R_i) = \{z \in \mathbb{C} : |z - a_{ii}| \leq R_i\}$$

Then every eigenvalue of A lies within at least one of these Gershgorin disks $D(a_{ii}, R_i)$. In other words, if $\sigma(A)$ denotes the spectrum (set of all eigenvalues) of A, then

$$\sigma(A) \subseteq \bigcup_{i=1}^{n} D(a_{ii}, R_i)$$

To further strengthen the stability analysis, Taussky's Theorem is introduced. Although the Gershgorin disk Theorem provides general regions in which eigenvalues reside, Taussky's Theorem offers a more precise characterization by establishing that if an eigenvalue lies on the boundary of the union of these disks, then it must lie on the boundary of every individual Gershgorin disk. This insight is particularly valuable when considering the structured and irreducible aggregate state matrix resulting from the interconnected thermal paths in the building.

Theorem 2 (Taussky's Theorem on Gershgorin Disks [62])**.** Let $A = [a_{ij}]$ be an $n \times n$ complex matrix that is **irreducible** [1]. Let $\mathcal{G} = \bigcup_{i=1}^{n} D(a_{ii}, R_i)$ be the union of the Gershgorin disks for A, where $D(a_{ii}, R_i) = \{z \in \mathbb{C} : |z - a_{ii}| \leq R_i\}$ and $R_i = \sum_{j \neq i} |a_{ij}|$.

If λ is an eigenvalue of A that lies on the boundary of the region $\mathcal{G}$ (i.e., $\lambda \in \partial\mathcal{G}$), then λ must lie on the boundary of *for every* Gershgorin disk $D(a_{ii}, R_i)$. That is,

$$|\lambda - a_{ii}| = R_i \quad \text{for all } i = 1, \ldots, n$$

Lemma 1. The aggregate state matrix A for the multizone thermal system, derived from the Laplacian truncated scaled negative graph, is Hurwitz. Consequently, the system dynamics $\dot{x} = Ax + Bu + Ed$ is stabilizable and detectable.

Proof. The aggregate state matrix is given by $A = -C^{-1}L'$, where C is a diagonal matrix with positive thermal capacitances $C_i > 0$, and L' is the principal submatrix of the full graph Laplacian L_G corresponding to the dynamic state nodes (walls x_w and zones x_z). The matrix L' represents the thermal conductances between these state nodes. Since physical building models ensure thermal paths connect the state nodes, the graph G restricted to these nodes is connected.

First, it must be established that all eigenvalues of A have non-positive real parts using the Gershgorin Disk Theorem (Theorem 1). The diagonal entries of A are $A_{ii} =$

[1] A square matrix A is said to be *irreducible* if it is not possible to permute its rows and columns into a block upper triangular form. Equivalently, for every partition of the index set into two nonempty subsets, there exists at least one nonzero element connecting them; this is tantamount to saying that the directed graph associated with A is strongly connected.

$-(1/C_i)L'_{ii}$. Since $L'_{ii} \geq \sum_{j\neq i} |L'_{ij}|$ (as L' is diagonally dominant, being derived from a Laplacian structure), the diagonal entries L'_{ii} are non-negative, making $A_{ii} \leq 0$. The radius of the i-th Gershgorin disk for A is $R_i(A) = \sum_{j\neq i} |A_{ij}| = (1/C_i) \sum_{j\neq i} |L'_{ij}|$. The rightmost point of the i-th disk is $A_{ii} + R_i(A) = -(1/C_i)L'_{ii} + (1/C_i) \sum_{j\neq i} |L'_{ij}| = -(1/C_i)[L'_{ii} - \sum_{j\neq i} |L'_{ij}|]$. Since $L'_{ii} \geq \sum_{j\neq i} |L'_{ij}|$, this rightmost point is less than or equal to zero. Thus, all eigenvalues λ of A must satisfy $\mathrm{Re}(\lambda) \leq 0$.

Next, it must be shown that zero cannot be an eigenvalue. A Laplacian matrix (and its principal submatrix L') corresponding to a connected undirected graph is irreducible [67]. Since $A = -C^{-1}L'$ shares the same off-diagonal structure and C^{-1} is diagonal with positive entries, A is also irreducible.

Assume, for contradiction, that $\lambda = 0$ is an eigenvalue of A. Since $\mathrm{Re}(0) = 0$, it must lie on the boundary of the union of Gershgorin disks $\mathcal{G} = \bigcup D(A_{ii}, R_i(A))$. By Theorem 2 since A is irreducible, $\lambda = 0$ must lie on the boundary of *every* disk. This means $|0 - A_{ii}| = R_i(A)$ for all $i = 1, \ldots, n$. This condition $|A_{ii}| = R_i(A)$ implies $L'_{ii} = \sum_{j\neq i} |L'_{ij}|$ for all i. This equality holds if and only if node i has no connection to any external node (i.e., the i-th row sum of L' is zero). For this to hold for *all* i, the entire system represented by L' must be thermally isolated from the external environment.

However, the thermal model represents a building interacting with external temperatures (e.g., T_e, ground), meaning at least one state node i must have a thermal connection to an external node. For such a node k, the diagonal dominance is strict, $L'_{kk} > \sum_{j\neq k} |L'_{kj}|$, which means $|A_{kk}| > R_k(A)$. Therefore, the condition $|0 - A_{ii}| = R_i(A)$ cannot hold for all i. This contradiction proves that $\lambda = 0$ cannot be an eigenvalue of A.

Since all eigenvalues of A satisfy $\mathrm{Re}(\lambda) \leq 0$ and $\lambda \neq 0$, it follows that $\mathrm{Re}(\lambda) < 0$ for all eigenvalues. Thus, A is Hurwitz.

A system with a Hurwitz state matrix A is inherently stable. All unstable modes (those with $\mathrm{Re}(\lambda) \geq 0$) are non-existent, satisfying the conditions for stabilizability. Similarly, all unobservable modes must be stable (as all modes are stable), satisfying the conditions for detectability. □

4.6.6 Multi-zone Formulation Examples

This section provides two examples. The first is an example not shown yet in this paper and shows the relationship between modeling the system one node at a time to the augmented model. The second shows how the 5R5C model would be modeled in code.

Multi-zone Formulation Comparison

Fig. 4.7 shows the schematic of the thermal zones being modeled. The building consists of two rooms, each of which is connected to the external air. A variation of the 3R2C model for a wall is shown in the upper right corner. This modeled the wall with two lumps, one on each side. Each side of the wall touches a different zone. Note that one-indexing is used to number the nodes. The nodes N_1 to N_6 correspond to the wall nodes, N_7 and N_8 correspond to the zone nodes, and N_9 corresponds to the external air.

Fig. 4.8 shows the matrices that are used with the resistance and capacitance values shown in Fig. 4.7.

The temperature of zone 1 of Node 7 was derived from Kirkoff's current law.

$$C_{z1}\dot{x}_7 = \frac{1}{R_{w12}}(x_2 - x_7) + \frac{1}{R_{w21}}(x_4 - x_7) \tag{4.53}$$

$$\dot{x}_7 = \frac{1}{C_{z1}R_{w12}}x_2 + \frac{1}{C_{z1}R_{w21}}x_4 - \frac{1}{C_{z1}}(\frac{1}{R_{w12}} + \frac{1}{R_{w21}})x_7 \tag{4.54}$$

Node 7 zone 1 temperature derived from multi-zone modeling technique.

$$\dot{x}_7 = [-C_z L_{zw} T_w + -C_z L_{zz} T_z] \tag{4.55}$$

$$= -\frac{1}{C_{z1}}(-a_{72}x_2 + a_{73}x_3 + d77x_7) \tag{4.56}$$

$$= \frac{1}{C_{z1}R_{w12}}x_2 + \frac{1}{C_{z1}R_{w21}}x_4 - \frac{1}{C_{z1}}(\frac{1}{R_{w12}} + \frac{1}{R_{w21}})x_7 \tag{4.57}$$

Node 4 wall 2 temperature derived from Kirkoffs' current law.

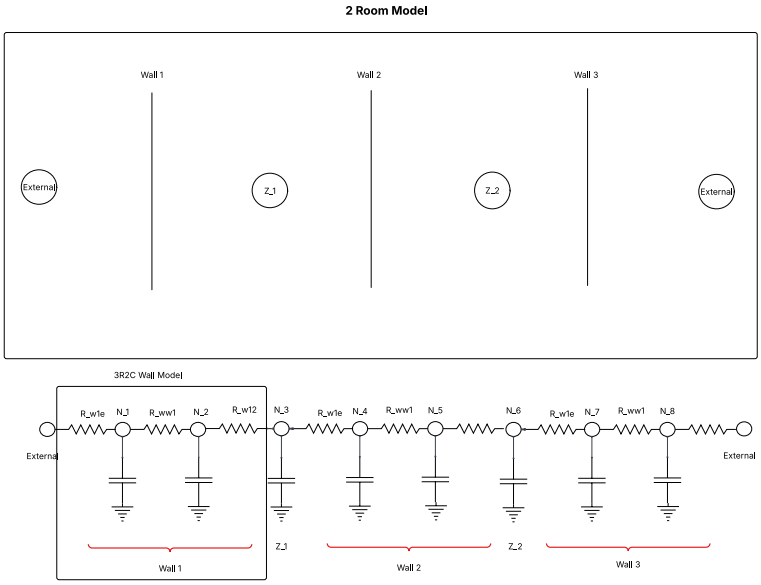

Fig. 4.7: Diagram of the system being modeled. Two zones and three walls.

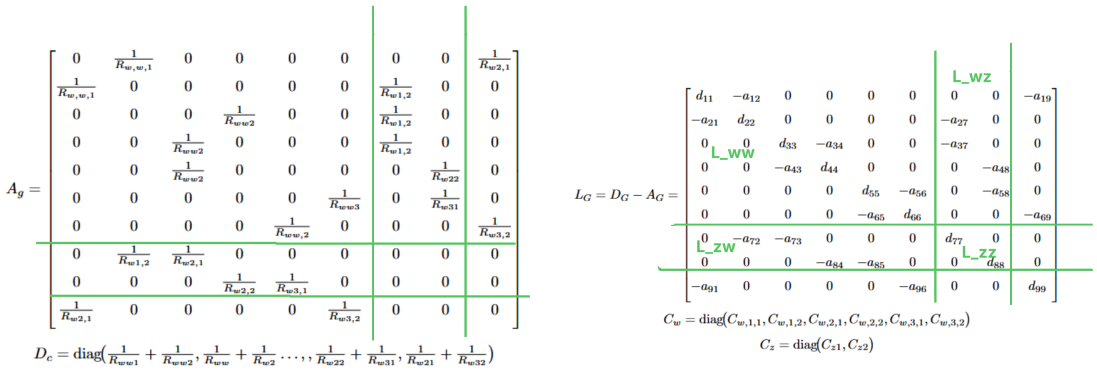

Fig. 4.8: Matrices for the graph

$$C_{w22}\dot{x}_4 = \frac{1}{R_{ww2}}(x_3 - x_4) - \frac{1}{R_{w22}}(x_8 - x_4) \tag{4.58}$$

$$\dot{x}_4 = \frac{1}{C_{w22}R_{ww2}}x_3 + \frac{1}{C_{w22}R_{w22}} - \frac{1}{C_{w22}}(\frac{1}{R_{w22}} + \frac{1}{R_{ww2}})x_4 \tag{4.59}$$

Temperature of wall 2 of node 4 derived from the multizone modeling technique.

$$\dot{x} = \left[-C_w^{-1}L_{ww}T_w - C_w^{-1}L_{wz}T_z + -C_w^{-1}L_e + T_e\right] \tag{4.60}$$

$$= -\frac{1}{C_{w222}}\left[-a_{43}x_3 + d_{44}x_4 - a_{48}x_8 + 0\right] \tag{4.61}$$

$$= -\frac{1}{C_{w222}}\left[-\frac{1}{R_{ww2}}x_3 + (\frac{1}{R_{ww2}} + R_{w_{12}})x_4 - \frac{1}{R_{w22}}x_8\right] \tag{4.62}$$

$$\dot{x}_4 = \frac{1}{C_{w22}R_{ww2}}x_3 + \frac{1}{C_{w22}R_{w22}} - \frac{1}{C_{w22}}(\frac{1}{R_{w22}} + \frac{1}{R_{ww2}})x_4 \tag{4.63}$$

4.7 A Multi-zone / Multi-floor Example

This section presents an extensive example with eight thermal zones on two floors. The resulting model has 50 state variables and 8 control inputs. The model is first discussed, followed by example results.

4.7.1 Two Floor / Eight Zone Model

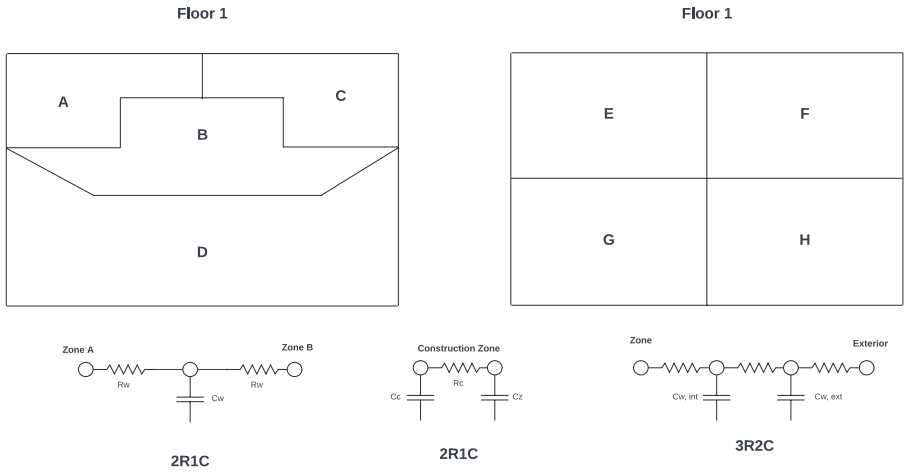

Fig. 4.9: (Top) Shows the floor plan with zones identified. (Bottom) Shows the 2R1C used to connect internal zones, the 1R2C used to connect a zone to internal construction, and the 3R2C used to connect internal zones to the exterior temperatures.

Fig. 4.9 shows the two-story building with eight zones, identified by $A, B, ..., H$ as well as the 2R1C, 1R2C, and 3R2C models used for the various connections. Each thermal zone is controlled directly and has an associated thermal capacitance.

An uncontrolled internal zone inside each thermal zone is connected via the 1R2C model to the "Internal Construction" zone used to model the capacitance of the internal furniture and walls. The thermal zone $(A, B, ..., H)$ forms once side of the resistor R_c, of the 1R2C model with capacitance C_z and the internal construction forms the other side with capacitance C_c. The internal construction results in 8 additional states, one for each thermal zone.

Each thermal zone is connected to other thermal zones through walls, ceilings, and floors via the 2R1C model, where the capacitor is used to model the thermal capacitance

of the wall or ceiling/floor. The connections are

$$\begin{array}{llll} A \leftrightarrow B & B \leftrightarrow D & B \leftrightarrow H & E \leftrightarrow F \\ A \leftrightarrow C & B \leftrightarrow E & C \leftrightarrow F & E \leftrightarrow G \\ A \leftrightarrow E & B \leftrightarrow F & D \leftrightarrow G & F \leftrightarrow H \\ B \leftrightarrow C & B \leftrightarrow G & D \leftrightarrow H & G \leftrightarrow H \end{array},$$

as well as a connection between the zones A, B, C, D and the ground, which is an external temperature. There is a zone associated with walls or ceiling / floor with capacitance C_w connected to each zone through a resistance R_w. Given the 16 internal zone-zone connections and the 4 ground connections, this results in an additional 20 states.

All zones, except B, have a connection with the outside temperature, another external zone. To model the fact that the internal side of the wall and the external side of the wall have very different temperatures, the 3R2C model is used to connect the zone with the external temperature. The capacitance closest to the internal thermal zone is labeled $C_{w,in}$ and that closest to the outside zone is labeled $C_{w,ext}$. Given seven connections to the outside with two zones in each wall, this results in 14 additional states.

In general, the model consists of 50 states, 8 inputs, and 2 disturbances (outer- and ground temperatures).

4.7.2 Optimization with Power Considerations

The problem is very similar to the recursive formulation in (4.29) with augmented form in (4.30). A significant change is made to (4.29) to allow for the inclusion of power and peak power considerations.

Given m control inputs at each time instant, k, the power consumed at that time instant is calculated as[2]

$$p_k = \sum_{l=0}^{m} u_{k,l}. \tag{4.64}$$

[2]Note that because the power is applied over a time interval, p_k actually corresponds to the average power used throughout that interval.

The peak power must be greater than or equal to any of the individual power values. Defining p as the vector of all power values and $\mathbb{1}$ to denote the a vector of ones, the peak power can be defined as

$$\mathbb{1}p_{peak} \geq p. \tag{4.65}$$

Although strictly speaking, the maximum power should be equal to the maximum power value, if a term is added to minimize the maximum power, then the optimization will for p_{peak} be reduced to its smallest value, the maximum element of p.

Defining the vector $y_o = [p_0; p_1; p_{N-1}; p_{peak}]$, the optimization problem to be solved can be written as

$$\begin{aligned}
\min_{x_k,u_k} & \left[\frac{1}{2}u_{e,0}^T R u_{e,0} + \frac{1}{2}\sum_{k=1}^{N-1}\left(u_{e,k}^T R u_{e,k} + x_{e,k}^T Q x_{e,k} + p_k^T O p_k\right)\right. \\
& \left. + \frac{1}{2}x_{e,N}^T P x_{e,N} + p_0^T O p_0 + c_o^T y_o\right] \\
\text{s.t. } & x_{x+1} = \bar{A}x_k + \bar{B}u_k + \bar{E}d_k \\
& x_{min,k} \leq x_k \leq x_{max,k} \quad k = 1, ..., N \\
& u_{min,k} \leq u_k \leq u_{max,k} \quad k = 0, ..., N-1 \\
& p_k = \sum_{l=0}^{m} u_{k,l} \quad k = 0, ..., N-1 \\
& \mathbb{1}p_{peak} \geq p
\end{aligned} \tag{4.66}$$

where Q, R, O and P are weighting matrices, c_o is a weighting vector, $\leq$ is interpreted element-wise, and the variables of optimization consist of $x_1, x_2, ..., x_N$, $u_0, u_1, ..., u_{N-1}$, $p_0, p_1, ..., p_{N-1}$, and p_{peak}. An additional difference between (4.66) and (4.29) is that the control and state limits are defined at each time interval. This allows different temperature zones to be used throughout the day.

The minimization in (4.66) can be transformed into an augmented form in a fashion similar to the creation of (4.30).

4.8 Results

4.8.1 Verification of Controllability and Observability

To ensure that our discrete-time 5R5C thermal model is suitable for the state-space controller design, we verify that the pair $(\bar{A}, \bar{B})$ is controllable and that the pair $(\bar{A}, C)$ is completely observable, where C is chosen to extract the internal temperature T_i. For a discrete-time system

$$x_{k+1} = \bar{A}x_k + \bar{B}u_k,$$

the controllability matrix is defined as

$$\mathcal{C} = \begin{bmatrix} \bar{B} & \bar{A}\bar{B} & \bar{A}^2\bar{B} & \cdots & \bar{A}^{n-1}\bar{B} \end{bmatrix},$$

and the system is controllable if $\text{rank}(\mathcal{C}) = n$, where n is the number of states (in our case, $n = 5$).

Similarly, for an output matrix

$$C = \begin{bmatrix} 0 & \cdots & 1 & \cdots & 0 \end{bmatrix},$$

which extracts the state corresponding to the controlled internal temperature, the observability matrix is given by

$$\mathcal{O} = \begin{bmatrix} C \\ C\bar{A} \\ C\bar{A}^2 \\ \vdots \\ C\bar{A}^{n-1} \end{bmatrix}.$$

The pair $(\bar{A}, C)$ is completely observable if $\text{rank}(\mathcal{O}) = n$.

In our simulation code, we compute and print these matrices using the `python-control` package. Listing 4.1 shows the relevant code snippet:

```
    tc = R5C5()
    Q_ws_0 = u_ff.item(tc.iu_Q_ws)  # 7 #  70.

    # Define storage
    t_vec = np.arange(start=0.0, stop=t_stop, step=time_step)
    n_times = t_vec.shape[0]
    x_mat = np.zeros((5, n_times))  # State
    u_mat = np.zeros((2, n_times))  # Input
    d_mat = np.zeros((2, n_times))  # Disturbance
    e_mat = np.zeros((2, n_times))  # Energy for each input
    e_tot_mat = np.zeros((n_times,))  # Total energy used over time
    mode_vec = np.zeros((n_times,))

    # Initialize for first step
    x_mat[:, [0]] = x0
    mode_vec[0] = mode_start  # start with resting

    # calculate discrete time matrices
    Abar, Bbar, Ebar = tc.discrete_matrices(dt=time_step)

    # Calculate the discrete-time control
    Gamma = ctl.ctrb(A=Abar, B=Bbar)
    print("Rank of controllability grammian: ", np.linalg.
matrix_rank(Gamma))

    Cbar = np.zeros((1, 5))
    Cbar[0, tc.ix_T_i] = 1.0

    Omega = ctl.obsv(A=Abar, C=Cbar)
    print("Rank of observability grammian: ", np.linalg.matrix_rank(
Omega))
```

Listing 4.1: Computation of Controllability and Observability Matrices

The printed output confirms that the controllability matrix $\mathcal{C}$ and the observability matrix $\mathcal{O}$ both have full rank (that is, rank 5). This result verifies that every state in the model is both controllable via the available inputs and observable through our measurement of the indoor temperature. Consequently, these conditions justify the subsequent use of state-space feedback design (e.g., via the Discrete-time Algebraic Riccati Equation) for controller synthesis.

4.8.2 Single-zone Control Observations

Several examples were run to show the utility of the proposed approach in keeping the temperature at 20 degrees Celsius. The first approach was a model-free threshold approach. The HVAC system commanded a constant value for the heating of the water storage and used the FCU to adjust the temperature. The FCU would turn on at 19 degrees and turn off at 21 degrees. The second control approach was a feedback control approach without an integral. The third added the integral, and the fourth added an optimal trajectory within the integral controller.

Table 4.2: Results with heating only

Approach	ΔC	kJ
Threshold	2.0	9000
FB-Cont	1.2	1000
Int-Cont	0.2	1000
Opt-Int-Cont	0.2	>1000

Three different scenarios were run. The first scenario was heating only (that is, no cooling allowed). For optimization techniques, this amounts to setting the minimum input

Table 4.3: Results with heating and cooling

Approach	ΔC	kJ
Threshold	7.5	1500
FB-Cont	1.5	1000
Int-Cont	0.01	6000
Opt-Int-Cont	0.2	900

Table 4.4: Optimal - Open and Closed Loop

Approach	ΔC	kJ
Open	0.8	1000
Closed	0.2	900

as 0. For feedback control techniques, this amounts to saturating the command control to less than zero. The second scenario was to allow for both heating and cooling. The final scenario was a comparison between open- and closed-loop design with the optimal trajectory.

In the first two scenarios, the external temperature starts at -10 degrees C and increases to 10 degrees around midday, and then returns to -10 degrees by the end of the day. In the third scenario, the optimizer assumes that same variation, but the temperature actually rises to 12 degrees C. In all scenarios, the controller is designed around the nominal temperature of -10 degrees.

Tables 4.2 and 4.3 and Fig 4.10 show the results of the first two scenarios. The threshold controller takes a significant amount of energy, especially in the heating-only scenario. The integrator adds a significant amount of fidelity to the desired temperature, reducing the error to 1/6 of the value under heating only and just a fraction of the error when both heating and cooling are allowed.

However, the integrator control effort is difficult to predict. To keep the error low in the heating and cooling scenarios, the integrator uses heat on one input and cooling on the other. Although this may produce a result with very little error, it uses a significant amount of energy. It is difficult to determine exactly how the integrator will affect the control effort.

On the other hand, the optimization-based controller has the ability to directly specify the allowable bounds. The Min / Max value for the optimization-based controller was actually chosen based on the error in the integrator for the heat-only scenario. Allowing more control flexibility (i.e., adding in cooling) allows the optimization-based controller to reduce its total effort, instead of the wild increase seen in the integrator controller. It is able to balance fidelity with energy usage regardless of the control freedom.

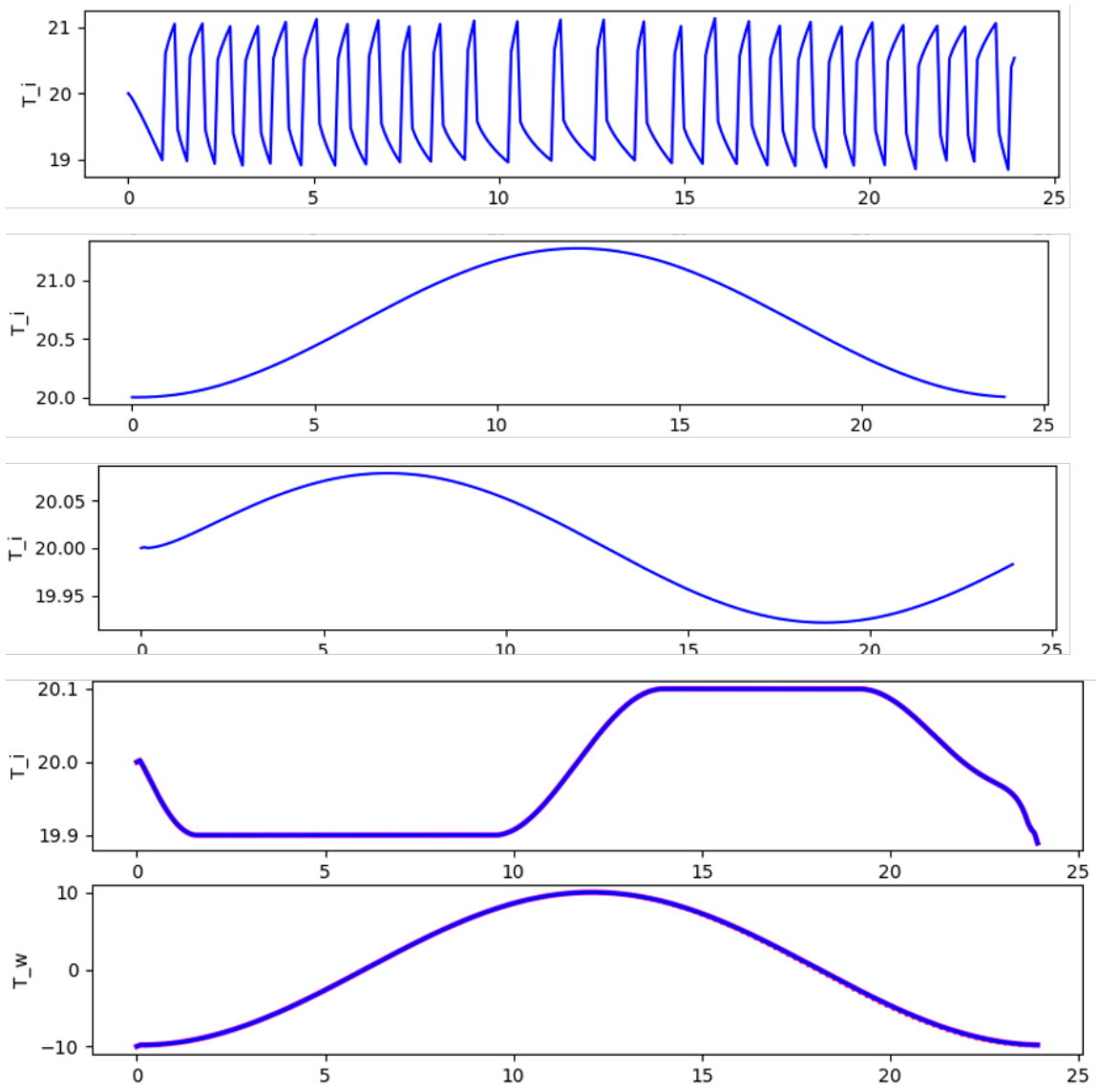

Fig. 4.10: Temperature profiles. The top four profiles are the result of the inside temperature for the different control strategies. From top-to-bottom: threshold, feedback control, state-space integral control, optimization-based integral control. The bottom figure shows the outside wall temperature, which lags slightly behind the external temperature.

An interesting observation can be made by comparing the temperature profiles shown in Fig. 4.10. Each control scheme has a very different profile. The shape of the pure feedback control almost mimics that of the external temperature. The effect of the integrator can be seen in the integrator profile. As the temperature of the day warms up, the integrator builds a value to then begin lowering the internal temperature back to the desired before the final peak external temperature. It still has a very reactive nature to it. However, the optimization-based controller has a much more deliberative profile. The internal temperature is allowed to drop to the lower threshold at first instead of naturally rising with the external temperature. Then, as the external temperature reaches its peak, the internal temperature is raised to the highest allowable threshold and stays there until it is allowed to drop right at the end of the day.

In the final scenario, a test was made to compare open- and closed-loop. The results can be seen in Table 4.4 and Fig. 4.11. The closed-loop controller can take advantage of the increased external temperature, which warms the internal temperature some during the middle of the day. It is therefore able to keep within the min and max bounds imposed by the optimization while using less energy. Fig. 4.11 shows that the actual temperature rises much higher than the desired temperature because the heating does not take into account the warmer external temperatures. Had it been a colder day than expected, it would have been possible to maintain the temperature settings, albeit with the use of more energy.

Fig. 4.12 shows the energy profiles for the open- and closed-loop control results. Note that the closed-loop effectively uses the integrator value to "pull down" the energy being put into both the water storage and the FCU. This allows the integrator value to converge back to zero and enables tracking of the desired temperature.

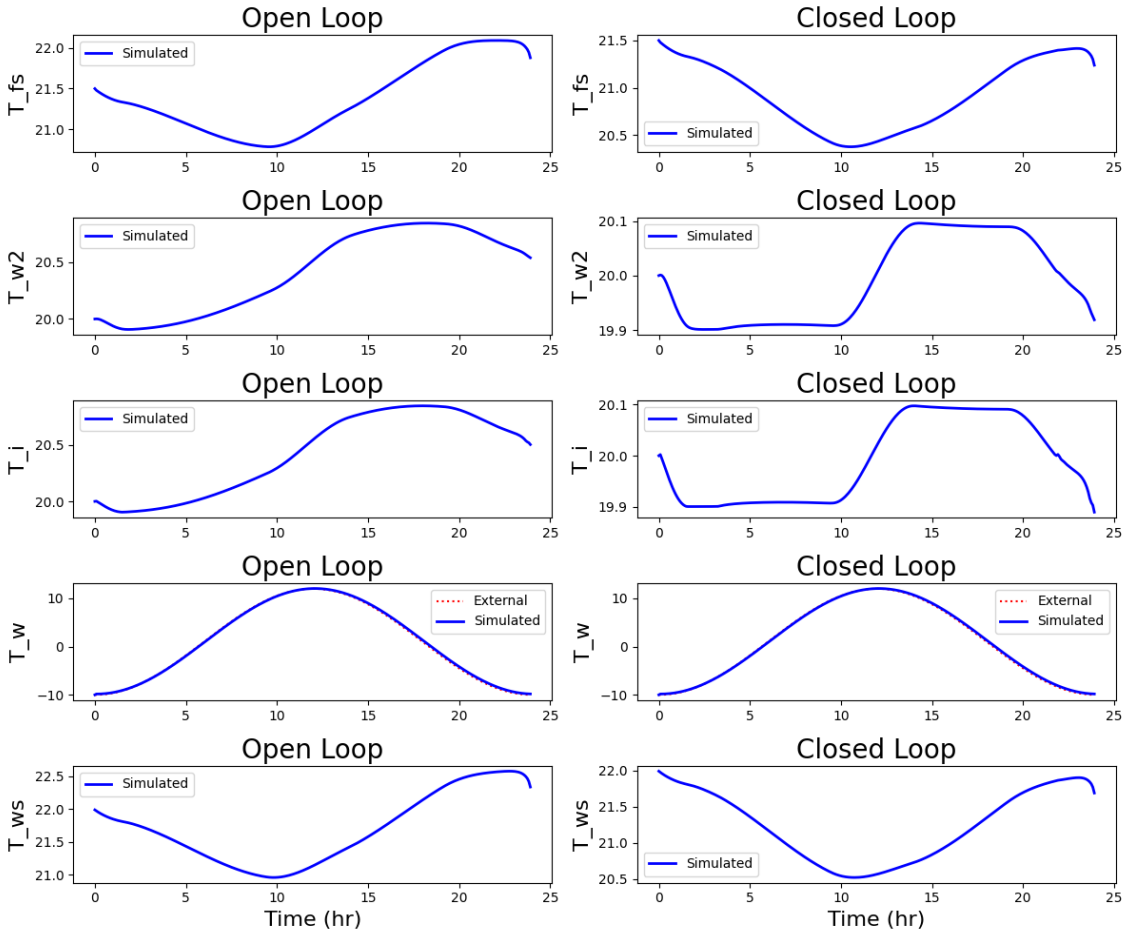

Fig. 4.11: Left shows the open-loop optimal control results, right shows the closed-loop optimal control results. From top to bottom of each side is the floor slab temperature, the internal wall temperature, the internal temperature, the exterior wall and external temperature, and the water storage temperature. The blue line shows the realized and the magenta line shows the desired.

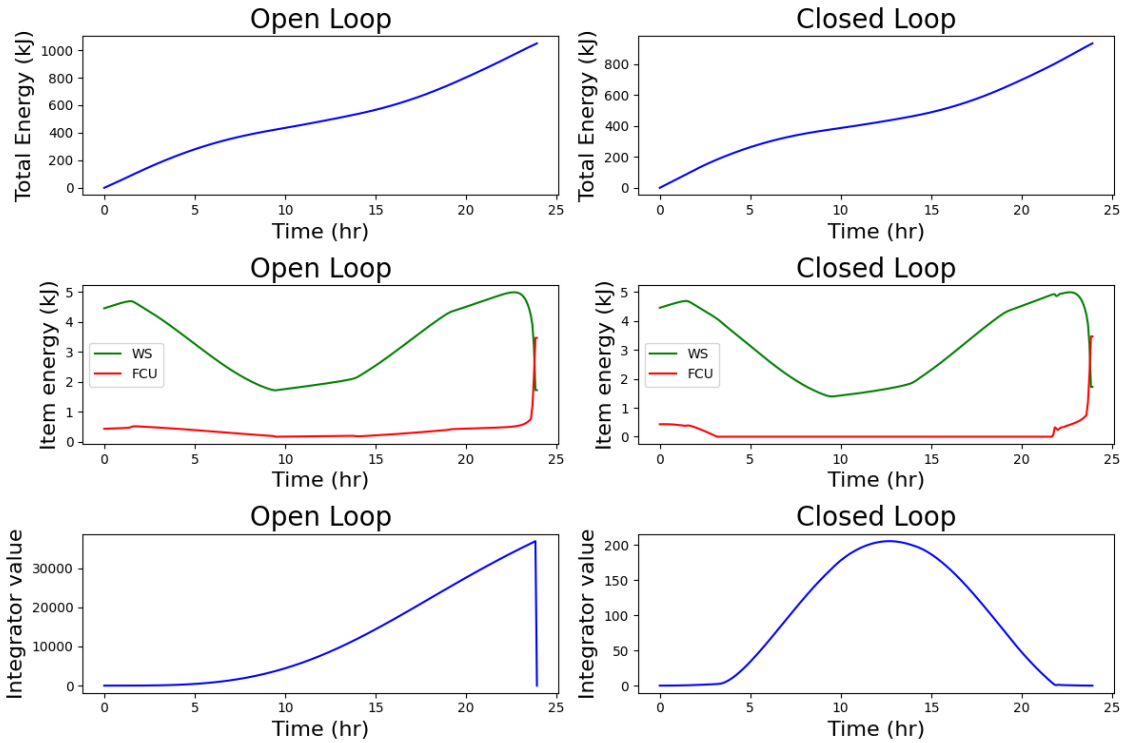

Fig. 4.12: Left shows the open-loop optimal control results, right shows the closed-loop optimal control results. From top to bottom of each side is the total energy used, the energy from water storage (green) and FCU (red), and the value of the integrator over time.

4.8.3 Multi Zone Results

Two scenarios are presented, one without peak power considered and the other with peak power. For each case, the optimization used a discretization of 15 minutes to develop a desired control and zone temperature trajectory. The resulting control and zone temperature trajectories are interpolated and passed into the model-free threshold, hysteresis control approach from Section 4.8.2 where the optimized zone temperatures are used as the desired temperature in the control. The hysteresis control evaluates the zone temperature in a 1-minute period.

No Peak Power Consideration

The first example does not consider the peak power minimization. Fig. 4.13 shows the power and energy usage planned by the optimizer, as well as the results of executing the controller. Fig. 4.14 shows the resulting temperature profiles for each of the zones.

Several observations are in order. First, note that the hysteresis controller is able to track the desired temperature fairly well, oscillating as can be expected around the desired temperature. The resulting energy consumed is nearly equivalent, just slightly higher for the hysteresis controller. Although energy and temperature are tracked fairly well, power usage is not. There are similar valleys in each, but the peak power usage occurs in different locations. In fact, the peak power usage in the hysteresis controller is the highest possible – there are times when every single zone is being heated simultaneously.

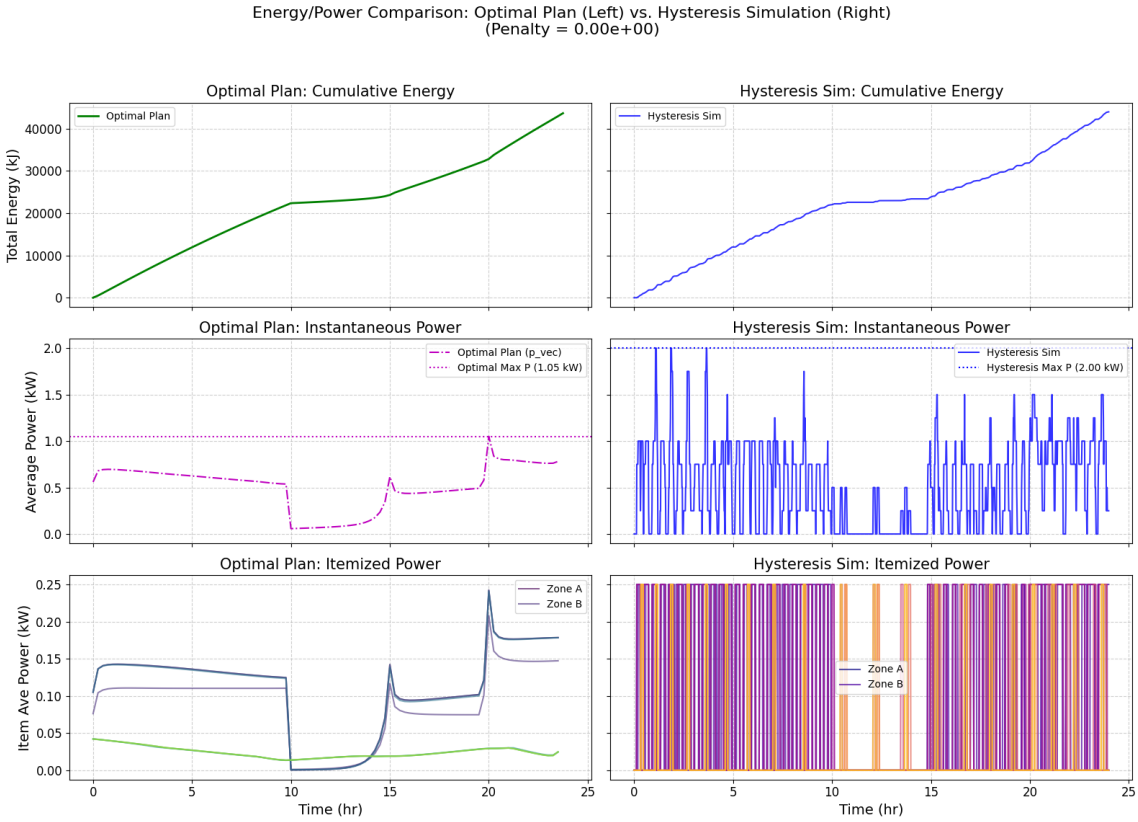

Fig. 4.13: Example not considering peak power minimization. Left is the desired result of the optimization, right shows the result using the hysteresis controller. Top shows energy consumption over time, middle shows the average power usage at each time step and max power used thus far, and the bottom shows the power used by each of the thermal zones.

Table 4.5 shows a comparison between a baseline constant temperature (set at 20 degrees C), The Optimal Plan, and the Optimal Plan limited by a thermostat controller.

Table 4.5 shows that significant energy savings are seen even when limited by thermostat controllers.

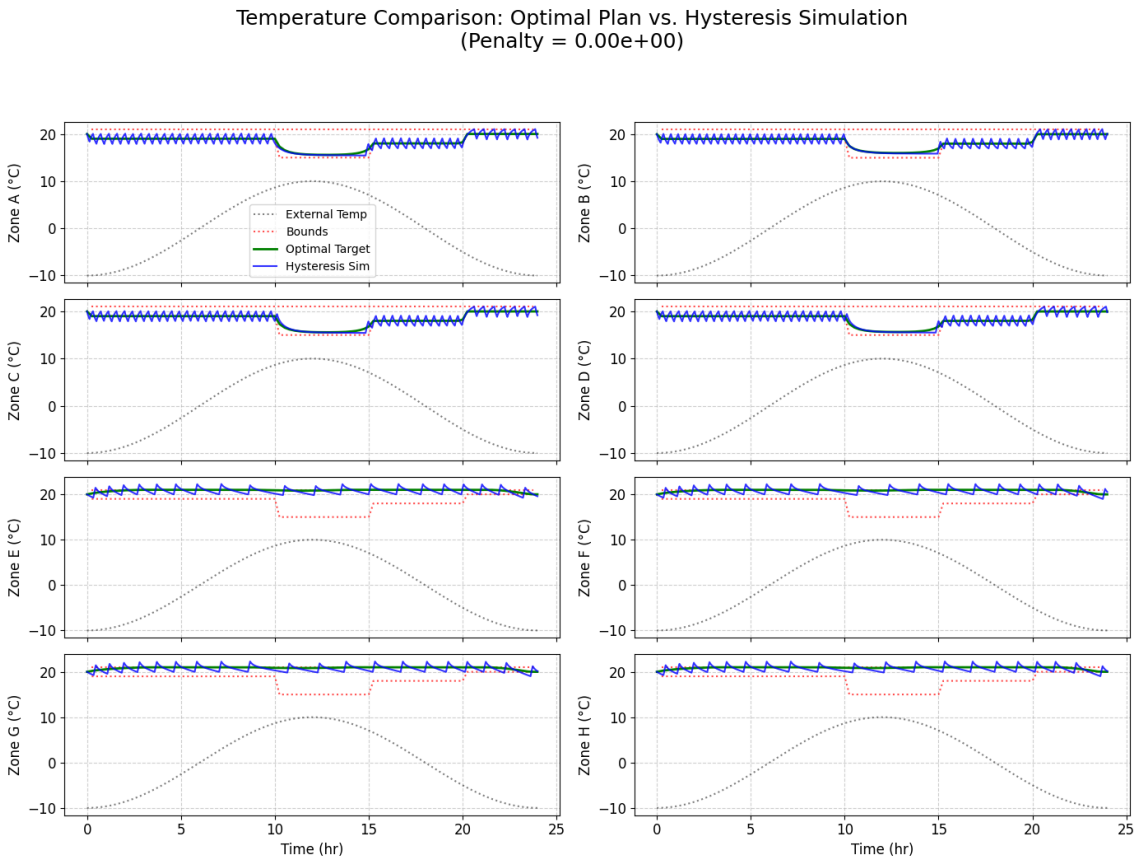

Fig. 4.14: Example not considering peak power minimization. The temperatures for each zone are depicted. Green shows the desired temperature from the optimizer, blue shows the resulting temperature from the controller, and red shows the temperature bounds for the zone as well as the external temperature.

Peak Power Considered

The second example considers the minimization of peak power with the results shown in Figs. 4.15 and 4.16. The optimizer is able to plan for a reduced peak power, nearly a 20% reduction over the scenario where peak power is not considered. This is achieved by warming slightly earlier from the lowest temperatures in zones A through D. This can be seen by comparing Fig. 4.14 and 4.16.

However, a similar trend occurs as before. Although temperature and energy are tracked fairly well, power usage is not. Therefore, while the optimizer can reduce the maximum power usage, the hysteresis controller still executes a profile with the highest

Key Performance Indicator (KPI)	**Baseline Constant Temperature Thermostat**	**Optimal Control Plan**	**Optimal Control with Thermostat Controller**
Peak Power (kW)	2.0	1.05	2.00
Total Energy (kJ)	64700	45000	49000
Temperature Tracking	Follows Constant Temperature with jittering	Target	Tracks target well with expected hysteresis jitter
Power Profile Adherence	Peaks the Power Almost Every time it Turns on	Target	Power profile is very non smooth and jittery with many peaks above optimal plan

Table 4.5: Comparison of Control Strategies Based on Key Performance Indicators Not Considering Peak Power

possible peak power.

Additional work will be needed to ensure that the peak power is appropriately addressed.

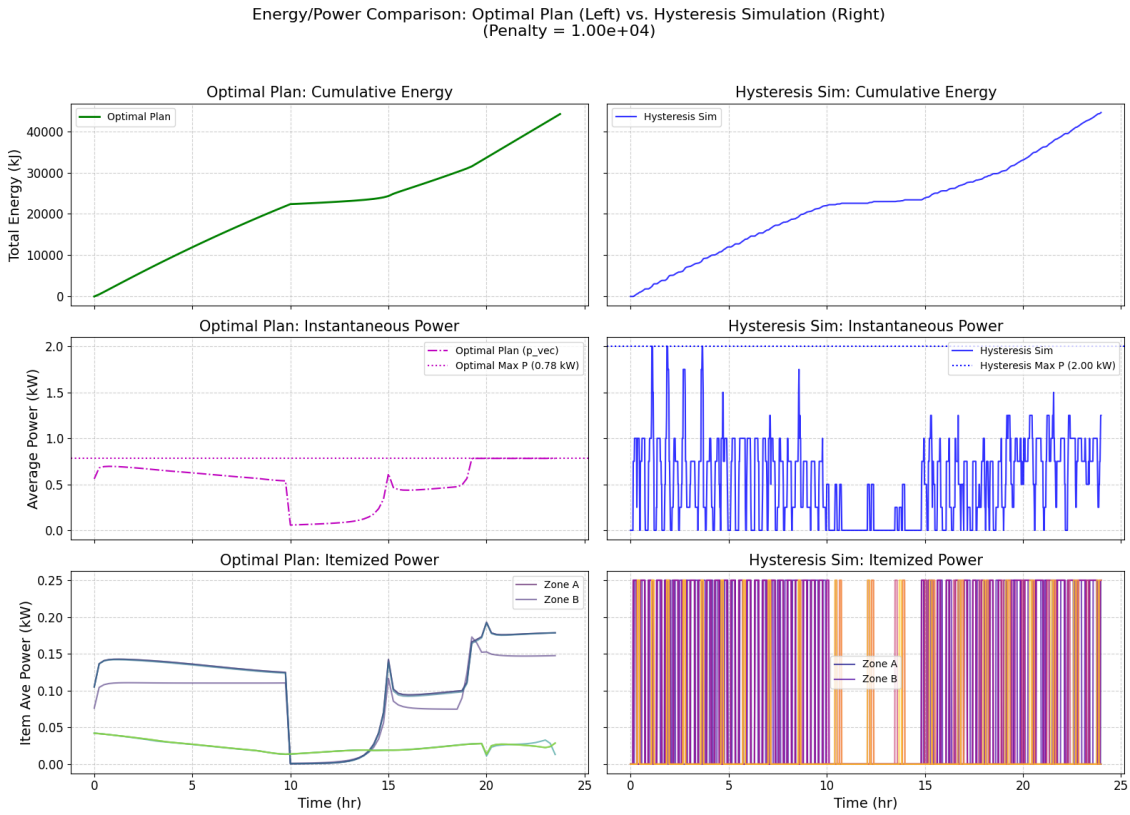

Fig. 4.15: Example considering peak power minimization. Left is the desired result of the optimization, right shows the result using the hysteresis controller. Top shows energy consumption over time, middle shows the average power usage at each time step and max power used thus far, and the bottom shows the power used by each of the thermal zones.

Table 4.6 shows the same performance indicators and control strategies as Table 4.5 but investagating if peak power can be shaved

From Table 4.6 shows that peak power is not able to be limited even when directly targeting these objectives. It does peak the system less often but overall peak shaving is not able to be achieved when limited by current thermostat controllers.

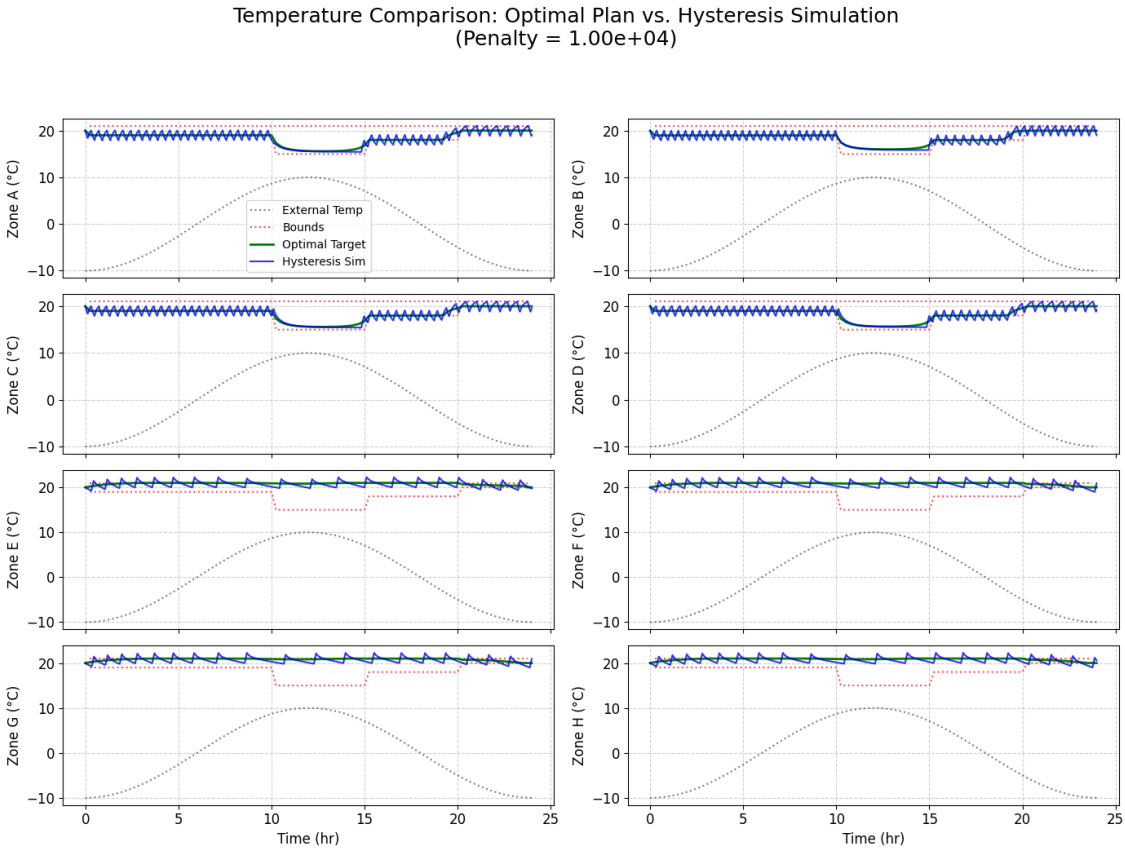

Fig. 4.16: Example considering peak power minimization. The temperatures for each zone are depicted. Green shows the desired temperature from the optimizer, blue shows the resulting temperature from the controller, and red shows the temperature bounds for the zone as well as the external temperature.

Key Performance Indicator (KPI)	**Baseline Constant Temperature Thermostat**	**Optimal Control Plan**	**Optimal Control with Thermostat Controller**
Peak Power (kW)	2.0	0.78	2.00
Total Energy (kJ)	64700	45000	49000
Temperature Tracking	Follows Constant Temperature with jittering	Target	Tracks Desired fairly well with expected hysteresis jitter
Power Profile Adherence	Peaks the Power Almost Every time it turns on	Target	Power profile is very non smooth and jittery with less peaks above optimal plan compared to not considering peak power

Table 4.6: Comparison of Control Strategies for Thermostat Systems Considering Peak Power

4.9 Conclusion

This work successfully demonstrated the application of lumped-sum RC modeling for building thermal control, including the development of a generalized graph-based formulation for multizone systems. A key outcome was the rigorous proof of the inherent stability of the resulting system models (Lemma 1). The investigation of feedback and optimal control strategies confirmed the utility of model-based approaches, with the optimal control framework effectively incorporating peak power constraints. However, simulations revealed challenges in translating optimized power trajectories into practice using simple hysteresis controllers, highlighting a significant area for future work. Further research should focus on developing advanced tracking controllers compatible with optimization objectives, along with exploring adaptive modeling techniques and experimental validation of the presented methods. The framework established provides a solid foundation for future advancements in energy-efficient building control.

CHAPTER 5

Regression HVAC Control

5.1 Introduction

The increasing integration of renewable energy sources and the proliferation of Distributed Energy Resources (DERs) are transforming modern power systems [7]. Microgrids, localized electrical systems that integrate DERs such as photovoltaic (PV) panels, energy storage systems (ESS), electric vehicle (EV) chargers and controllable loads are emerging as key enablers of this transition. They offer enhanced energy resilience, flexibility, and the potential for optimized local energy consumption, operating autonomously or in coordination with the main grid [68]. Among controllable loads, Heating, Ventilation and Air Conditioning (HVAC) systems represent a significant portion of the energy demand in buildings and offer substantial flexibility for demand-side management.

Effective microgrid management requires sophisticated strategies that coordinate diverse DERs while balancing multiple objectives, such as cost minimization, peak load reduction, and grid support. Although comprehensive frameworks for microgrid energy management have been proposed, often addressing forecasting, scheduling, and control [10], the practical integration and control of complex thermal loads, such as HVAC systems, remain a challenge. Traditional physics-based (white box) models can be accurate, but require detailed building information that is often unavailable [69], while purely data-driven (black box) approaches offer adaptability but must be carefully formulated for control purposes [39, 40].

This chapter focuses on addressing the challenge of HVAC control within a microgrid context by developing a regression-based optimization approach. Specifically, this chapter proposes a method for determining optimal temperature set point trajectories for existing HVAC systems. The core of this approach is a black-box regression model, trained

on simulation data, that captures the relationship between temperature setpoints, external conditions, and energy consumption. This model is then embedded within an optimization problem designed to balance potentially competing goals: minimizing HVAC energy consumption, reducing peak energy demand (average power over discrete intervals) and maintaining occupant comfort by limiting deviations from nominal desired temperatures.

The primary contribution of this work is the formulation and evaluation of a regression-based HVAC control strategy that controls standard temperature setpoints, making it compatible with existing building infrastructure. This work develops and evaluates a novel regression-based HVAC optimization approach designed to improve microgrid energy management flexibility, operational efficiency, and occupant comfort. The data-driven nature ensures adaptability without requiring complex physical models. Although developed as a component intended for integration into a larger hierarchical microgrid optimization framework, this chapter specifically details the development of the HVAC model, its associated optimization problem, and evaluates its performance independently using simulation. This provides a foundational element for more comprehensive microgrid energy management systems that aim to take advantage of the flexibility of the HVAC efficiently.

The remainder of this chapter is structured as follows. First, Table 5.1 provides an overview of terminology used in this chapter. Next, Section 5.2 details the development of the regression-based HVAC control model and formulates the associated optimization problem. Section 5.3 describes the simulation environment and the data used for model training. Section 5.4 presents the simulation results that evaluate the performance of the proposed HVAC optimization approach under different objective weightings. Finally, Section 5.5 concludes the chapter and discusses potential future work.

5.2 A Regression-based HVAC Optimization Approach

This section will define the dynamics and control of the HVAC system so that it can be incorporated into the total optimization process. The buildings of interest for this research already have thermostats and HVAC systems in place. The goal is to develop a control system that can be implemented on top of the existing system. This section will develop

Regression Control Problem Terminology	
A_E	Energy State Matrix
B_E	Energy Control Matrix
C_E	Energy Output State Matrix
A_T	Temperature State Matrix
B_T	Temperature Control Matrix
C_T	Temperature Output State Matrix
x_k	Aggregate state consisting of energy and desired temperature state
$u_{T,k}$	Input to the temperature system at time k
g_k	Energy gain of the system from time step $k-1$ to k
e_k	Total energy consumed from time 0 to time k
g_{max}	Maximum energy gain overall time
e_{tot}	Total energy consumed overall time, equivalent to e_N
$T_{d,k}$	Desired temperature at time k
$T_{n,k}$	The nominal set-point temperature for each zone.
w_p	Scalar Weighting term for maximum power
w_T	Scalar Weighting term for temperature deviation

Table 5.1: Chapter 5 Problem Terminology

a regression-based optimization approach that can be used to define a control system that can be implemented on top of the existing system, i.e., controlling temperature setpoints. The regression component will utilize data collected from real-life data or simulations.

The regression model approach aims to provide an informed trajectory of temperature set points for the HVAC system. An optimization problem is defined and solved to determine the desired temperature set-points that balance overall energy consumption, average power usage in any given interval, and deviation from the desired nominal temperature.

5.2.1 Developing the Control Model

A primary element in optimization is modeling energy usage over time, given the desired temperature setpoints and the external temperature. The dynamic model combines two components, as shown in Figure 5.1. The first component is the desired temperature model, which can range from a simple pass-through of the desired temperature to a more complex model incorporating derivatives of the desired temperature. The second component is the regression-based energy model, which predicts energy consumption based on the desired temperature and historical external temperature data. The primary goal of optimization

will be to determine $u_{T,k}$ to balance energy consumption with deviations from the desired temperature, average power usage, and peak power usage.

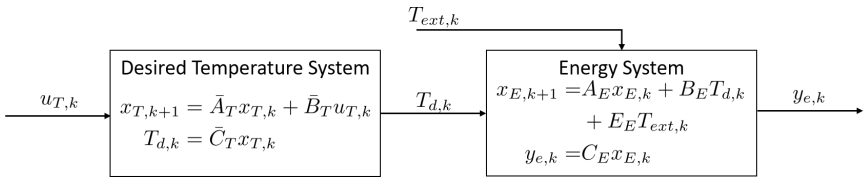

Fig. 5.1: Shows the interaction between the desired temperature model and the regression-based energy model where $u_{T,k}$ is a chosen input at time k and $T_{ext,k}$ contains the expected external temperature(s).

The desired temperature model is used first to create the setpoint for the temperature within the energy model. The details of the regression-based energy model are then provided, followed by the statement of the aggregate system.

Desired Temperature Dynamics

The desired temperature model consists of a state, $x_T \in \mathbb{R}^{n_{xT}}$, input, $u_T \in \mathbb{R}^{n_T}$, and the vector of desired temperatures for the system, $T_d \in \mathbb{R}^{n_T}$, where n_{xT} is the number of states for the system and n_T is the number of temperature zones. A linear model is used to represent the dynamics of the system as

$$\begin{aligned} \dot{x}_T &= A_T x_T + B_T u_T \\ T_d &= C_T x_T \end{aligned}. \tag{5.1}$$

Any form of (A_T, B_T, C_T) can be used to model the desired temperature dynamics. However, it is important to recall that this model is purely used to define the desired temperature trajectory for the energy model, i.e., it is used to define the input to the energy model. In this work, a multiintegrator model is used to model the temperature dynamics. This allows for constraints to be placed on the time evolution of the desired temperature, from constraints on the temperature itself to constraints on the rate of change of the temperature, the acceleration of the temperature, and so forth.

Assuming that there is a desired constraint on the m^{th} derivative of T_d, the temperature state can be expresses as

$$x_T = \left[T_d; \quad \dot{T}_d; \quad ...; \quad T_d^{(m-1)} \right]. \tag{5.2}$$

Allowing I_n be the $n \times n$ identity matrix, the dynamics of the desired temperature model can be expressed as

$$\begin{aligned} \dot{x}_T &= \begin{bmatrix} 0 & I_{n_T} & 0 & ... & 0 \\ 0 & 0 & I_{n_T} & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \ddots & I_{n_T} \\ 0 & 0 & 0 & ... & 0 \end{bmatrix} x_T + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ I_{n_T} \end{bmatrix} u_T \\ &= A_T x_T + B_T u_T \\ T_d &= \begin{bmatrix} I_{n_T} & 0 & ... & 0 \end{bmatrix} x_T, \\ &= C_T x_T \end{aligned} \tag{5.3}$$

where each 0 in (5.3) is a $n_T \times n_T$ matrix of zeros. The constraints for T_d and its $m-1$ derivatives can be set on x_T, with the constraints on T_d^m corresponding to the constraints on u_T.

A discrete-time model is obtained by performing a zero-order hold and discretizing the continuous-time model via the discretization process described in section 6.4 in [64]. This discretization process is done on a continuous time system of the form $\dot{x} = Ax + Bu$ is discretized as $x_{k+1} = \bar{A}x_k + \bar{B}u_k$ over the time interval T, where

$$\bar{A} = e^{AT}, \; \bar{B} = \int_0^T e^{A(T-\tau)} d\tau B, \tag{5.4}$$

and e^A represents the matrix exponential.

The discrete model can be expressed with the overbar notation as

$$\begin{aligned} x_{T,k+1} &= \bar{A}_T x_{T,k} + \bar{B}_T u_{T,k} \\ T_{d,k} &= \bar{C}_T x_{T,k}. \end{aligned} \tag{5.5}$$

Forming a Regression Model

Similarly, the regression-based energy model consists of a state, $x_E \in \mathbb{R}^{n_{xE}}$, external temperature disturbance, $T_{ext} \in \mathbb{R}^{n_{ext}}$, and energy consumption, $e \in \mathbb{R}$, where n_{xE} is the number of energy states and n_{ext} is the number of external temperature values affecting the system. The model takes the form

$$\begin{aligned} \mathbf{x}_{E,k+1} &= A_E x_{E,k} + B_E T_{d,k} + E_E T_{ext,k} \\ e_k &= C_E x_{E,k} \end{aligned}. \tag{5.6}$$

The two inputs for the regression at each time step are $T_{d,k}$ and $T_{ext,k}$. The output is e_k.

Forming the Aggregate System Model

Aggregating these two systems results in the aggregate state x, which is the concatenation of x_E and x_T.

$$x_k = \begin{bmatrix} x_{E,k} \\ x_{T,k} \end{bmatrix}. \tag{5.7}$$

The aggregate system model is then given by

$$\begin{aligned} x_{k+1} &= \begin{bmatrix} x_{E,k+1} \\ x_{T,k+1} \end{bmatrix} \\ &= \begin{bmatrix} A_E & B_E \bar{C}_T \\ 0 & \bar{A}_T \end{bmatrix} x_k + \begin{bmatrix} 0 \\ \bar{B}_T \end{bmatrix} u_{T,k} + \begin{bmatrix} E_E \\ 0 \end{bmatrix} T_{ext,k} \\ x_{k+1} &= A_{aug} x_k + B_{aug} u_{T,k} + E_{aug} T_{ext,k} \end{aligned} \tag{5.8}$$

5.2.2 Problem Formulation

With the aggregate dynamics in hand, the optimization problem can be formulated. Several additional variables must be defined. The first is the aggregate state (x_k) consisting of the energy and desired temperature states. Next is the input ($u_{T,k}$), which is the input temperature system at time k. The gain g_k is defined as the energy gain of the system from time step $k-1$ to k. Energy (e_k), is the total energy consumed from time 0 to time k. This is calculated iteratively as the sum of e_{k-1} and g_k. The maximum gain (g_{max}), is the overall maximum energy gain time. The total energy (e_{tot}) is the sum of all the energy used over time. The desired temperature ($T_{d,k}$) is the desired temperature at time k. Finally, the nominal temperature, ($T_{n,k}$) is the nominal set-point temperature for each zone. If other factors were not considered, the desired temperatures would be set as the nominal temperatures.

Since energy dynamics is obtained from a linear regression model, the change in energy from one time step to the next, calculated by the energy state, could be negative. However, g_k is constrained to be nonnegative and at least as large as The energy gain from the regression model. This is done to ensure that e_k is monotonically increasing in relation to k. Note that a choice would be to set e_k as the output of the energy state and calculate g_k based on the difference between e_{k-1} and e_k. However, as mentioned, the approach taken has the benefit of ensuring that the energy gain is not negative.

The optimization problem is formulated to balance energy consumption with the deviation from the nominal temperature, the average power usage, and the average peak power usage. This is done through a weighted sum of objectives based on these quantities. The weights $w_{e,k}$, $w_{T,k}$, $w_{g_{max}}$, and w_{cons} are used to balance the importance of the total energy consumed up to the time step k, the deviation from the nominal temperature at the time step k, the energy gain at the time step k, the overall maximum energy gain time and the total energy consumed in general time, respectively. The weight matrix R penalizes non-zero values of $u_{T,k}$.

The variables of optimization are defined as

$$y = \begin{bmatrix} x_1; x_2; ...; x_N; u_0; u_1; ...; u_{N-1}; g_1; g_2; ...; g_N; \\ g_{max}; e_{tot}; T_{d,0}; T_{d,1}; ...; T_{d,N-1} \end{bmatrix}, \tag{5.9}$$

with $T_{n,k}$, $w_{e,k}$, $w_{T,k}$, $w_{g_{max}}$, w_{cons}, and R being fixed parameters of the optimization. The optimization problem is then of the form

$$\begin{aligned} \min_y & \sum_{k=0}^{N-1} (w_{e,k} e_k^2 + u_k^T R u_k) + \\ & \sum_{k=1}^{N} (w_{T,k}(T_{d,k} - T_{d,n})^2 + w_{e,k} g_k) + \\ & w_{g_{max}} g_{max} + w_{cons} e_{tot} \end{aligned} \tag{5.10}$$

s.t.

$$\begin{aligned} x_{k+1} &= A_{aug} x_k + B_{aug} u_k + E_{aug} T_{e,k} \\ g_{max} &\geq g_k \quad k = 1, ..., N \\ g_k &\geq 0 \quad k = 1, ..., N \\ g_k &\geq C_E [I \ 0](x_k - x_{k-1}) \quad k = 1, ..., N \\ e_k &= e_{k-1} + g_k \quad k = 1, ..., N-1 \\ e_0 &= 0 \\ e_{tot} &= e_{N-1} + g_N \\ T_{d,k} &= \bar{C}_T x_{T,k} \quad k = 0, ..., N-1 \\ l_{T,k} &\leq T_{d,k} \leq b_{T,k} \quad k = 0, ..., N-1 \end{aligned} \tag{5.11}$$

5.3 Experimental Setup

This section outlines the experimental setup used to generate data for the training and validation of the regression model, focusing on realistic simulation environments and controlled thermal dynamics scenarios. The setup takes advantage of both detailed simulation

platforms and simplified physics-based models to create diverse datasets. Specifically, simulations were performed using the Building Optimization Testing Framework (BOPTEST), which integrates EnergyPlus and Modelica for accurate HVAC and building envelope modeling, and using configurable lumped-parameter RC models. These complementary approaches enable comprehensive analysis and robust validation across varying degrees of system complexity, from single-zone environments to multi-zone buildings with intricate thermal interactions.

5.3.1 HVAC Simulation Platform (BOPTEST)

To perform a regression analysis on an HVAC system, a sandbox developed based on first principles is required to provide a ground-truth state for model fitting. In this work, BOPTEST [70] serves as the simulation platform, using EnergyPlus [34] to model the geometry of the building and the thermal envelope, while Modelica [36] simulates the response of the HVAC system. BOPTEST offers a variety of test cases, each representing a real-world building located in different regions. These test cases feature different configurations of the HVAC system and building geometries, allowing a comprehensive evaluation.

For the regression analysis, two test cases were utilized. The first test case, BESTEST Air, represents a single-room building based on the BESTEST Case 900 model, located near Denver, CO, USA. The building has a floor area of 48 m^2, with four exterior walls oriented in the cardinal directions and a flat roof. The south facing wall includes two windows, each measuring 6 m^2. The use of the building corresponds to a two-person office with a light internal load density.

The second test case, a two-zone hydronic apartment, represents a two-room, one-bathroom apartment in Milan, Italy. This apartment has a floor area of 44.5 m^2 and is well-insulated, with an external transmittance of 0.46 $\frac{W}{m^2K}$. The south-east facing façade has a total window area of 8 m^2. The heating system consists of two floor-radiant heating circuits, each serving one room, with heat supplied by a 5-kW air-source heat pump.

5.3.2 Data Generation using Physics-Based RC Models

In addition to potentially using data directly from BOPTEST simulations, physics-based lumped parameter RC models developed in Chapter 4 were utilized as simulation testbeds to generate controlled datasets for regression model training. These models provide a configurable environment based on thermal resistance (R) and capacitance (C) networks, analogous to electrical circuits, to simulate heat transfer. The specific models used were as follows:

- **The 5R5C Single-Zone Model:** This model, shown schematically in Figure 5.2, represents a single thermal zone with significant detail, incorporating five resistive and five capacitive elements. This 5R5C model explicitly models the thermal dynamics of air in the interior room, an internal mass (representing the furniture and interior walls), the structure of the exterior wall, a heated floor slab, and the water storage system that supplies the heating of the slab. Inputs include heating power sources for the air (e.g., Fan Coil Unit - FCU) and the water storage, external temperature, and solar gains. This model allows for simulating a zone with both fast-acting (air-based) and slow-acting (slab-based) thermal responses.

- **The Two-Floor, Eight-Zone Model:** This simulates a more complex building environment with eight distinct thermal zones distributed over two floors, illustrated in Figure 5.3. It includes thermal capacitances for each zone's air, internal mass within each zone, and the walls/floors/ceilings connecting adjacent zones and separating zones from the exterior (including ground contact). The connections are modeled using various RC substructures (also shown in Figure 5.3) to capture different heat transfer paths (e.g., detailed external wall models, simpler internal wall models). Overall, this results in a higher-order system (approximately 50 thermal states) driven by individual controls for each of the eight zones and influenced by external ambient and ground temperatures. It serves to test the approach in a multizone setting with thermal coupling between zones.

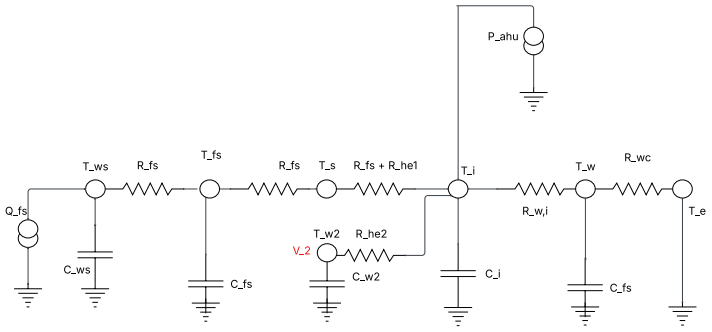

Fig. 5.2: Schematic of the 5R5C single-zone lumped-parameter thermal model used for simulation.

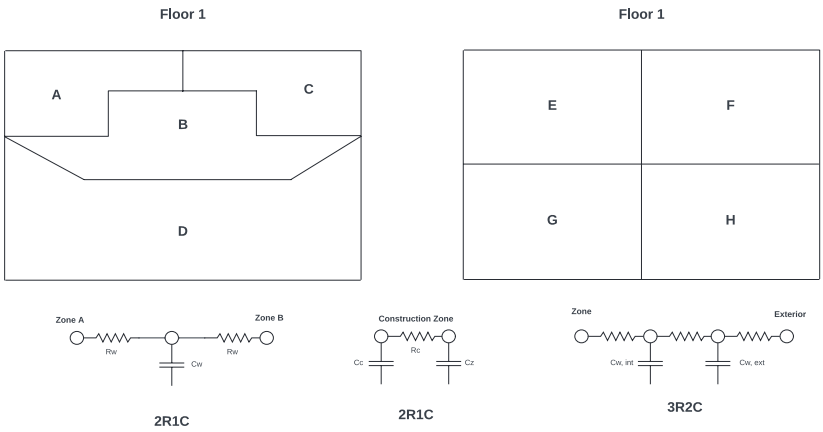

Fig. 5.3: Layout of the two-floor, eight-zone building model (Top) and the RC sub-models used for connections (Bottom).

These RC models are implemented and simulated using Python. The simulations were driven by representative input profiles, including time-varying external temperatures (T_e), solar radiation (Q_S), and sequences of target temperature setpoints ($T_{d,k}$ or $T_{n,k}$ from Section 5.2.2) for the zone(s). Basic control logic (e.g., hysteresis control based on the target setpoints, or direct application of heating/cooling power commands derived from setpoint errors) was applied within the simulation to mimic HVAC operation attempting to track

these setpoints.

During the simulations, the key outputs recorded were the time series data of the applied temperature set point (s), the external temperature (s) and the resulting HVAC power consumption (e.g., Q_{fcu} and Q_{ws} for the 5R5C model, or the sum of individual zone powers for the multizone model). This simulated data, representing the system's input-output behavior under various operating conditions, was then processed and formatted. The objective was to create data sets where the input for the regression model is the desired temperature set point (s) and the external temperature (s) at time k, and the output is the corresponding energy consumed or power (terms related toe_k or g_k from Section 5.2.2) required by the simulated HVAC system. This generated data provides the foundation for training the black-box regression models discussed next.

5.3.3 Regression Process

The workflow for building the HVAC energy regression models consists of four primary stages: (i) data acquisition, (ii) feature and target definition, (iii) model identification and selection, and (iv) validation and performance evaluation.

(i) Data acquisition. Data is generated through simulations using the BOPTEST platform or custom RC models. At each discrete timestep k, the following information is collected:

a) commanded temperature set-points for all controllable zones ($T_{d,k}$),

b) corresponding outside-air temperature ($T_{e,k}$).

Each data set covers a full representative year (8,760 hours) with data sampled at a resolution of 15 minutes, resulting in $N = 35,040$ records per experiment.

(ii) Feature and target definition. For the regression analysis, the simulation data collected is structured into inputs and outputs:

- **Inputs (features):**

  i) zone temperature set-points, $T_{d,k} \in \mathbb{R}^{n_T}$,

ii) external temperature, $T_{e,k} \in \mathbb{R}$,

- **Output (target):** HVAC energy consumed during each interval, g_k [kWh].

All variables are standardized to zero mean and unit variance ($\mu = 0, \sigma = 1$) before identification to ensure consistency and numerical stability.

(iii) Model identification and selection. Data is input into MATLAB's System Identification Toolbox MATLAB [71], exploring several linear structures, including ARX, ARMAX, and state-space models via the subspace method (MNSID). Models are identified by iteratively adjusting the model structure and order to achieve the best fit between predicted and observed energy consumption.

Rather than imposing a strict minimum threshold, model selection involves balancing prediction accuracy, complexity, and robustness. Initially, models with FIT above 90% in validation datasets are prioritized due to their reliability and suitability for integration into the optimization framework. However, recognizing that certain HVAC systems inherently exhibit complex and nonlinear behaviors, lower FIT scores may be acceptable, provided the model captures critical dynamic trends and demonstrates adequate predictive capability. The models selected for further optimization application are explicitly justified based on their validation performance and relevance to system dynamics.

The identified discrete-time state-space realization,

$$x_{E,k+1} = A_E x_{E,k} + B_E \begin{bmatrix} T_{d,k} \\ T_{e,k} \end{bmatrix}, \quad g_k = C_E x_{E,k},$$

is directly incorporated into the optimization formulation outlined in Section 5.2.2.

(iv) Validation and performance evaluation Model validation is carried out using unseen data sequences to quantify predictive precision. The primary metrics for evaluation include the percentage FIT and the root mean square error (RMSE). Models are assessed not only

for their overall accuracy but also for their ability to represent specific system behaviors and energy dynamics relevant to HVAC optimization tasks.

5.4 Results

This section presents two key outcomes of the regression-based HVAC control study. First, predictive performance is quantified using percentage FIT for four different testbeds: the 5R5C single zone RC model, the two-story eight-zone RC model, the BESTEST Case 900 BOPTEST simulation and the BOPTEST case of a two-zone hydronic apartment (Subsection 5.4.1). Second, optimization trade-offs between total energy consumption, peak power demand, and temperature deviation are compared between multiple objective weightings (Subsection 5.4.2).

5.4.1 Regression Model Performance Evaluation

Predictive accuracy is quantified using FIT (%) a. Here, FIT denotes the percentage match between the predicted and actual cumulative energy. All references to "simulated energy response" denote the cumulative energy output e_k over the validation interval.[1]

5R5C RC Model

The regression model for the 5R5C RC system in a single zone achieves a FIT of 90.26% (Figure 5.4), indicating a strong predictive capacity suitable for optimization tasks. The model prediction closely follows the simulated energy use, with minor deviations during initial transients. The ARX/ARMAX structures effectively capture system dynamics, demonstrating suitability for optimization embedding.

[1] In each figure, "energy response" refers to the cumulative energy output e_k from the regression model over time.

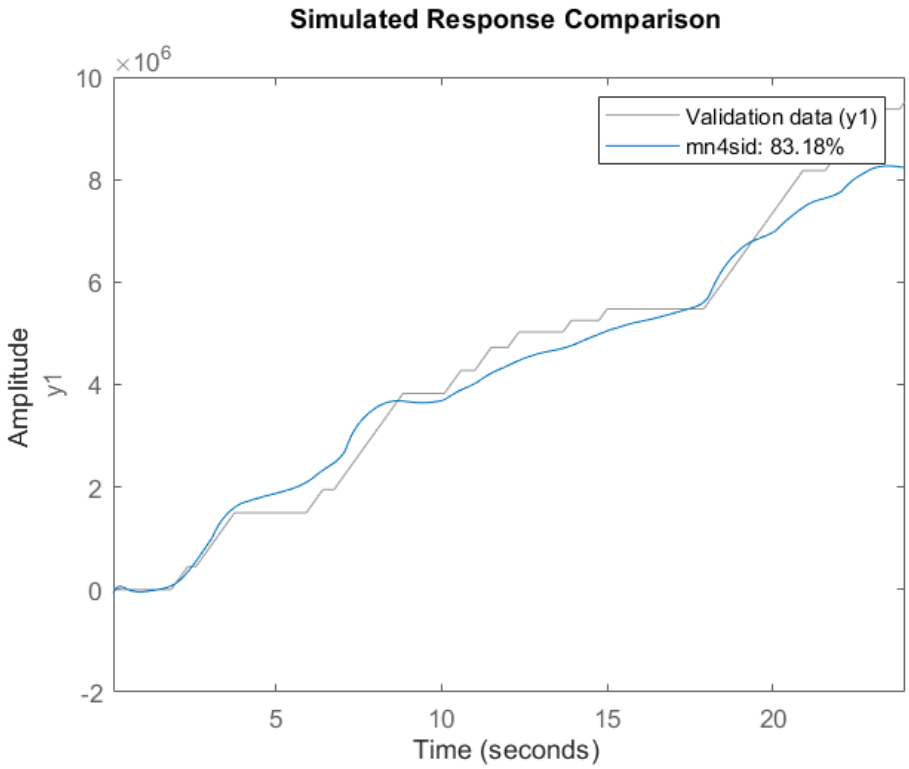

Fig. 5.4: Comparison of the simulated energy response (validation data, y1) and the regression model prediction for the 5R5C single-zone thermal model over a validation interval. The model demonstrates a FIT of 83.18%.

Two-Floor, Eight-Zone RC Model

The linear regression model identified for the more complex eight-zone system achieves a high FIT value of 98.78% (Figure 5.5). Although predictive accuracy is generally excellent, interzone thermal coupling introduces minor inaccuracies during midday intervals affected by solar gains. These dynamics highlight the inherent complexity of multizone models; however, the overall accuracy remains suitable for optimization purposes.

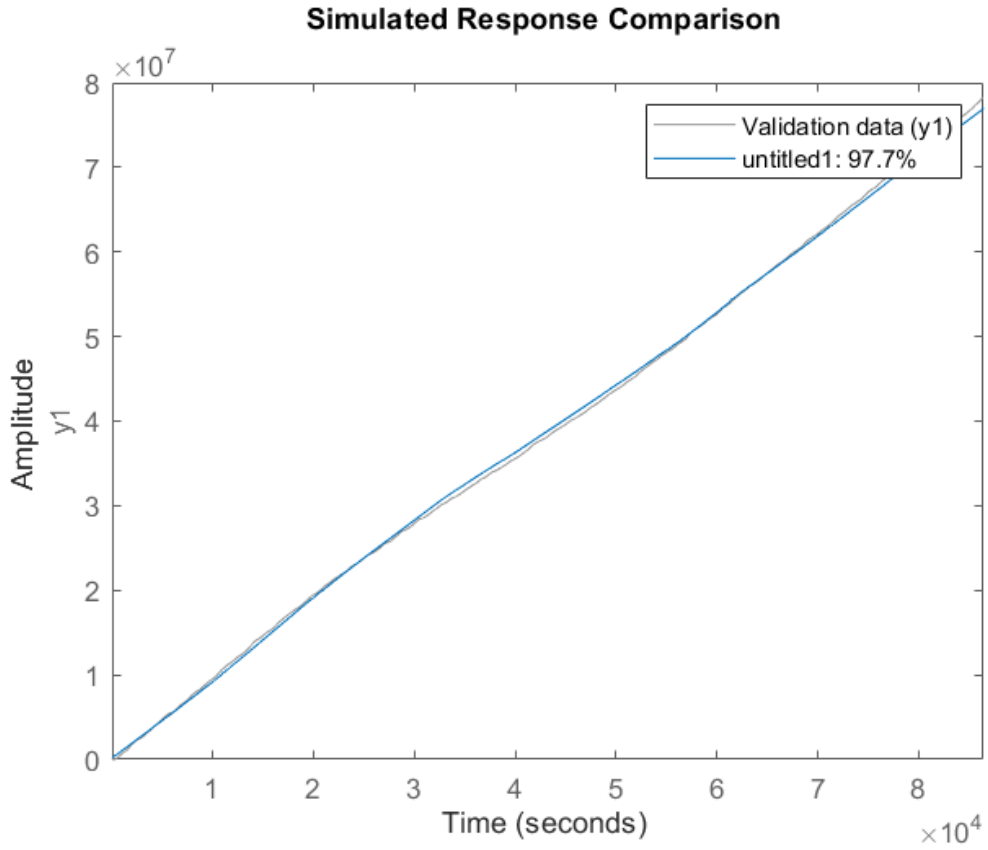

Fig. 5.5: Comparison of the simulated energy response (validation data, y1) and the regression model prediction for the two-floor, eight-zone thermal model over the validation interval, demonstrating a FIT of 97.7%.

BESTEST Case 900 (BOPTEST)

For the BESTEST Case 900 simulation, the ARMAX models achieved FITs of 74.9% and 82.3% (mean 78.6%), capturing seasonal and daily cumulative energy patterns within ± 3 kWh / day, errors small relative to typical 12 kWh/day loads. Although FIT is lower than in the RC testbeds, this level of accuracy suffices to inform optimization of the demand side over multiday horizons (Figure 5.6).

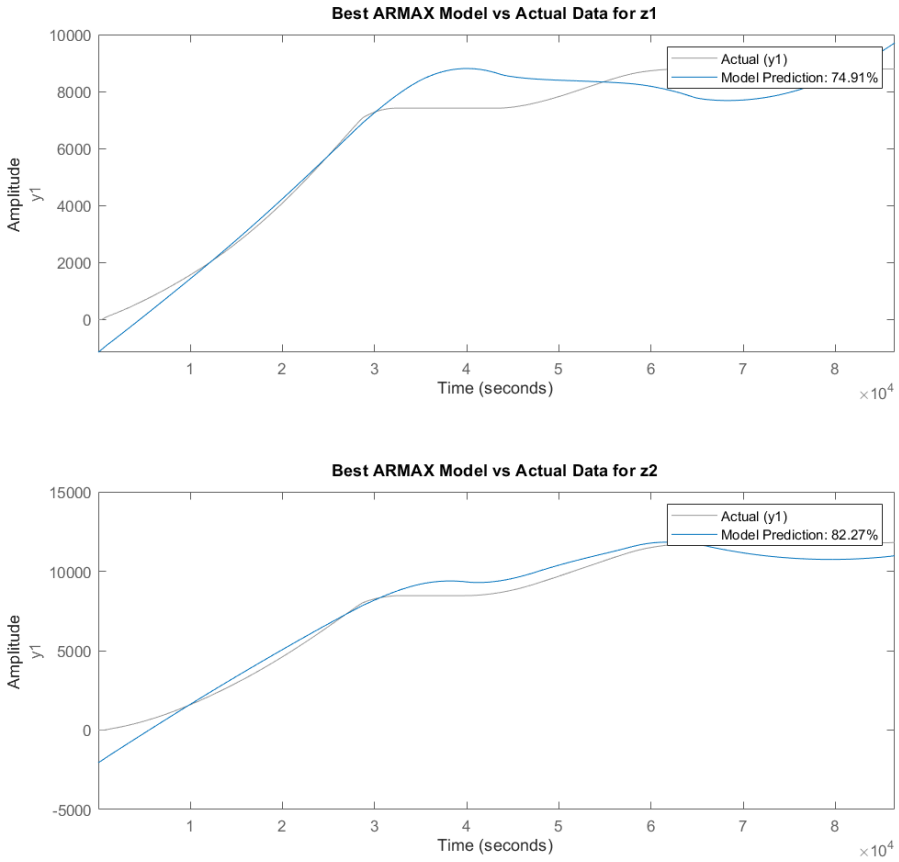

Fig. 5.6: Validation of the ARMAX regression model on BESTEST Case 900 data: actual vs. predicted cumulative energy for Zone 1 (top, FIT 74.9%) and Zone 2 (bottom, FIT 82.3%).

Two-Zone Hydronic Apartment (BOPTEST)

Nonlinear dynamics,thermal stratification, hydronic loop delays and interzone coupling cause linear regression models to fail in the two-zone hydronic apartment, which achieved FITs of 33.1% and 73.3% (Figure 5.7). Such low accuracy produces misleading set-point optimizations (see Section 5.4.5), motivating future work on nonlinear or hybrid identification methods (see Conclusion).

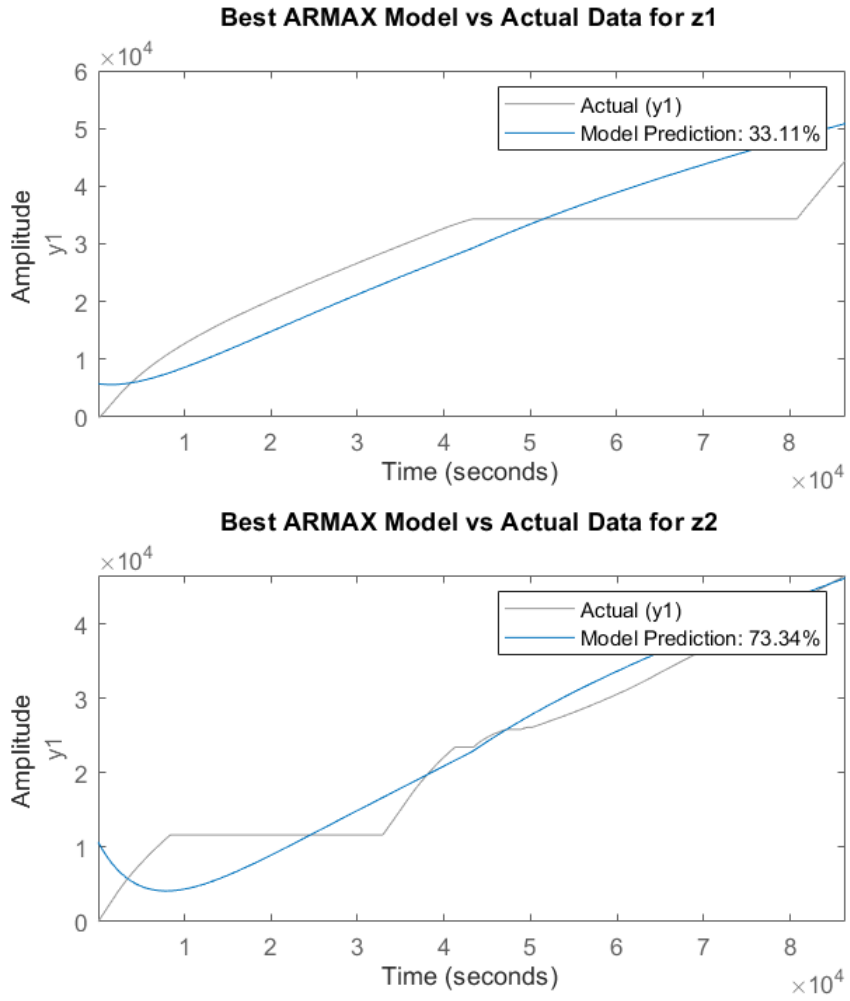

Fig. 5.7: ARMAX regression model validation on the two-zone hydronic apartment: actual versus predicted cumulative energy for Zone 1 (top, FIT 33.1%) and Zone 2 (bottom, FIT 73.3%), illustrating the degraded model accuracy.

5.4.2 Optimization Results and Analysis

Control trade-offs are evaluated for four scenarios: minimizing maximum energy gain, minimizing total energy consumption, minimizing temperature deviation and a combined weighted objective using the regression model and weight definitions from Section 5.2.2 ($w_{e,k}$ for energy, $w_{g,\max}$ for peak gain, $w_{T,k}$ for comfort deviation, and w_{cons} for cumulative energy). The results highlight how the adjustment of these weights affects overall energy

use, peak demand, and occupant comfort.

5R5C Model Scenarios

Each of the following examples has the same nominal temperature profile and external temperature values. They differ in the weight placed on the various terms in the objective function.

Minimizing Maximum Energy Gain This strategy (Figure 5.8) effectively reduces peak power demand but results in higher cumulative energy use, highlighting a clear trade-off where peak management adversely impacts overall efficiency.

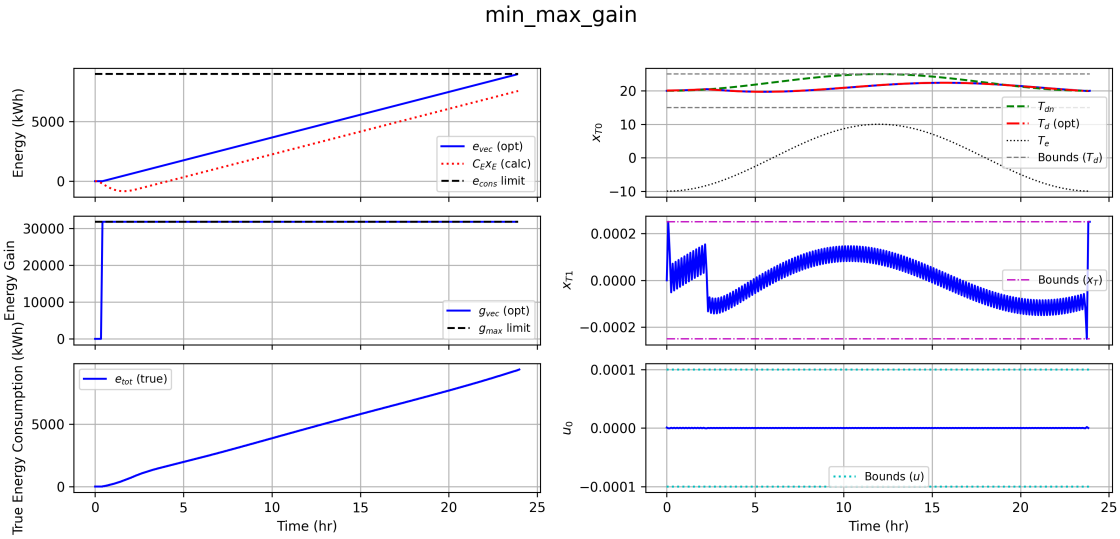

Fig. 5.8: Regression results for minimizing max energy gain. Left column (top to bottom): Optimization cumulative energy (kWh) profile (e_{vec} blue, $C_E x_E$ red dotted, e_{cons} limit black dashed); Optimization energy gain (g_{vec} blue, g_{max} limit black dashed); Total cumulative simulated energy (kJ) from controller (sim_e_{tot} green dash-dot). Right column (top to bottom): Desired Temperature (x_{T0}) dynamics (°C) (Optimized T_d red dash-dot, Nominal T_{dn} green dashed, External T_e black dotted, T_d bounds grey dashed); Temperature State x_{T1} (e.g., Rate) (Optimized state blue, bounds magenta dash-dot); Temperature Input u_0 (e.g., Acceleration) (Optimized input blue, bounds cyan dotted). X-axis is Time (hr).

Minimizing Total Energy Consumption Minimizing total energy use significantly lowers cumulative energy but increases peak demand considerably (Figure 5.9). The results underscore that prioritizing energy efficiency can exacerbate peak loads.

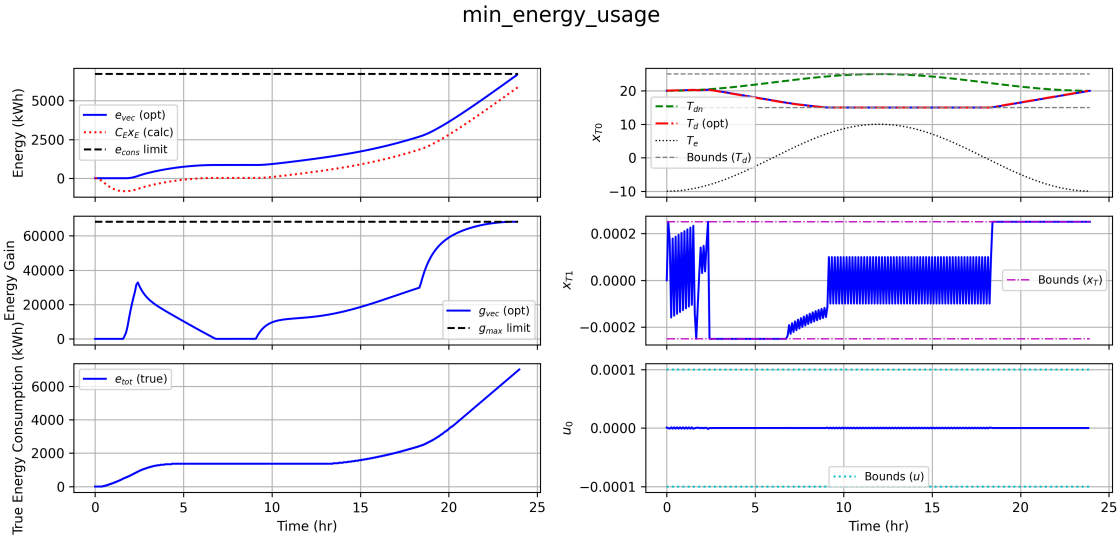

Fig. 5.9: Regression results for minimizing energy usage. Left column (top to bottom): Optimization cumulative energy (kWh) profile (e_{vec} blue, $C_E x_E$ red dotted, e_{cons} limit black dashed); Optimization energy gain (g_{vec} blue, g_{max} limit black dashed); Total cumulative simulated energy (kJ) from controller (sim_e_{tot} green dash-dot). Right column (top to bottom): Desired Temperature (x_{T0}) dynamics (°C) (Optimized T_d red dash-dot, Nominal T_{dn} green dashed, External T_e black dotted, T_d bounds grey dashed); Temperature State x_{T1} (e.g., Rate) (Optimized state blue, bounds magenta dash-dot); Temperature Input u_0 (e.g., Acceleration) (Optimized input blue, bounds cyan dotted). X-axis is Time (hr).

Weighted Optimization Analysis All three scenarios use the same baseline temperature and external temperature profiles (see Section 5.2.2 for weight definitions). They differ only in how the objective function balances peak reduction, cumulative energy, and comfort deviation.

1. **Moderate peak control:** $w_{g,\max} = 10$, $w_{\text{cons}} = 1$, $w_{e,k} = 0$, $w_{T,k} = 1/N$.

2. **Strong peak control:** $w_{g,\max} = 100$, $w_{\text{cons}} = 1$, $w_{e,k} = 0$, $w_{T,k} = 1/N$.

3. **Strong comfort weighting:** $w_{g,\max} = 100$, $w_{\text{cons}} = 1$, $w_{e,k} = 0$, $w_{T,k} = 100/N$.

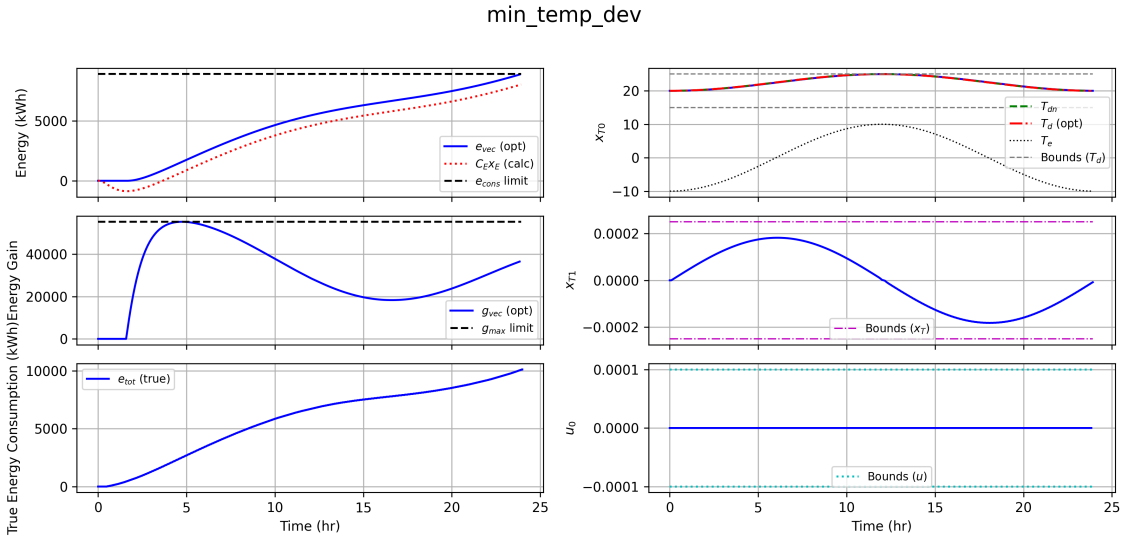

Fig. 5.10: Regression results for minimizing deviation from nominal temperature. Left column (top to bottom): Optimization cumulative energy (kWh) profile (e_{vec} blue, $C_E x_E$ red dotted, e_{cons} limit black dashed); Optimization energy gain (g_{vec} blue, g_{max} limit black dashed); Total cumulative simulated energy (kJ) from controller (sim_e_{tot} green dash-dot). Right column (top to bottom): Desired Temperature (x_{T0}) dynamics (°C) (Optimized T_d red dash-dot, Nominal T_{dn} green dashed, External T_e black dotted, T_d bounds grey dashed); Temperature State x_{T1} (e.g., Rate) (Optimized state blue, bounds magenta dash-dot); Temperature Input u_0 (e.g., Acceleration) (Optimized input blue, bounds cyan dotted). X-axis is Time (hr).

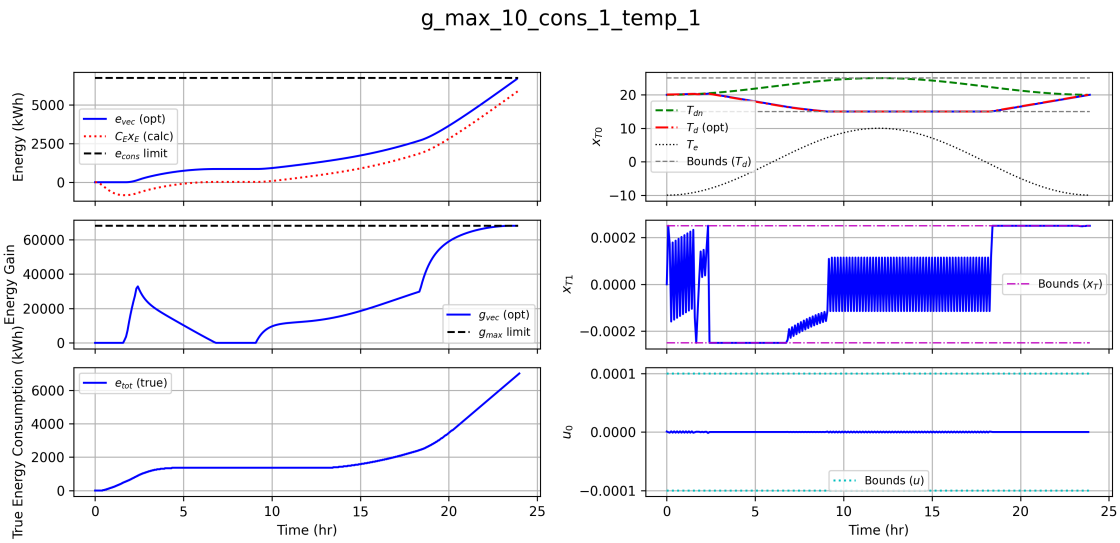

Fig. 5.11: Regression results balanced minimization - $w_{g,max} = 10, w_{cons} = 1$, $w_{e,k} = 0, w_{T,k} = 1/N$. Left column (top to bottom): Optimization cumulative energy (kWh) profile (e_{vec} blue, $C_E x_E$ red dotted, e_{cons} limit black dashed); Optimization energy gain (g_{vec} blue, g_{max} limit black dashed); Total cumulative simulated energy (kJ) from controller (sim_e_{tot} green dash-dot). Right column (top to bottom): Desired Temperature (x_{T0}) dynamics (°C) (Optimized T_d red dash-dot, Nominal T_{dn} green dashed, External T_e black dotted, T_d bounds grey dashed); Temperature State x_{T1} (e.g., Rate) (Optimized state blue, bounds magenta dash-dot); Temperature Input u_0 (e.g., Acceleration) (Optimized input blue, bounds cyan dotted). X-axis is Time (hr).

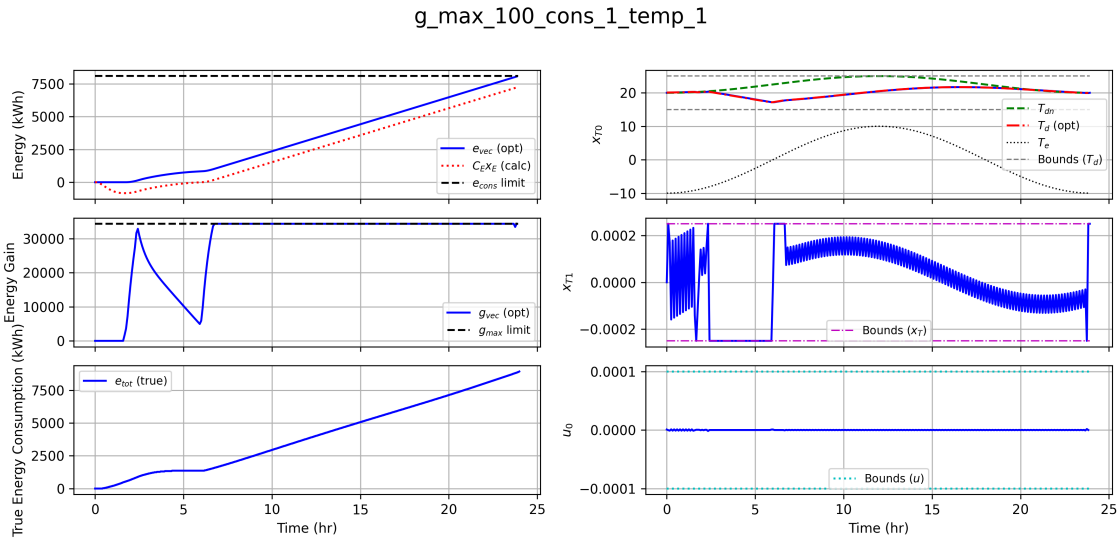

Fig. 5.12: Regression results balanced minimization - $w_{g,max} = 100, w_{cons} = 1$, $w_{e,k} = 0, w_{T,k} = 1/N$. Left column (top to bottom): Optimization cumulative energy (kWh) profile (e_{vec} blue, $C_E x_E$ red dotted, e_{cons} limit black dashed); Optimization energy gain (g_{vec} blue, g_{max} limit black dashed); Total cumulative simulated energy (kJ) from controller (sim_e_{tot} green dash-dot). Right column (top to bottom): Desired Temperature (x_{T0}) dynamics (°C) (Optimized T_d red dash-dot, Nominal T_{dn} green dashed, External T_e black dotted, T_d bounds grey dashed); Temperature State x_{T1} (e.g., Rate) (Optimized state blue, bounds magenta dash-dot); Temperature Input u_0 (e.g., Acceleration) (Optimized input blue, bounds cyan dotted). X-axis is Time (hr).

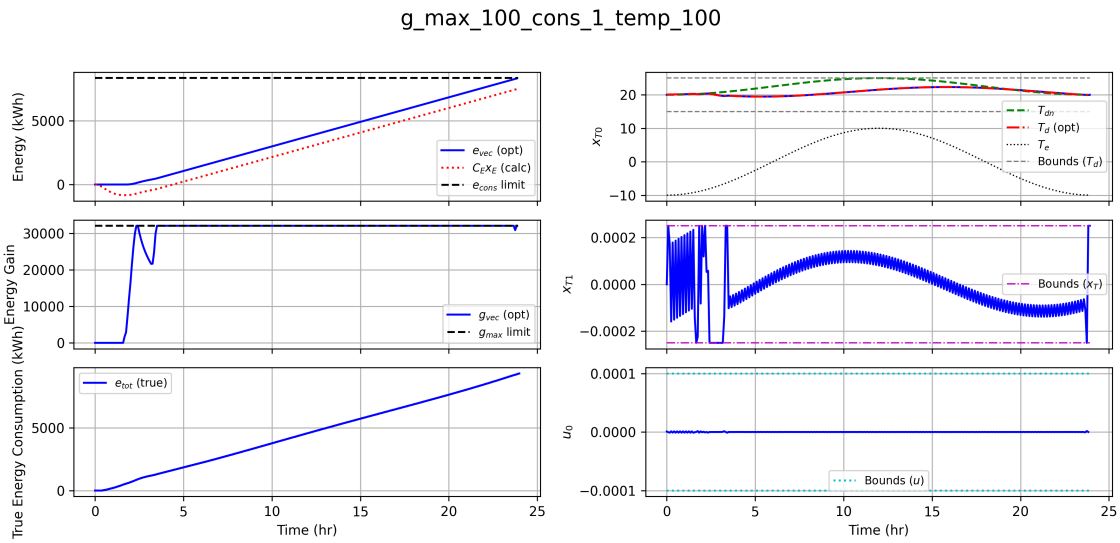

Fig. 5.13: Regression results balanced minimization - $w_{g,max} = 100, w_{cons} = 1$, $w_{e,k} = 0, w_{T,k} = 100/N$. Left column (top to bottom): Optimization cumulative energy (kWh) profile (e_{vec} blue, $C_E x_E$ red dotted, e_{cons} limit black dashed); Optimization energy gain (g_{vec} blue, g_{max} limit black dashed); Total cumulative simulated energy (kJ) from controller (sim_e_{tot} green dash-dot). Right column (top to bottom): Desired Temperature (x_{T0}) dynamics (°C) (Optimized T_d red dash-dot, Nominal T_{dn} green dashed, External T_e black dotted, T_d bounds grey dashed); Temperature State x_{T1} (e.g., Rate) (Optimized state blue, bounds magenta dash-dot); Temperature Input u_0 (e.g., Acceleration) (Optimized input blue, bounds cyan dotted). X-axis is Time (hr).

- **Moderate peak control** ($w_{g,\max} = 10$, $w_{\text{cons}} = 1$, $w_{e,k} = 0$, $w_{T,k} = 1/N$): peak power reduced by $\approx 5\%$ (from 4.0 kW to 3.8 kW), total energy increased by $\approx 2\%$ (to 18.4 kWh/day).

- **Strong peak control** ($w_{g,\max} = 100$, $w_{\text{cons}} = 1$, $w_{e,k} = 0$, $w_{T,k} = 1/N$): Maximum power reduced by $\approx 12\%$ (to 3.5 kW), total energy increased by $\approx 4\%$ (to 19.2 kWh/day).

- **Strong comfort weighting** ($w_{g,\max} = 100$, $w_{\text{cons}} = 1$, $w_{e,k} = 0$, $w_{T,k} = 100/N$): the temperature deviation narrowed from $\pm 1.5\,^{\circ}$C to $\pm 0.8\,^{\circ}$C; peak power fell by $\approx 8\%$ (to 3.7 kW), the energy rose by $\approx 6\%$ (to 19.0 kWh/day).

In general, increasing $w_{g,\max}$ yields diminishing returns: initial peak reductions are efficient (moderate case), but pushing $w_{g,\max}$ to 100 drives energy and comfort costs higher. Introducing a high-comfort weight recovers tighter temperature control at the expense of slightly higher energy, suggesting a balanced middle ground for applications where both demand charges and occupant comfort are critical.

5.4.3 Results for Regression-based Optimization Approach - Two Floor Model

An example of the two-floor model is presented in this section. The weights are defined as $w_{g,max} = 100, w_{cons} = 1$, $w_{e,k} = 0, w_{T,k} = 100/N$. As can be seen in Figure 5.14, the energy consumption is slightly different from the actual. There is a large zone in the middle where the predicted energy usage is flat, while the real energy usage has a linear slope. The sharp energy-use spikes at the beginning and end are in good agreement.

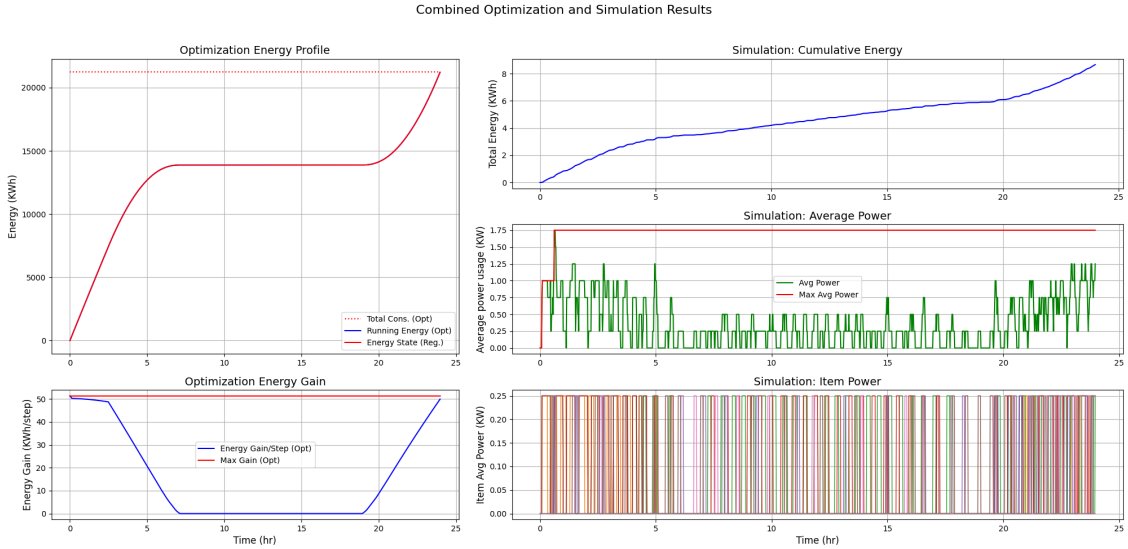

Fig. 5.14: Regression results balanced minimization - $w_{g,max} = 100, w_{cons} = 1$, $w_{e,k} = 0, w_{T,k} = 100/N$. Left column (top to bottom): Optimization energy profile (red-dotted: total consumption, blue: running energy, red: energy state); Optimization energy gain (blue: gain/step, red: max gain). Right column (top to bottom): Simulation cumulative energy (KWh); Simulation average power (green: average, red: running max KW); Simulation average power per zone (KW).

5.4.4 Regression-based Optimization Results: BESTEST Case 900

This subsection evaluates the optimization performance using the regression model derived from the BESTEST Case 900 BOPTEST simulation data over a 14-day period. Three scenarios, defined by different objective function weightings (Table 5.2), are compared to analyze the trade-offs between temperature tracking fidelity, energy consumption, and peak power demand. The `Baseline` scenario prioritizes strict temperature tracking ($w_{T,k} = 1$). The `Balanced (Peak-Focused)` scenario significantly reduces the weight on temperature tracking ($w_{T,k} = 1e-1$) to prioritize peak power mitigation, while the `Minimized Energy Consumption` scenario further relaxes the temperature constraint ($w_{T,k} = 1e-5$) to focus purely on reducing overall energy usage. Initial analysis indicated that peak power (g_{max}) could not be effectively controlled using this regression-based approach with the available inputs; therefore, its corresponding weight ($w_{g,max}$) was not varied in these scenarios.

Name	$w_{g,max}$	$w_{T,k}$	$w_{e,k}$
Baseline	1	1	1
Balanced (Peak-Focused)	1	1×10^{-1}	1
Minimized Energy Consumption	1	1×10^{-5}	1

Table 5.2: Revised Weighting values for BESTEST Case 900 Simulations. Note: The $w_{e,k}$ value for Minimized Energy Consumption was assumed to be 1 based on the original table; please verify.

Baseline: Tracking Desired Temperature Trajectory

In the baseline scenario (Figs. 5.15–5.16), the controller adheres strictly to the desired temperature setpoint ($w_{T,k} = 1$). This results in frequent on–off cycling of the HVAC system to maintain tight control, as seen by the indoor temperature closely following the heating setpoint in Figure 5.15. Consequently, peak power demand reaches approximately 3.5 kW (Table 5.3), driven primarily by morning and evening recovery periods influenced by outdoor temperature fluctuations. Total daily energy consumption remains relatively steady, averaging around 35-40 kWh per day (Figure 5.16, bottom), reflecting the energy cost of maintaining precise temperature control. The peak energy consumption occurred on day 1, reaching approximately 40 kWh (Table 5.3).

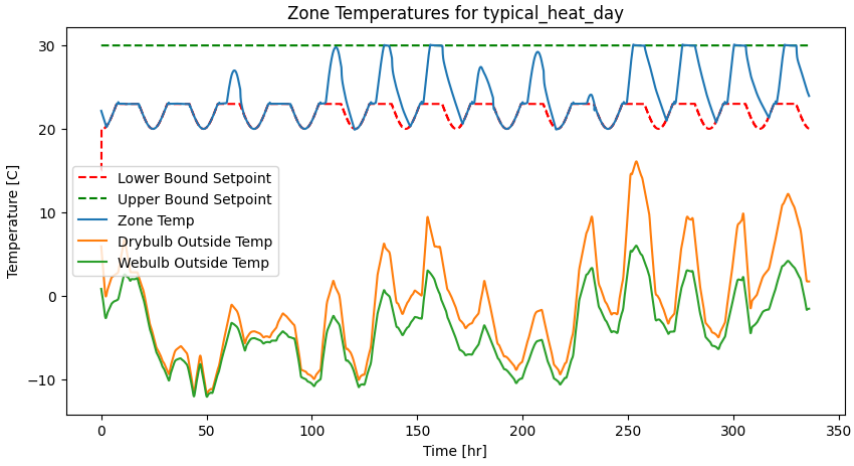

Fig. 5.15: Baseline temperature tracking over 14 days. Dashed green=fixed upper (cooling) setpoint, dashed red=heating setpoint (lower bound, controlled), solid blue=simulated zone air temperature, solid orange=outdoor dry-bulb temperature, solid green=outdoor wet-bulb temperature.

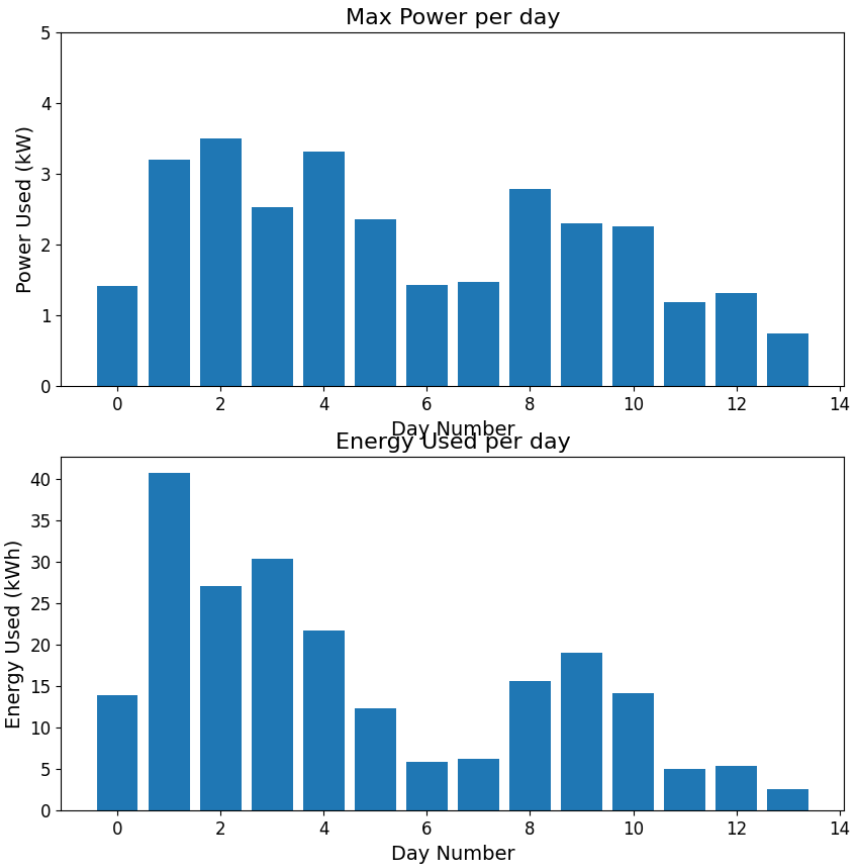

Fig. 5.16: Baseline daily HVAC performance: (top) peak power demand exhibits pronounced morning and evening spikes; (bottom) total daily energy consumption remains relatively steady, reflecting the cost of tight setpoint control.

Objective: Balance of Energy Consumption and Temperature Profile (Peak-Focused)

The `Balanced (Peak-Focused)` profile (Figs. 5.17–5.18), which reduces the weighting on temperature tracking ($w_{T,k} = 1e-1$), prioritizes flattening the power profile. It employs preconditioning strategies ahead of expected peak load periods, significantly reducing the maximum daily peak power demand by approximately 20% compared to the baseline, down to about 2.8 kW (Table 5.3, Figure 5.18, top). However, this peak reduction comes at the cost of increased overall energy consumption, approximately 6% higher than baseline over the 14 days (Table 5.3), likely due to longer HVAC runtimes for preconditioning (Figure 5.18,

bottom). Despite the increase in total energy, the energy use on the peak consumption day was reduced by approximately 5 kWh compared to the baseline (Table 5.3). This scenario allows for moderate temperature deviations from the setpoint, as seen in Figure 5.17.

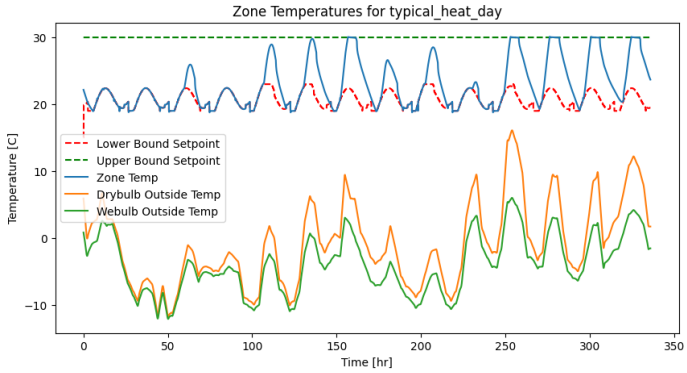

Fig. 5.17: Temperature profile when minimizing peak power demand. Dashed green=fixed upper setpoint, dashed red=heating setpoint (lower bound, optimized for pre-conditioning), solid blue=simulated zone air temperature, solid orange=outdoor dry-bulb temperature, solid green=outdoor wet-bulb temperature.

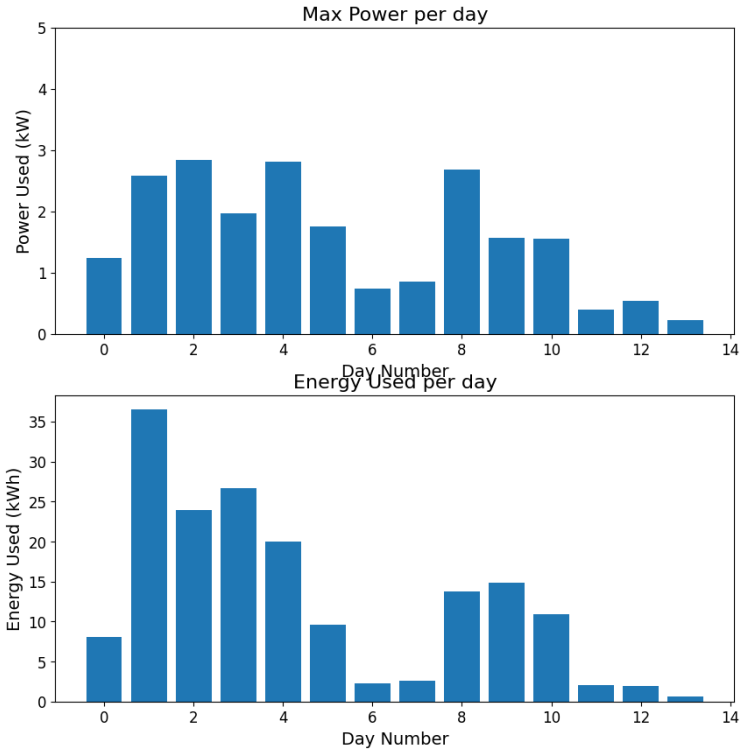

Fig. 5.18: HVAC power under peak-minimization: (top) daily peak demand is reduced by roughly 20 percent versus baseline; (bottom) total daily energy use increases by approximately 6 percent due to longer runtimes.

Objective: Minimize Total Energy Consumption

Finally, when minimizing total energy consumption (Figs. 5.19–5.20), the optimizer, guided by a very low temperature tracking weight ($w_{T,k} = 1e-5$), allows for greater temperature dropouts during off-peak hours (Figure 5.19). This strategy successfully reduces overall energy use by approximately 15% compared to the baseline (Table 5.3). The peak energy consumption day saw a reduction of about 10 kWh, down to 30 kWh (Table 5.3). However, this energy saving forces the HVAC system to work harder during recovery periods, leading to significantly higher instantaneous peak demands during these times, although the overall maximum peak power over the 14 days (approx. 3.0 kW) was slightly lower than the baseline (Table 5.3, Figure 5.20). The average daily energy consumption is reduced compared to both the baseline and balanced profiles.

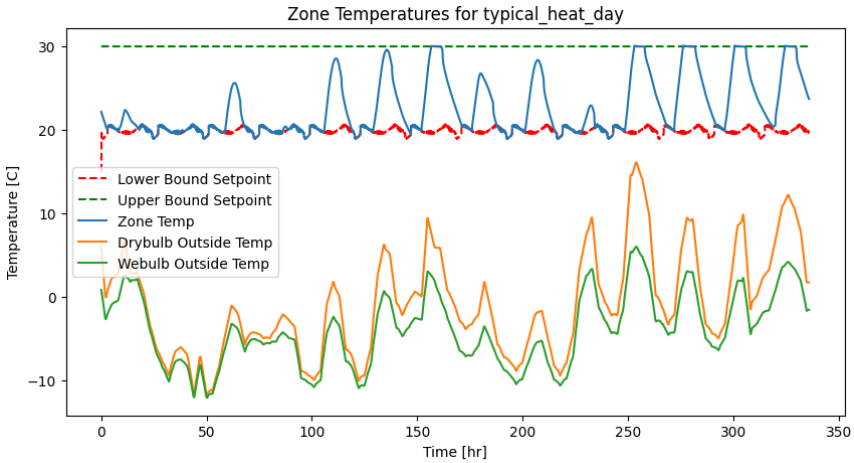

Fig. 5.19: Temperature profile when minimizing total energy consumption. Dashed green=fixed upper setpoint, dashed red=heating setpoint (lower bound, optimized for wider setbacks), solid blue=simulated zone air temperature, solid orange=outdoor dry-bulb temperature, solid green=outdoor wet-bulb temperature.

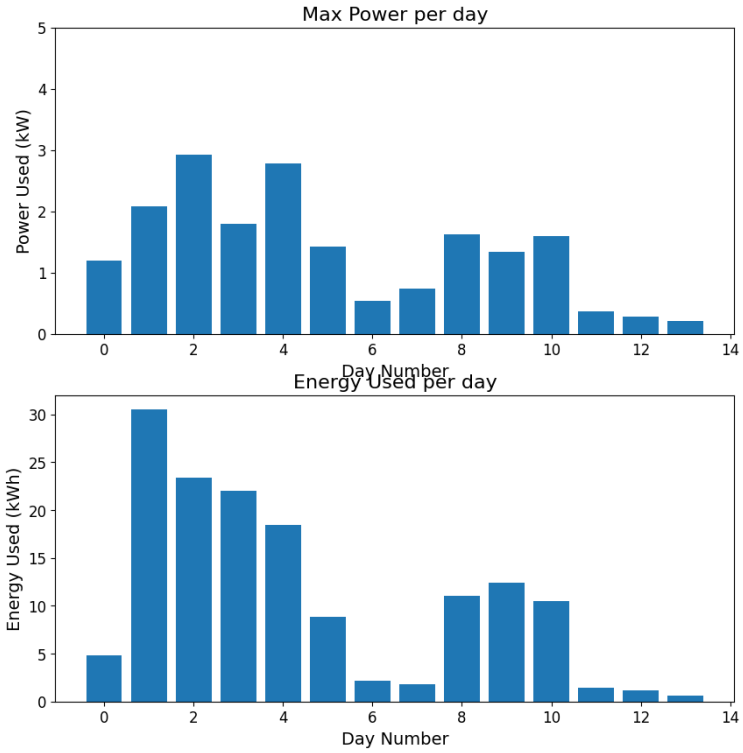

Fig. 5.20: HVAC power under energy-only objective: (top) peak demand increases sharply during recovery periods, (bottom) overall daily energy usage is reduced compared to baseline.

Discussion

Key Performance Indicator (KPI)	Baseline	Balanced (Peak-Focused)	Minimized Energy
Max Peak Power (kW)	≈ 3.5*	≈ 2.8*	≈ 3.0*
Total Energy Use (14 days, kWh)	211	171	142
Peak Energy Day (kWh)	≈ 40	≈ 35*	≈ 30
Approx. Temp Deviation Range (°C)	Low*	Moderate*	High*

Table 5.3: Summary of Key Performance Indicators (KPIs) over the 14-day BESTEST Case 900 Simulation. Note: Values marked with * are estimates based on figures and text percentages;

These results (summarized in Table 5.3) clearly illustrate the inherent trade-offs in HVAC control optimization using this regression model. The `Baseline` scenario achieves tight set-point control and stabilizes comfort but incurs the highest energy cost (211 kWh) and the highest peak power demand ($\approx$ 3.5 kW). The `Minimized Energy Consumption` strategy achieves the lowest total energy use (142 kWh) but at the expense of wider temperature swings and potentially sharp recovery peaks, with a maximum peak power ($\approx$ 3.0 kW) only slightly below the baseline. The `Balanced (Peak-Focused)` control provides a middle ground, achieving the lowest maximum peak power ($\approx$ 2.8 kW, a reduction of 20% from baseline) while also reducing total energy consumption compared to the baseline (171 kWh vs 211 kWh), albeit allowing moderate temperature deviations. The choice between these strategies depends on the specific priorities, such as minimizing operational costs (`Minimized Energy`), reducing demand charges (`Balanced (Peak-Focused)`), or maximizing occupant comfort (`Baseline`).

5.4.5 Regression-based Optimization Results: Two-Zone Hydronic Apartment

This section evaluates the regression-based optimization strategy on the two-zone hydronic apartment from BOPTEST. Due to the relatively low identification accuracy (FIT scores of 33% and 73%), we examine how the optimizer, tuned to heavily penalize peak power, performs against a simple rule-based square-wave heating schedule. Figures 5.21 and 5.22 show the optimized zone temperature trajectories and the resulting daily peak power and energy consumption, while Figures 5.24 and 5.25

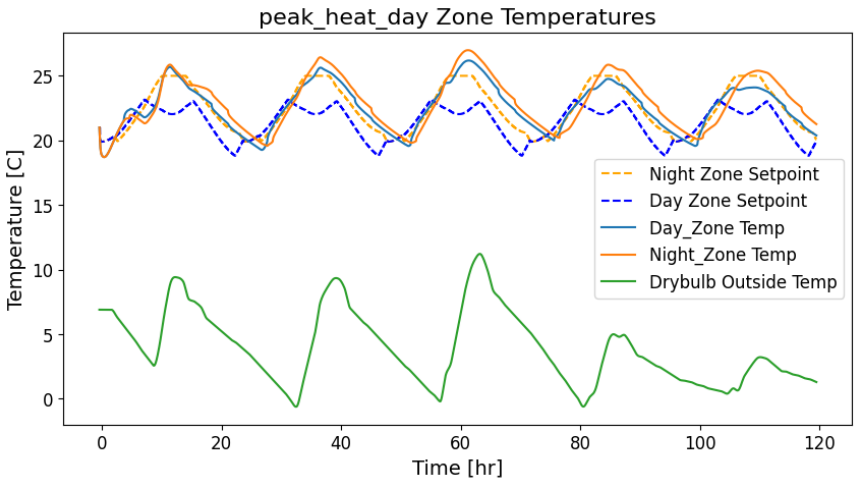

Fig. 5.21: Two-Zone Hydronic (Optimizer): Zone air temperatures under the regression-based controller. Dashed green = fixed upper (cooling) setpoint, dashed red = heating setpoint (lower bound, optimized), solid blue = Zone 1 temperature, solid cyan = Zone 2 temperature, solid orange = outdoor dry-bulb temperature, solid green = outdoor wet-bulb temperature.

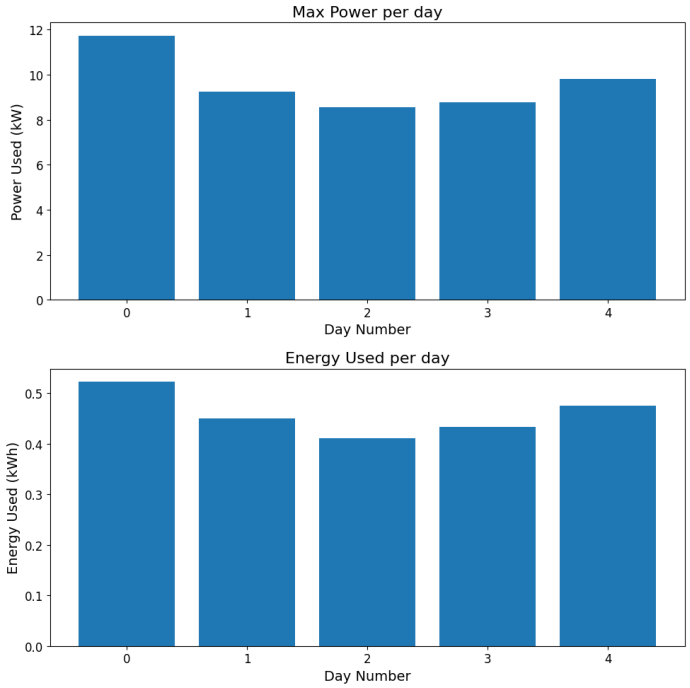

Fig. 5.22: Two-Zone Hydronic (Optimizer): Daily HVAC performance under regression-based control. (Top) peak power demand each day (kW); (Bottom) total daily energy consumption (kWh).

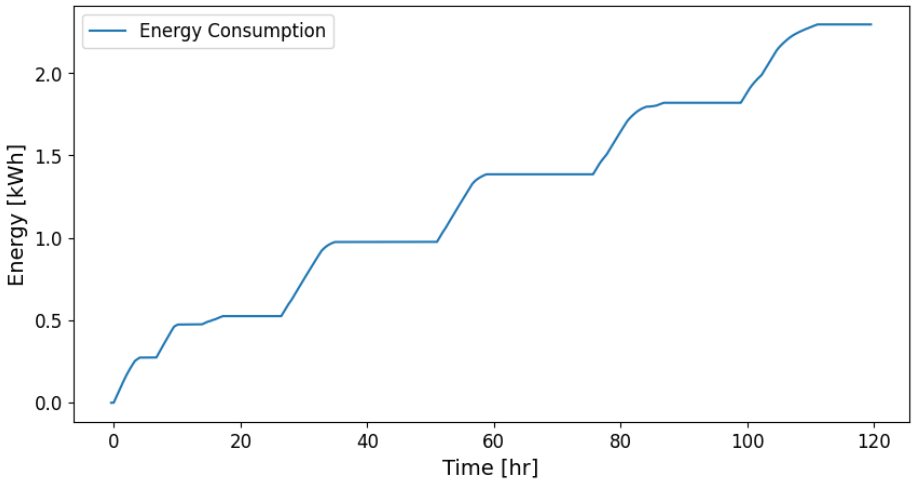

Fig. 5.23: Two-Zone Hydronic (Optimizer): The cumulative energy over all days in the simulation

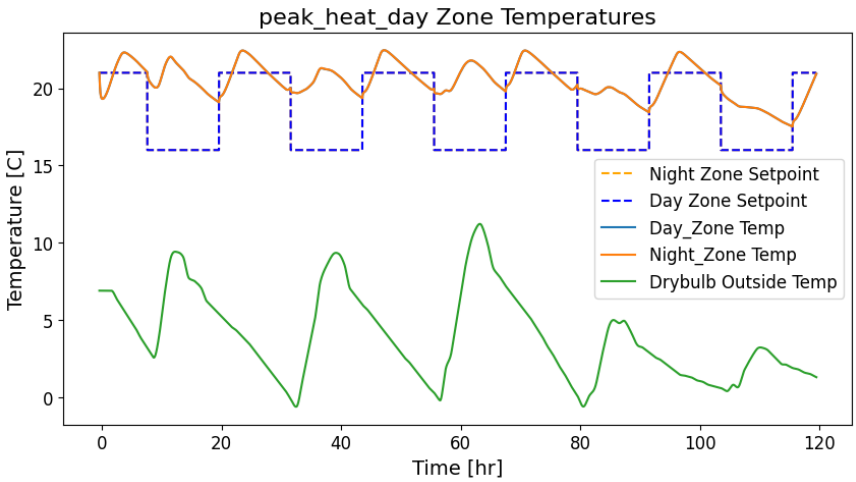

Fig. 5.24: Two-Zone Hydronic (Baseline): Zone air temperatures under rule-based square-wave control. Dashed green = fixed upper setpoint, dashed red = heating setpoint (lower bound), solid blue = Zone 1 temperature, solid cyan = Zone 2 temperature, solid orange = outdoor dry-bulb temperature, solid green = outdoor wet-bulb temperature.

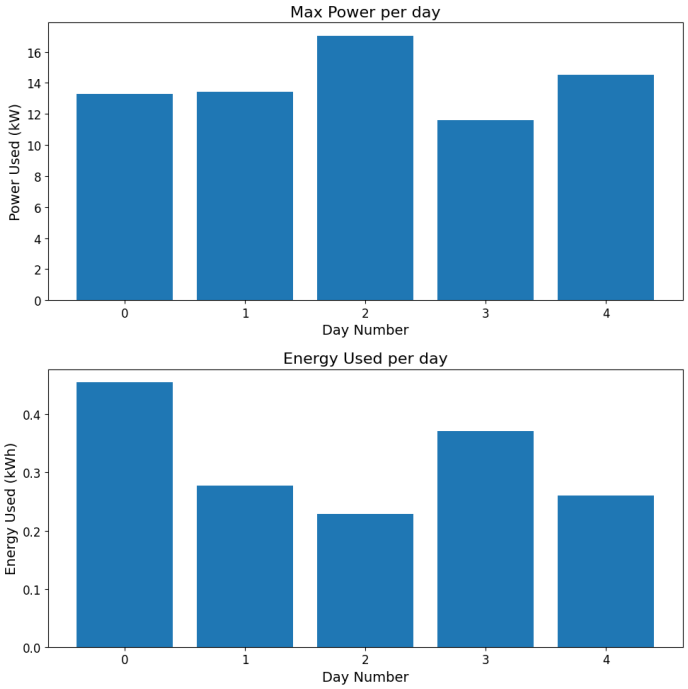

Fig. 5.25: Two-Zone Hydronic (Baseline): Daily HVAC performance under rule-based control. (Top) peak power demand each day (kW); (Bottom) total daily energy consumption (kWh).

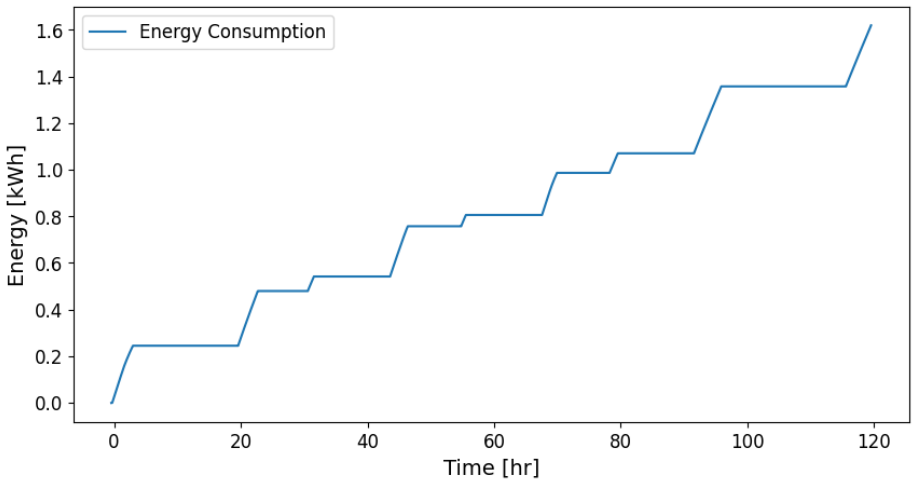

Fig. 5.26: Two-Zone Hydronic (Baseline): The cumulative energy over all days in the simulation

Because the linear regression model for this hydronic apartment only achieved FIT scores of 33% and 73%, its embedded optimizer yields mixed results compared to the simple square-wave baseline. Under the regression-based controller (Figs. 5.21–5.22), average daily peak power is reduced by approximately 14%-55% relative to the baseline rule-based control (Figs. 5.24–5.25). However, total energy consumption increases by approximately 37.5%. The zone comfort is able to follow a better temperature profile relative to the baseline. The commanded profile slowly lowers the temperature at night and gradually increases for the morning. These results demonstrate that optimization built on an underperforming model can worsen both energy use, highlighting the critical need for more accurate or nonlinear system identification before deploying model-based control in complex multi-zone hydronic systems.

5.5 Conclusion

A regression-based optimization strategy for HVAC set point control in microgrids was developed and rigorously evaluated on four distinct testbeds. The approach demonstrated significant success when applied to linear RC models (the 5R5C single zone and the two-story eight-zone), achieving high FIT scores (above 90%) and clearly illustrating the trade-offs

between minimizing energy consumption and reducing peak power demand. These results confirm that high-fidelity linear models can reliably inform optimization decisions for such systems.

Application to the BESTEST Case 900 BOPTEST simulation, a more realistic building model, yielded moderate FIT scores (mean 78.6%). Despite the lower accuracy compared to the RC models, the identified ARMAX models captured essential daily and seasonal energy patterns. Optimization based on these models successfully demonstrated the ability to manage trade-offs: the `Balanced (Peak-Focused)` strategy reduced peak power by approximately 20% compared to the `Baseline` while also reducing total energy use, whereas the `Minimized Energy` strategy achieved the lowest energy consumption (142 kWh vs. 211 kWh baseline) at the cost of wider temperature variations and less peak reduction. This illustrates that even moderately accurate linear regression models can provide valuable guidance for demand-side management over multi-day horizons in standard building types.

In contrast, the strategy performed poorly on the complex two-zone hydronic apartment case, where the linear regression models achieved low FIT scores (33.1% and 73.3%). This inaccuracy, stemming from significant nonlinearities, thermal stratification, and hydronic loop delays inherent in the system, led to suboptimal control actions. While peak power was reduced compared to a simple baseline, the optimization resulted in a substantial increase in total energy consumption (approx. 37.5%) and potentially compromised comfort. This outcome underscores the critical dependence of model-based control effectiveness on model fidelity, particularly for complex, nonlinear HVAC systems.

CHAPTER 6

Conclusion

Effective coordination of distributed energy resources (DER) is crucial to realize the full benefits of microgrids in modern energy systems, including enhanced stability, cost reduction, and better integration of renewables. Addressing the complexities of forecasting, modeling diverse DERs such as HVAC systems, and optimizing operations with multiple objectives remains a significant challenge. This thesis has presented a comprehensive framework that makes significant progress toward managing these challenges. The primary contributions and accomplishments are:

1. **Developed and evaluated a hierarchical Receding Horizon Control (RHC) framework** to coordinate energy storage systems (ESS) and electric vehicle (EV) chargers. This framework utilizes XGBoost forecasting and prioritizes operational objectives such as peak power reduction (demonstrating significant reductions in simulation , load smoothing, and State-of-Charge (SOC) management, while integrating demand response capabilities.

2. **Developed and analyzed a physics-based HVAC control methodology** that employs lumped sum resistor-capacitor (RC) modeling. This included a novel 5R5C model for water storage systems, a general graph-based technique for multizone system modeling, a proof of model stability, and application of optimal control (QP), while also identifying practical challenges in translating optimal power plans to simpler controllers.

3. **Developed and integrated a black-box HVAC control methodology** using linear regression models (ARMAX) identified via the BOPTEST platform. This allows for optimized temperature set-point control based on learned energy dynamics within the broader hierarchical framework.

These contributions provide a robust and adaptable approach to coordinating multiple DERs within a microgrid, balancing grid stability objectives with efficient resource utilization. Dual HVAC control strategies offer flexibility depending on available data and modeling preferences.

Although this work presents significant advances, future research could focus on improving the regression process by incorporating more advanced system identification algorithms such as Hammerstein-Wiener models or Koopman operators. Further validation using hardware-in-the-loop testing or real-world deployment would also strengthen the findings. Additionally, exploring alternative hierarchical objective structures or incorporating more sophisticated stochastic optimization techniques remain promising avenues for future investigation.

REFERENCES

[1] H. Viot, A. Sempey, L. Mora, J. Batsale, and J. Malvestio, "Model predictive control of a thermally activated building system to improve energy management of an experimental building: Part i—modeling and measurements," *Energy and Buildings*, vol. 172, pp. 94–103, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378778818311435

[2] T. Dewson, B. Day, and A. Irving, "Least squares parameter estimation of a reduced order thermal model of an experimental building," *Building and Environment*, vol. 28, no. 2, pp. 127–137, 1993.

[3] S. F. Fux, A. Ashouri, M. J. Benz, and L. Guzzella, "Ekf based self-adaptive thermal model for a passive house," *Energy and Buildings*, vol. 68, pp. 811–817, 2014.

[4] A. Tindale, "Third-order lumped-parameter simulation method," *Building Services Engineering Research and Technology*, vol. 14, no. 3, pp. 87–97, 1993.

[5] F. Belić, Ž. Hocenski, and D. Slišković, "Thermal modeling of buildings with rc method and parameter estimation," in *2016 International Conference on Smart Systems and Technologies (SST)*. IEEE, 2016, pp. 19–25.

[6] V. Chandan and A. G. Alleyne, "Decentralized predictive thermal control for buildings," *Journal of Process Control*, vol. 24, no. 6, pp. 820–835, 2014.

[7] G. B. Gharehpetian, H. R. Baghaee, and M. M. Shabestary, "Chapter 1 - introduction," in *Microgrids and Methods of Analysis*, G. B. Gharehpetian, H. R. Baghaee, and M. M. Shabestary, Eds. Academic Press, 2021, pp. 1–6. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128161722000018

[8] "Confronting duck curve: How to address over-generation of solar energy." [Online]. Available: https://www.energy.gov/eere/articles/confronting-duck-curve-how-address-over-generation-solar-energy

[9] "Frequently asked questions (faqs) - u.s. energy information administration (eia)." [Online]. Available: https://www.eia.gov/tools/faqs/faq.php?id=1174&t=1

[10] S. Ahmad, M. Shafiullah, C. B. Ahmed, and M. Alowaifeer, "A review of microgrid energy management and control strategies," *IEEE Access*, vol. 11, pp. 21 729–21 757, 2023.

[11] M. B. Sigalo, A. C. Pillai, S. Das, and M. Abusara, "An energy management system for the control of battery storage in a grid-connected microgrid using mixed integer linear programming," *Energies*, vol. 14, no. 19, 2021. [Online]. Available: https://www.mdpi.com/1996-1073/14/19/6212

[12] I. R. da Silva, R. de A.L. Rabêlo, J. J. Rodrigues, P. Solic, and A. Carvalho, "A preference-based demand response mechanism for energy management in a microgrid," *Journal of Cleaner Production*, vol. 255, p. 120034, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0959652620300810

[13] R. Mannini, T. Darure, J. Eynard, and S. Grieu, "Predictive energy management of a building-integrated microgrid: A case study," *Energies*, vol. 17, no. 6, 2024. [Online]. Available: https://www.mdpi.com/1996-1073/17/6/1355

[14] E. Rezaei and H. Dagdougui, "Optimal real-time energy management in apartment building integrating microgrid with multizone hvac control," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 11, pp. 6848–6856, 2020.

[15] S. Sahoo, S. Mishra, S. Jha, and B. Singh, "A cooperative adaptive droop based energy management and optimal voltage regulation scheme for dc microgrids," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 4, pp. 2894–2904, 2020.

[16] M. Dashtdar, M. Bajaj, and S. M. S. Hosseinimoghadam, "Design of optimal energy management system in a residential microgrid based on smart control," *Smart Science*, vol. 10, no. 1, pp. 25–39, 2022.

[17] T. Shekari, A. Gholami, and F. Aminifar, "Optimal energy management in multi-carrier microgrids: an milp approach," *Journal of Modern Power Systems and Clean Energy*, vol. 7, no. 4, pp. 876–886, 2019.

[18] B. Lu and M. Shahidehpour, "Short-term scheduling of battery in a grid-connected pv/battery system," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 1053–1061, 2005.

[19] M. R. Sandgani and S. Sirouspour, "Coordinated optimal dispatch of energy storage in a network of grid-connected microgrids," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 3, pp. 1166–1176, 2017.

[20] B. Wang, C. Zhang, and Z. Y. Dong, "Interval optimization based coordination of demand response and battery energy storage system considering soc management in a microgrid," *IEEE Transactions on Sustainable Energy*, vol. 11, no. 4, pp. 2922–2931, 2020.

[21] A. Hassan, S. Acharya, M. Chertkov, D. Deka, and Y. Dvorkin, "A hierarchical approach to multienergy demand response: From electricity to multienergy applications," *Proceedings of the IEEE*, vol. 108, no. 9, pp. 1457–1474, 2020.

[22] S. Leonori, A. Martino, F. M. Frattale Mascioli, and A. Rizzi, "Microgrid energy management systems design by computational intelligence techniques," *Applied Energy*, vol. 277, p. 115524, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261920310369

[23] F. Rodríguez, A. M. Florez-Tapia, L. Fontán, and A. Galarza, "Very short-term wind power density forecasting through artificial neural networks for microgrid control," *Renewable Energy*, vol. 145, pp. 1517–1527, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0960148119310870

[24] M. Rodríguez, D. Arcos-Aviles, and W. Martínez, "Fuzzy logic-based energy management for isolated microgrid using meta-heuristic optimization algorithms," *Applied Energy*, vol. 335, p. 120771, 2023.

[25] D. M. Lopez-Santiago, E. Caicedo Bravo, G. Jiménez-Estévez, F. Valencia, P. Mendoza-Araya, and L. G. Marín, "A novel rule-based computational strategy for a fast and reliable energy management in isolated microgrids," *International Journal of Energy Research*, vol. 46, no. 4, pp. 4362–4379, 2022.

[26] V. V. Babu, J. P. Roselyn, and P. Sundaravadivel, "Multi-objective genetic algorithm based energy management system considering optimal utilization of grid and degradation of battery storage in microgrid," *Energy Reports*, vol. 9, pp. 5992–6005, 2023.

[27] A. Hussain and H.-M. Kim, "A rule-based modular energy management system for ac/dc hybrid microgrids," *Sustainability*, vol. 17, no. 3, p. 867, 2025.

[28] A. El Zerk and M. Ouassaid, "Real-time fuzzy logic based energy management system for microgrid using hardware in the loop," *Energies*, vol. 16, no. 5, p. 2244, 2023.

[29] H. Binsalim, S. Badaam, A. Ahmed, and M. Majid, "Battery management system for enhancing the performance and safety of lithium-ion batteries," 01 2024, pp. 215–220.

[30] J. Hernández, F. Sanchez-Sutil, F. Muñoz-Rodríguez, and C. Baier, "Optimal sizing and management strategy for pv household-prosumers with self-consumption/sufficiency enhancement and provision of frequency containment reserve," *Applied Energy*, vol. 277, p. 115529, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261920310412

[31] B. Sun, "A multi-objective optimization model for fast electric vehicle charging stations with wind, pv power and energy storage," *Journal of Cleaner Production*, vol. 288, p. 125564, 2021.

[32] M. Khodayar, M. E. Khodayar, and S. M. J. Jalali, "Deep learning for pattern recognition of photovoltaic energy generation," *The Electricity Journal*, vol. 34, no. 1, p. 106882, 2021, special Issue: Machine Learning Applications To Power System Planning And Operation. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1040619020301743

[33] Y. Li, Z. O'Neill, L. Zhang, J. Chen, P. Im, and J. DeGraw, "Grey-box modeling and application for building energy simulations - a critical review," *Renewable and Sustainable Energy Reviews*, vol. 146, p. 111174, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032121004639

[34] D. Crawley, C. Pedersen, L. Lawrie, and F. Winkelmann, "Energyplus: Energy simulation program," *Ashrae Journal*, vol. 42, pp. 49–56, 04 2000.

[35] A. Ferguson, "Esp-r," Apr 2006. [Online]. Available: https://github.com/ESP-rCommunity/ESP-rSource

[36] Modelica Association, *Modelica—A Unified Object-Oriented Language for Systems Modelling, Language Specification, Version 3.3*, Modelica Association, 2012. [Online]. Available: http://www.modelica.org

[37] B. T. Kulakowski, J. F. Gardner, and J. L. Shearer, *Dynamic Modeling and Control of Engineering Systems*, 3rd ed. Cambridge University Press, 2007.

[38] R. L. Woods and K. L. Lawrence, *Modeling and simulation of dynamic systems*. Prentice Hall Upper Saddle River, NJ, 1997.

[39] F. Feng and Z. O'Neill, "Identifying models of hvac systems using arimax," in *2019 ASHRAE Annual Meeting, Kansas City, MO*, 2019.

[40] Y. Zhang, Z. O'Neill, B. Dong, and G. Augenbroe, "Comparisons of inverse modeling approaches for predicting building energy performance," *Building and Environment*, vol. 86, pp. 177–190, 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360132314004405

[41] J. Penman, "Second order system identification in the thermal response of a working school," *Building and Environment*, vol. 25, no. 2, pp. 105–110, 1990. [Online]. Available: https://www.sciencedirect.com/science/article/pii/036013239090021I

[42] S. Wang and X. Xu, "Parameter estimation of internal thermal mass of building dynamic models using genetic algorithm," *Energy Conversion and Management*, vol. 47, no. 13, pp. 1927–1941, 2006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0196890405002311

[43] Y. W. Law, T. Alpcan, V. C. Lee, A. Lo, S. Marusic, and M. Palaniswami, "Demand response architectures and load management algorithms for energy-efficient power grids: a survey," in *2012 seventh international conference on knowledge, information and creativity support systems*. IEEE, 2012, pp. 134–141.

[44] A. T. Dahiru, D. Daud, C. W. Tan, Z. T. Jagun, S. Samsudin, and A. M. Dobi, "A comprehensive review of demand side management in distributed grids based on real estate perspectives," *Environmental science and pollution research*, vol. 30, no. 34, pp. 81 984–82 013, 2023.

[45] D. Stanelyte, N. Radziukyniene, and V. Radziukynas, "Overview of demand-response services: A review," *Energies*, vol. 15, no. 5, 2022. [Online]. Available: https://www.mdpi.com/1996-1073/15/5/1659

[46] D. Dieziger, "Saving money by understanding demand charges on your electric bill." [Online]. Available: https://www.fs.usda.gov/t-d/pubs/htmlpubs/htm00712373/index.htm

[47] N. OConnell, P. Pinson, H. Madsen, and M. OMalley, "Benefits and challenges of electrical demand response: A critical review," *Renewable and Sustainable Energy Reviews*, vol. 39, pp. 686–699, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032114005504

[48] D. Keane and A. Goett, "Voluntary residential time-of-use rates: lessons learned from pacific gas and electric company's experiment," *IEEE Transactions on Power Systems*, vol. 3, no. 4, pp. 1764–1768, 1988.

[49] P. Jazayeri, A. Schellenberg, W. Rosehart, J. Doudna, S. Widergren, D. Lawrence, J. Mickey, and S. Jones, "A survey of load control programs for price and system stability," *IEEE Transactions on Power Systems*, vol. 20, no. 3, pp. 1504–1509, 2005.

[50] F. C. Schweppe, M. C. Caramanis, R. D. Tabors, and R. E. Bohn, *Spot pricing of electricity*. Springer Science & Business Media, 2013.

[51] D. Isogawa, H. Ohashi, and T. Anai, "The role of advance notice in shaping industrial response to time-varying electricity prices," 2024.

[52] M. Albadi and E. El-Saadany, "A summary of demand response in electricity markets," *Electric Power Systems Research*, vol. 78, no. 11, pp. 1989–1996, 2008. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378779608001272

[53] D. Stanelyte, N. Radziukyniene, and V. Radziukynas, "Overview of demand-response services: A review," *Energies*, vol. 15, no. 5, 2022. [Online]. Available: https://www.mdpi.com/1996-1073/15/5/1659

[54] C. Z. El-Bayeh, I. Mougharbel, M. Saad, A. Chandra, S. Lefebvre, and D. Asber, "Novel multilevel soft constraints at homes for improving the integration of plug-in electric vehicles," pp. 1–9, 2018.

[55] A. Vijay and A. Hawkes, "Demand side flexibility from residential heating to absorb surplus renewables in low carbon futures," *Renewable Energy*, vol. 138, pp. 598–609, 2019.

[56] S. Polu, "Residential power usage 3 years data - time-series," 2020. [Online]. Available: https://www.kaggle.com/datasets/srinuti/residential-power-usage-3years-data-timeseries

[57] A. Kannal, "Solar power generation data," 2020. [Online]. Available: https://www.kaggle.com/datasets/anikannal/solar-power-generation-data

[58] J.-M. D., "Smart meters in london," May 2022. [Online]. Available: https://www.kaggle.com/datasets/jeanmidev/smart-meters-in-london

[59] P. Zippenfenig, "Open-meteo.com weather api," latest release of the Open-Meteo weather API. Repository: https://github.com/open-meteo/open-meteo. [Online]. Available: https://open-meteo.com/

[60] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," *CoRR*, vol. abs/1603.02754, 2016. [Online]. Available: http://arxiv.org/abs/1603.02754

[61] F. Borrelli, A. Bemporad, and M. Morari, *Predictive control for linear and hybrid systems*. Cambridge University Press, 2017.

[62] "Gershgorin theorem - encyclopedia of mathematics." [Online]. Available: https://encyclopediaofmath.org/wiki/Gershgorin_theorem

[63] B. Stellato, G. Banjac, P. Goulart, A. Bemporad, and S. Boyd, "OSQP: an operator splitting solver for quadratic programs," *Mathematical Programming Computation*, vol. 12, no. 4, pp. 637–672, 2020. [Online]. Available: https://doi.org/10.1007/s12532-020-00179-2

[64] J. Hespanha, *Linear Systems Theory: Second Edition*. Princeton University Press, 2018, see Section 6.5. [Online]. Available: https://books.google.com/books?id=O3OYDwAAQBAJ

[65] University of California, Berkeley EECS, "Fundamentals of hvac controls," Course Material/Technical Note, 2009, accessed from https://people.eecs.berkeley.edu/~culler/cs294-f09/m197content.pdf.

[66] K. Ogata, *Discrete-time control systems*, 2nd ed. Prentice Hall, 1995.

[67] R. Horn, "C. johnson, matrix analysis. cambridge," 1985.

[68] P. Siano, "Demand response and smart grids—a survey," *Renewable and Sustainable Energy Reviews*, vol. 30, pp. 461–478, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032113007211

[69] L. Von Krannichfeldt, K. Orehounig, and O. Fink, "Combining physics-based and data-driven modeling for building energy systems," *arXiv preprint arXiv:2411.01055*, 2024.

[70] D. Blum, J. Arroyo, S. Huang, J. Drgoňa, F. Jorissen, H. T. Walnum, Y. Chen, K. Benne, D. Vrabie, M. Wetter, and et al., "Building optimization testing framework (boptest) for simulation-based benchmarking of control strategies in buildings," *Journal of Building Performance Simulation*, vol. 14, no. 5, pp. 586–610, 2021. [Online]. Available: https://doi.org/10.1080/19401493.2021.1986574

[71] T. M. Inc., "Matlab system identification toolbox," 2024.

[72] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, "Scikit-learn: Machine learning in python," *Journal of machine learning research*, vol. 12, no. Oct, pp. 2825–2830, 2011.

[73] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems 32*. Curran Associates, Inc., 2019, pp. 8024–8035. [Online]. Available: http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf

[74] N. Somu, G. Raman M R, and K. Ramamritham, "A deep learning framework for building energy consumption forecast," *Renewable and Sustainable Energy Reviews*,

vol. 137, p. 110591, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032120308753

APPENDICES

APPENDIX A

Forecasting Methodology Development

A.1 Introduction

This appendix provides an overview of the development of the forecasting methodology implemented in Chapter 3.

A.2 Data Processing and Feature Engineering

The primary data source was the Smart meters in London dataset [58], which contained half-hourly household power usage and hourly weather data (via the Dark Sky API), along with bank holiday information.

Weather data were augmented to a half-hour frequency using quadratic interpolation; descriptive columns adopted values from preceding rows. The relevant columns were extracted and the descriptors were converted to numerical values, resulting in the following characteristics: visibility, temperature, dew point, pressure, wind speed, type of precipitation, and humidity.

The power consumption data was normalized by calculating the average power consumption per household for each time interval (total consumption divided by the number of active households). This average consumption was then scaled to values between zero and one using scikit-learn's MinMaxScaler [72].

Feature engineering involved decomposing timestamps into year, month, day, hour, and minute components to capture seasonality and daily patterns. Binary indicators for UK holidays and weekends were added to reflect periods of likely higher home occupancy and consumption.

The data set was chronologically divided into 70% for training, 10% for validation, and 20% for testing to prevent data leakage and provide realistic estimates of future performance.

A.3 Machine Learning Models

Several models were evaluated to forecast average household power consumption.

XGBoost

XGBoost [60], a commonly used gradient boost method for time series, was used as a baseline. Feature importance analysis (Figure A.1) was performed using XGBoost's built-in functionality, initially with a shortened input window (most recent observation) to better assess the contribution of nonpower features such as weather and date, as recent power usage heavily influenced predictions. Subsequently, the model was trained using full power window, weather, and time data.

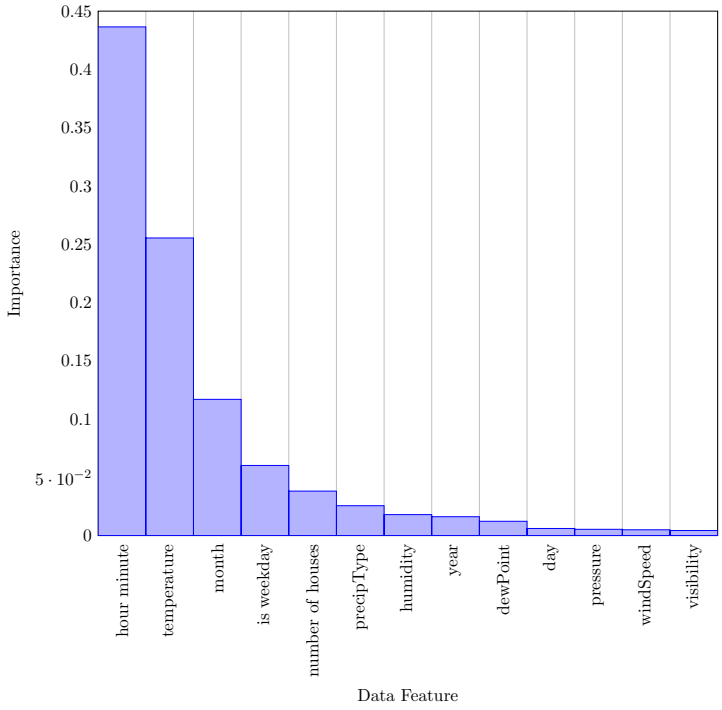

Fig. A.1: Importance scoring of each state.

Figure A.1 visually depicted the contributions of each feature to the predictions and formed a basis for the input features to each of the machine learning models. Next, I used the full power window, weather, and time data to predict the average power consumption

to serve as a baseline model for the neural network based models.

Feed Forward Networks

Neural networks were implemented in PyTorch [73].

Deep Residual Network (ResNet): Due to the complexity of time series forecasting, a deep network architecture was explored. Residual connections were used across 10 layers (approx. 1000 nodes/layer) to mitigate vanishing gradients. The architecture (Figure A.2) concatenated layer outputs with inputs but resulted in millions of parameters.

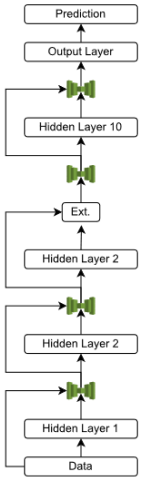

Fig. A.2: Deep Residual Network Architecture

Convolutional Network (CONV): A 1D convolutional layer was added before the dense layers to extract features from the time-dependent input window, leveraging temporal locality, as in [74]. The architecture was depicted in Figure A.3.

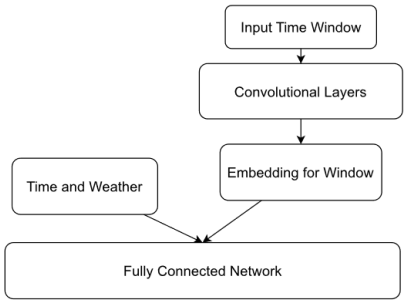

Fig. A.3: CONV Network Architecture

Convolutional LSTM (CONV-LSTM)

Building on the CONV network, an LSTM layer was incorporated to explicitly model the sequential dependency in the 24-hour (48 half-hour steps) output window. Looking at each half-hour prediction, all subsequent predictions depended on previous ones, which made LSTM suitable due to its cyclical inference method. The architecture (Figure A.4) used a CONV layer for initial feature extraction, followed by a linear layer feeding into the LSTM, whose output was then decoded by linear layers into the final forecast.

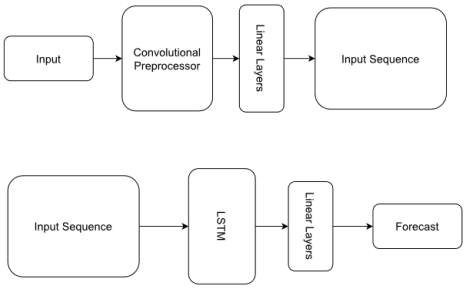

Fig. A.4: CNN-LSTM Architecture

A.4 Results and Discussion

Model performance was evaluated visually using predictions for the last day of the test set (furthest from training data) and quantitatively using L1 loss, L2 loss, and explained variance over the entire test set.

XGBoost: Followed the actual consumption curve well and matched peak magnitudes accurately but tended to over-predict during the convex region between morning and evening

peaks, which likely indicated fitting to the average of highly variable behavior in this period (Figure A.5).

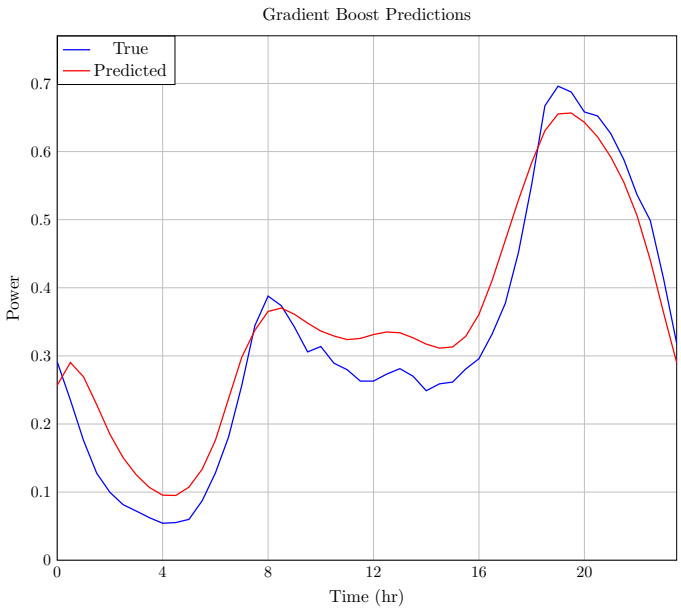

Fig. A.5: Final 24 hour prediction results from Gradient Boosted Trees.

Deep ResNet: Largely underestimated consumption and produced overly smooth approximations, possibly due to the vast number of parameters and lack of explicit feature pre-processing that hindered learning before overfitting occurred (Figure A.6).

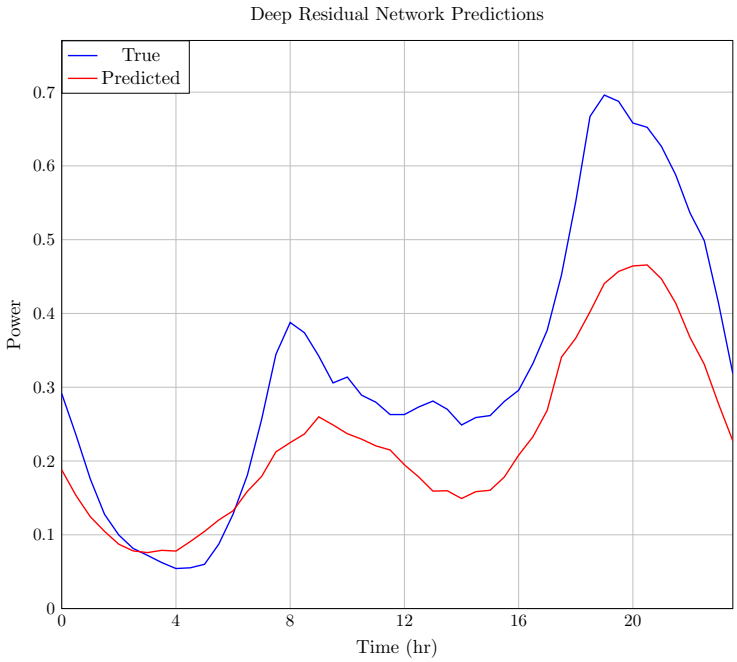

Fig. A.6: Final 24 hour prediction results for the Deep Residual Network.

CONV Network: Showed significant improvement over the ResNet, validating the utility of the CONV layer for feature extraction. It followed the general curve well but underestimated peak magnitudes compared to XGBoost (Figure A.7).

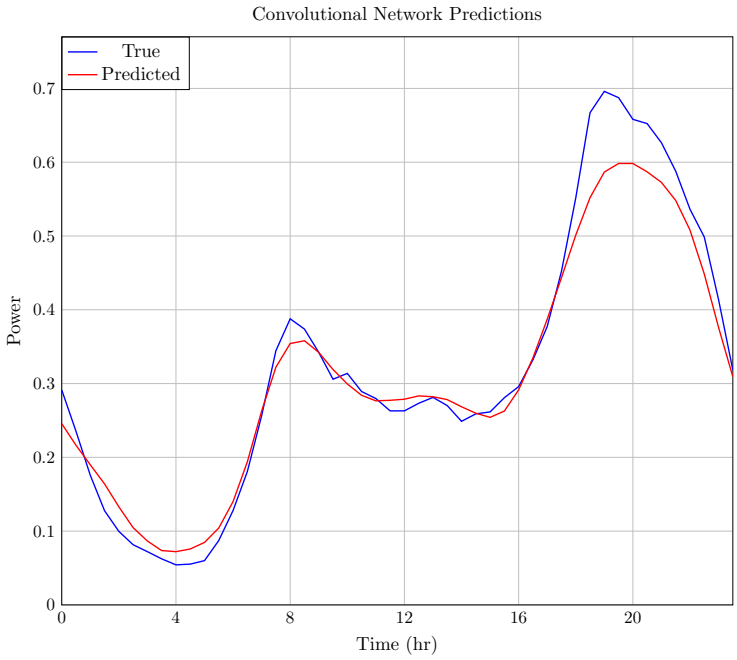

Fig. A.7: Final 24 hour prediction for network with convolutional pre-processing

Figure A.3 showed the result of adding the convolutional preprocessor before the dense layers of the forecasting network. As was seen, the CONV layer greatly increased the accuracy of a simple feed forward network. The CONV network followed the curve of the actual power consumption, validating my assumption that a convolution through the input window performed meaningful feature extraction. However, the CONV network still underestimated the peaks in the output window whereas the XGBoost did not.

CONV-LSTM: Achieved better accuracy at peaks compared to the CONV network, confirming the value of LSTM for modeling output sequence dependencies (Figure A.8).

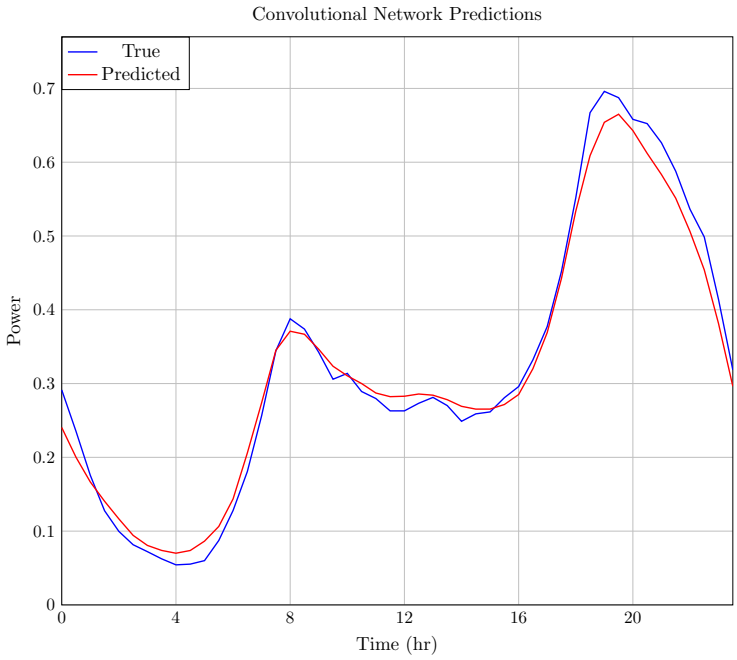

Fig. A.8: Final 24 hour prediction for CNN-LSTM network

Figure A.8 showed the results of the CONV-LSTM forecaster. As was seen, the CONV-LSTM network was significantly more accurate at the peaks of the output than the CONV network. This validated my assumption that explicit knowledge of the previous predictions provided by the LSTM was valuable for power consumption forecasting. The results also solidified the usefulness of a CONV layer in time series prediction problems because the accuracy of the CONV-LSTM improved upon the CONV network's accuracy.

Figure A.9 in Appendix A showed all of the above results on a single graph for better comparison. Additionally, figure A.10 showed the error for each method to the true output.

Accuracy Metrics: Table A.1 summarized the performance metrics. Based on L1 and L2 loss, the ranking (best to worst) was CONV-LSTM, XGBoost, CONV Network, Deep ResNet. However, XGBoost slightly outperformed CONV-LSTM in explained variance, potentially indicating the CONV-LSTM might have been slightly overfit.

Training Details: The Adam optimizer was used for neural networks. Dropout (0.05) was necessary for the CONV-LSTM to improve validation accuracy. Training parameters

Table A.1: Loss results of All Methods

Loss Type	XGBoost	Deep Residual Network	CONV Network	CONV-LSTM
L1 Loss	3.268E-2	8.434E-2	3.520E-2	2.293E-2
L2 Loss	1.765E-3	1.244E-2	2.100E-3	1.637E-3
Explained Variance	.9654	.8011	.9476	.9601

were in Table A.2.

Table A.2: Table of training parameters for each model

Model	Optimizer / Estimators	lr	Weight Decay	Dropout
XGBoost	1000	0.01	n/a	n/a
Deep Residual	Adam	1E-4	1E-5	No
CONV Net	Adam	1E-4	1E-4	No
CONV-LSTM	Adam	1E-4	1E-5	0.05

Table A.2 showed the training parameters for the various models. For all of the neural network models, I used the Adam optimizer in Pytorch because of its versatility and adaptive properties. The LSTM was the only model requiring dropout because of its tendency to overfit to the data. I found that dropout in the linear layers and a small amount of dropout in the LSTM greatly improved the validation accuracy of CONV-LSTM.

A.5 Comparison Plots

Final Model Selection: Although CONV-LSTM showed the best accuracy based on L1/L2 loss, XGBoost [60] captured trends well and tended to over-predict consumption. For the main hierarchical control framework aimed at minimizing power consumption, a conservative forecast (predicting slightly more consumption than actual) was preferable. Therefore, the XGBoost algorithm was selected for implementation in Chapter 3.

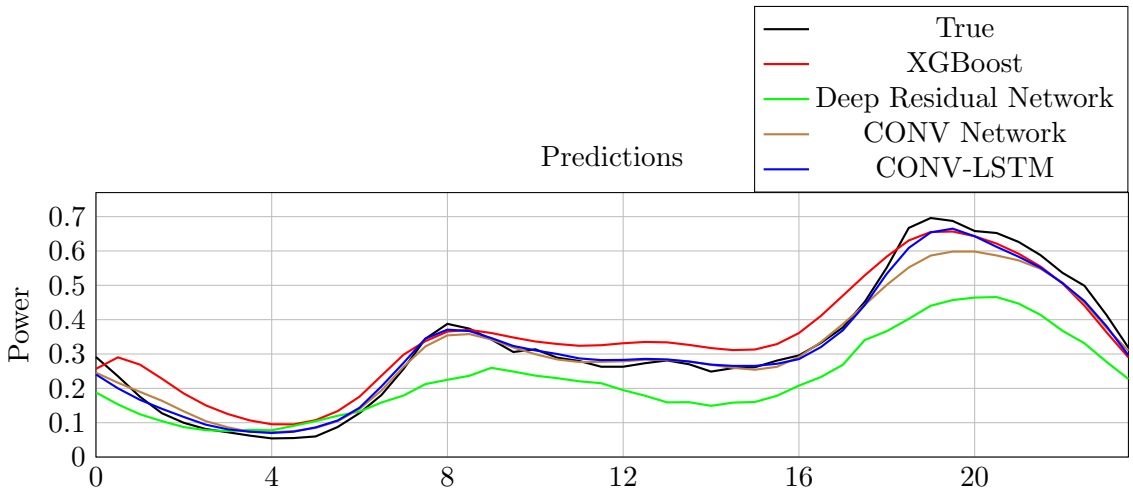

Fig. A.9: Last 24 hour prediction results for all models.

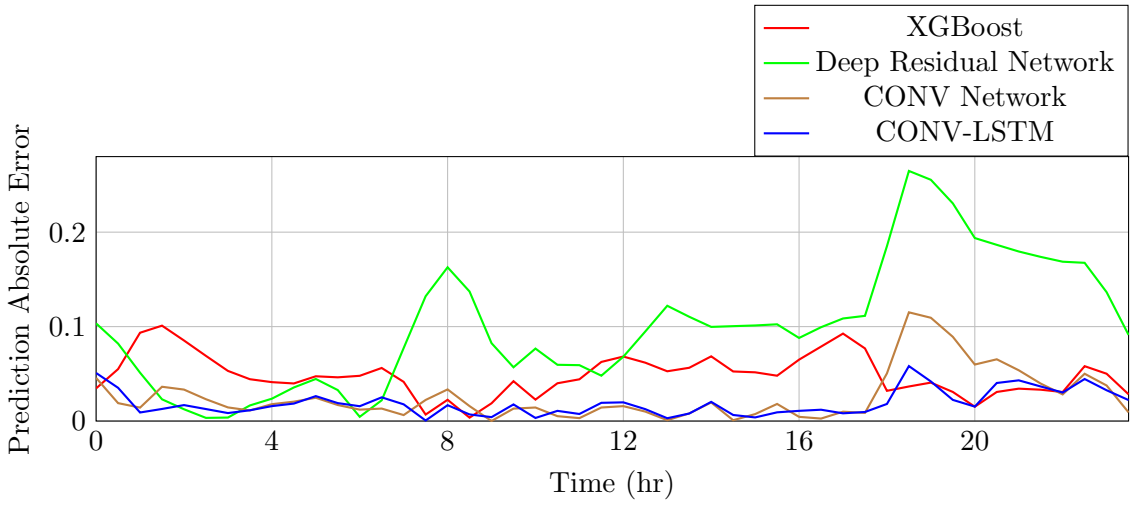

Fig. A.10: Error in prediction results for all models.