Towards Model-Driven Heat Pump Control in a Multi-Story Building

Imran Riaz Hasrat[a,**], Kamilla Heimer Andersen[b], Peter Gjøl Jensen[a], Rasmus Lund Jensen[b], Kim Guldstrand Larsen[a] and Jiří Srba[a]

[a]Department of Computer Science, Aalborg University, Aalborg 9220, Denmark
[b]Department of The Built Environment, Aalborg University, Aalborg 9220, Denmark

ARTICLE INFO

Keywords:
energy flexibility
radiant floor heating
heat pump control
model identification
strategy synthesis
model approximation

ABSTRACT

Domestic heating systems exhibit significant energy flexibility when integrated with heat pumps and hot-water buffer tanks, especially with fluctuating day-ahead energy prices. However, optimizing these systems in large buildings with shared resources poses crucial challenges. While most existing methods target single-room or single-family houses, this study delves into complexities within a three-story building housing six apartments. The building's heating system comprises a hot water buffer tank, mixing loop, floor heaters, and a Ground Source Heat Pump (GSHP) controlled by a weather-compensated control strategy (WCS). Our approach aims to tackle challenges like integrating real sensor data, scalability, varying weather effects, and diverse resident heat use preferences. We employ CTSMR software to identify thermal behaviour and use reinforcement learning to design an intelligent Uppaal Stratego controller. Our results reveal a 43% reduction in energy costs while maintaining comfort levels compared to WCS. The temporal validity of estimated thermal models is also analyzed.

1. Introduction

1.1. Motivation

Space heating is the primary use of energy across European countries [1] for domestic building environments, with renewable energy sources meeting a significant portion of this demand [2]. Integrating renewable energy sources into space heating systems is essential for meeting sustainability goals and minimizing environmental impact. However, the inherent variability in renewable energy availability poses substantial challenges to exploiting these sources effectively for heating purposes. Despite the growing trend in the adoption of renewable energy, there remains unused potential due to the limited capabilities of traditional heating systems. Developing advanced heating solutions capable of adapting to the variable availability of renewable energy sources promises to optimize heating operations and reduce reliance on conventional energy sources.

Multi-story buildings are becoming more popular in urban areas due to their capacity to tackle urbanization issues, maximize space efficiency, and support sustainability. Efficient heating control in such large buildings becomes more challenging as the number of apartments increases. The diverse heating needs and usage patterns of occupants in different apartments pose significant challenges. Additionally, apartments are subject to varying wind and weather conditions depending on their position within the building. Moreover, the buildings' dynamics change over time, further complicating the situation. The overarching question is whether introducing more intelligent control solutions for such buildings can efficiently address these challenges. The solutions must consider the evolving and unpredictable factors that affect the operation of heating systems, especially when apartments share common resources like a heat pump and a buffer tank. Various control solutions for space heating systems have been proposed recently, primarily focusing on single rooms [3] or individual houses [4, 5]. However, approaches that can scale up to control the complex heating system of multi-story buildings with multiple apartments still need to be explored. These approaches must confront the modelling complexities and achieve effective heating control in buildings with occupants who have diverse heating demands and habits. Moreover, the approaches must also account for stochastic weather effects on the performance of the heating systems. Furthermore, they must address challenges related to real sensor data, evolving thermal dynamics, storage tank models, and other computational complexities found in large-scale buildings.

This study uses sensor data collected from an existing multi-apartment building equipped with a heat pump-based hydronic floor heating system. Initially, we simplify the complexity of this large system while preserving essential information. Our approach involves decomposing the system apartment-wise rather than room-wise. We adopted this method because we observed that most of the rooms within an apartment often have similar indoor temperatures, reducing unnecessary complexities. We describe the thermal dynamics of the apartments using Ordinary Differential Equations (ODEs), which capture the intricate thermal behaviour of the building. These equations consider all energy balance factors and incorporate constant resistance and capacitance parameters that require estimation. We employ grey-box modelling facilitated by CTSMR software [6] to identify apartment thermal dynamics. In addition to the heat pump, the building's heating system includes a hot water buffer tank and a mixing loop. The thermal model, buffer tank, and mixing loop model are then utilized to develop an intelligent controller using UPPAAL STRATEGO [7] tool. We use the term 'intelligent' for our controller because it learns strategies using machine learning techniques. It optimizes the trade-off between cost and comfort by considering weather forecasts and future electricity prices in its

imranh@cs.aau.dk (I.R. Hasrat)

learning process. Additionally, it offers online and automatic strategy synthesis. This controller integrates advanced control strategies to enhance user comfort and uses the energy flexibility potential of the building's heating system. The primary contributions of the paper include:

1. Transitioning from a small family house to a larger multi-apartment real-world building, we address additional practical challenges related to scalability and applicability. We utilize historical data continuously recorded from different heating components (through installed sensors) in the building where the occupants are already residing in the apartments.
2. We employ data-driven grey-box modelling approaches in CTSMR to estimate the thermal dynamic model and assess its accuracy against various standard error metrics.
3. We analyze the temporal validity of the estimated thermal model over extended time horizons and recommend regular re-estimation of the constant coefficients as the building dynamics evolve over time.
4. We design a UPPAAL STRATEGO controller to synthesize near-optimal strategies for controlling the heat pump, followed by a performance comparison with an already implemented weather-compensated strategy in the building to determine how much heating cost reductions real inhabitants/occupants of the apartments can gain.

1.2. Related Work

The literature discusses various approaches to tackling the domestic heating problem. We specifically focus on the quality of modelling for the main components, including thermal dynamics, thermal storage devices, and controllers, to ensure the efficient operation of domestic heating systems.

Model identification is a crucial component in optimizing heating systems as it facilitates identifying a model that reflects the thermal behaviour of that system. Incorrect model identification poses significant practical challenges for heating system optimizations [8]. Several studies [9, 10, 11, 12, 13, 14, 5, 15, 16, 17] have proposed grey-box modelling [18] as a sophisticated approach for model identification, given its reliance on minimal data and a few mathematical equations. Volger-Finck et. al. [9] represented the thermal dynamics of heating systems as a set of restricted linear equations and performed their model identification using the MATLAB System Identification toolbox [19]. They observed a significant reduction in CO_2 emissions by demonstrating the effectiveness of their approach on three family houses of varying sizes and ages. The studies [10, 11, 12] developed reduced-order energy building models to forecast the thermal behaviour of buildings. Attoue et. al. [13] explored different thermal conditions to analyze the optimal order of grey-box models for predicting short-term building thermal behaviour. They investigated how the heating conditions are important in shaping the performance evaluation of grey-box models. Moreover, Nespoli et. al. [14] utilized grey-box models to estimate building thermal dynamics for HVAC control and load management. Furthermore, studies conducted in [5, 15, 16] presented the thermal dynamics of an experimental single-family house as a set of Ordinary Differential Equations (ODEs) and employed CTSMR software to estimate the hidden heat transfer coefficients including resistance and capacitance. We modify the model identification approach outlined in [16] to scale it to model the thermal behaviour of a real multi-apartment building with occupants exhibiting diverse heating demands and habits. Existing studies witness that thermal dynamics estimated in one season do not perform well in another [20, 21, 22]. However, determining the exact duration for which the estimated thermal dynamics remain reliable still requires further exploration. Therefore, we analyze the temporal validity of the estimated heat coefficient to determine the frequency at which we should re-estimate these coefficients.

Hot water buffer tanks are another crucial component that target energy flexibility in heating systems. These tanks receive hot water from the heating element, such as a boiler or heat pump, and cold water returning from the radiators. The studies [23, 24, 25, 26, 27] utilized a straightforward one-mass model, which involves considering the average temperature of the tank water. The one-mass model often leads to significant discrepancies in temperature estimation, especially for larger tank sizes. Some other studies [28, 29] proposed two-mass models, which proved a better approach than one-mass models as the temperature calculation accounted for an additional water layer within the tank. Furthermore, the studies [30, 31] introduced a fixed number of multiple layers in the buffer tanks, which proved beneficial but could still not accommodate tanks of all sizes. Hasrat et. al. [32] introduced a dynamic buffer tank model, showcasing the ability to incorporate any number of virtual water layers to manage temperature stratification, thereby adapting to the size of the tank. Existing approaches have been shown to model buffer tanks for single rooms or small family houses. However, we scale the dynamic approach proposed by [32] and examine its applicability to when multiple apartments share the same buffer tank. The novelty of this approach lies not in introducing a larger tank, but in synthesizing the control for a shared tank used by six apartments to meet their heating needs. The challenge arises from the competition for this shared resource, necessitating global synchronization or communication between the apartments and the controller. Developing strategies for the controller to determine the optimal amount of water to store in the shared tank, without exceeding the heat demand capacity, while ensuring proper synchronization, is a challenging task.

Implementing effective control strategies is pivotal to exploring the energy flexibility potentials in heating systems. Model Predictive Control (MPC) is the most common and robust choice for executing such strategies. Larsen et. al. [4] focused on enhancing user comfort in a family house with a floor heating system. They employed the UPPAAL STRATEGO tool [7] and introduced an online and compositional MPC synthesis approach. Similarly, Hasrat [16] also applied a similar online strategy synthesis approach to operate the

heat pump in an experimental house, aiming to optimize a multi-objective involving energy cost and user comfort. They introduced various control strategies to achieve this multi-objective optimization goal. On the other hand , the approach proposed by Golmohammadi et. al. [5] utilized the same experimental house but designed an Economic Model Predictive Controller (EMPC) explicitly focusing on optimizing energy costs. Moreover, the studies [33, 34] presented an experimental investigation and control optimization of a ground source heat pump (GSHP) system, resulting in improved system stability and substantial energy savings. Furthermore, Potocnik et. al. [35], optimized the operation of weather-controlled air-to-water heat pumps to enhance thermal comfort in residential buildings, with findings showing improved indoor comfort levels through offline optimization of generalized heating curves incorporating outdoor temperature and solar radiation. Similarly, Chen et. al. [36] introduced an MPC to enhance comfort and energy efficiency in radiant floor heating systems. Additionally, Park et. al. [37] concentrated on optimizing heating system operation under various load conditions by predicting optimal airflow rates. The existing MPC approaches were designed to enhance heating systems in small experimental or real family houses. However, we adapt the online strategy synthesis approach proposed in [16] for a larger real building comprising six apartments with real occupants. Our proposed controller considers variable electricity prices and weather forecasts, learning control decisions to adjust heat use in response to fluctuating energy prices.

Reinforcement learning (RL) [38] is an emerging control technique that can be utilized to optimize energy flexibility in domestic heating systems, offering an alternative to MPC. RL enables systems to learn and adapt through trial and error in complex environments. Wang et. al. [39] highlighted RL's advantages for building controls under varying energy demands but noted challenges like long training times and limited real-world applications. They recommended focusing on faster training, improved control robustness, and developing open-source benchmarks for RL in real buildings. Zhao et. al. [40] showed that deep reinforcement learning (DRL) could optimize household device scheduling, resulting in substantial cost reductions. Langer et. al. [41] demonstrated DRL's superiority in smart home heating, achieving high self-sufficiency and reduced CO_2 emissions. Lissa et. al. [42] used DRL for smart home energy management, achieving significant energy savings and enhanced renewable energy use while maintaining comfort.

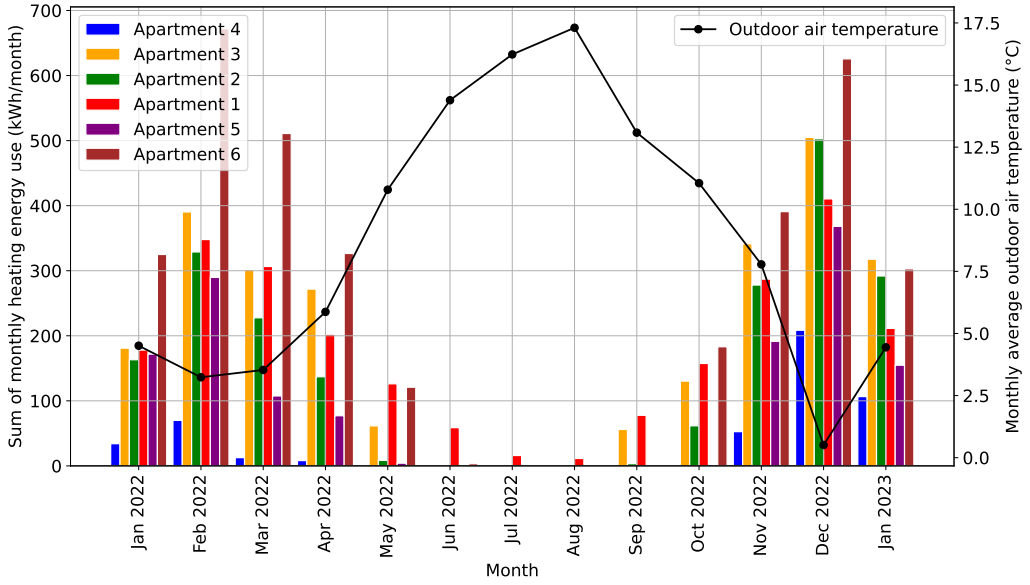
Combining RL with MPC leverages the strengths of both techniques. RL enhances MPC by adapting to changes, handling nonlinearities, and exploring new strategies, making it scalable. Arroyo et. al. [43] proposed a hybrid RL-MPC algorithm, which improved performance and constraint satisfaction in simulations compared to using RL or MPC alone. Inspired by the effectiveness of hybrid RL-MPC approach, we apply it to a domestic heating system. Additionally, we propose an online strategy synthesis approach in which strategies are periodically synthesized at regular intervals. This approach effectively addresses the challenge of long training times, accelerating the training process and rendering it applicable in real-time scenarios.

In summary, current methods for addressing heating problems lack a comprehensive tool-chain for end-to-end solutions in domestic heating applications. This tool-chain should include model identification, MPC design, and advanced control strategy synthesis to optimize energy cost and comfort. Most existing methods rely on offline learning [44, 45, 46], which struggles to adapt to real-time changes like user behaviour and weather. Thus, there is a need for online, automated learning approaches that use real-time data for periodic control decisions. There is limited research on combining RL and MPC for domestic heating [43, 47, 48], suggesting potential synergies are unexplored. Current solutions often use static thermal dynamics models [5, 16], ignoring changes over time, indicating a need for more dynamic and frequently updated models. Additionally, many studies neglect the role of buffer tanks and mixing loops in heating systems. Developing adaptive models and intelligent control for these components could improve energy flexibility and comfort. A generic buffer tank model applicable to various tank sizes remains a necessary advancement.

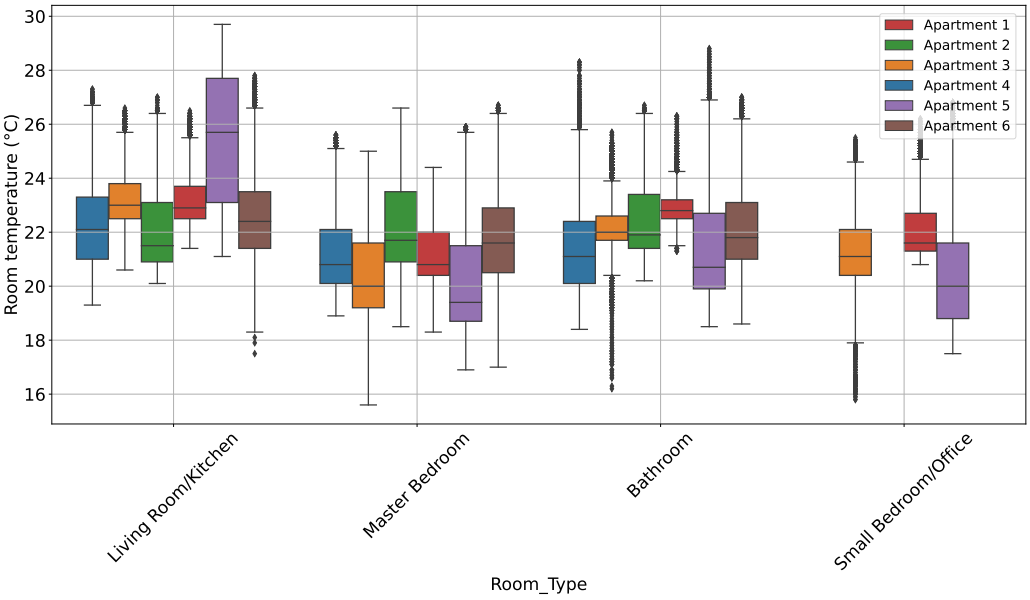
In order to address the above-mentioned issues, we propose a comprehensive tool-chain/framework for modelling and controlling different components of domestic heating systems e.g., heat pumps, buffer tanks, and mixing loops. The tool-chain combines the strengths of grey-box models, the modelling capabilities of stochastic timed automata, the control design of MPC, and the power of RL. Grey-box models are convenient because they incorporate the empirical adaptability of black-box models with the mathematical aspects of white-box models in a balanced manner. Online training with RL helps reduce CPU utilization time and memory requirements for strategy synthesis, compared to offline training where a single strategy is synthesized over a much longer period. Stochastic timed automata are useful for modelling the behavior of complex and stochastic systems, such as heating systems. The objective of the tool-chain is to surpass current state-of-the-art approaches in lowering energy costs while maintaining comfort.

1.3. Structure of the Paper

The paper is structured as follows: Section 2 introduces the case study, highlighting its significance and context. Section 3 outlines the methodology, introduces the tools utilized, and details the thermaldynamics (mathematical modelling) of the building. Section 3 also describes the identification process in CTSMR and the design of the UPPAAL STRATEGO controller. Section 4 presents the experimental evaluation, discussing the implementation and results of the experiments conducted, providing insights and interpretations of the results. Finally, Section 5 includes concluding remarks, summarizing the key findings.



(a) Monthly heating energy use by apartments along with outdoor air temperature



(b) Yearly room temperature by room type and apartment for 2022

Figure 1: An overview of heating use and indoor temperatures for apartments

2. Case Study - A Multi-Story Building

This case study is a multi-story residential building built in 1949/50, located in an urban area in Northern Denmark [49, 50]. The building was renovated in 2012/13 to be considered a low-energy building. The building is three stories high with a basement and attic with a total floor area of 2160 m^2. There are 24 apartments divided over five entrances. The deep renovation included the installation of Ground Source Heat Pumps (GSHPs) in the basement of each entrance, which supplies hydronic floor heating to all rooms. This study includes one staircase (located on the right-side with one external wall facing the West orientation) consisting of six apartments, three apartments of 55 m^2 (to the right) and 72 m^2 (to the left). Table 1 outlines key details for each apartment, including its number, floor placement, and total rooms.

The available building data is from January 19, 2022 to January 19, 2023 with a 5-minute resolution, extracted through a dedicated REST API:

1. Power (kW, apartment level)
2. Room temperature (°C, room level)
3. Heating set point (°C, room level)
4. Mass flow (m^3/s, apartment level)

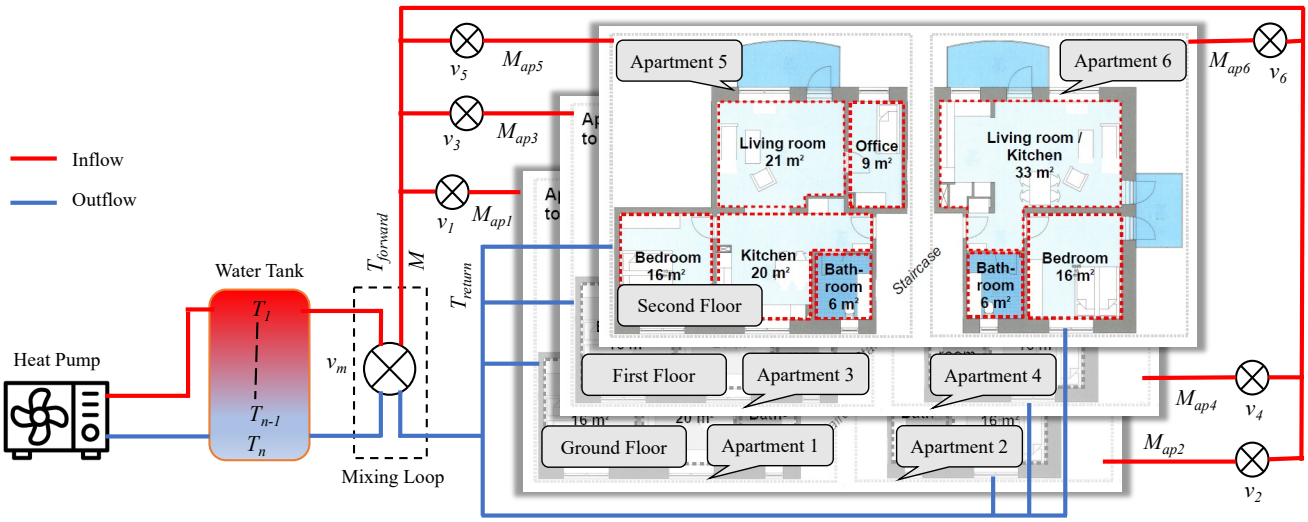


Figure 2: Overview of the designated section of the building and its associated heating system

Table 1
Placement and room count for apartments

Apartment	No. of Rooms	Location
Apartment 1	4	ground floor
Apartment 2	3	ground floor
Apartment 3	4	first floor
Apartment 4	3	first floor
Apartment 5	4	second floor
Apartment 6	3	second floor

5. Supply water temperature (°C, apartment level)
6. Return water temperature (°C, apartment level)

The available weather data is downloaded from Denmark Meteorological Institute (DMI) open API (`https://www.dmi.dk/kontakt/frie-data/`) with 60-minute resolution:

1. Outdoor air temperature (°C)
2. Global solar radiation (Wh/m^2)

Figure 1a illustrates the monthly heating energy use and average outdoor air temperature, highlighting variations in heating usage across apartments. These differences result from the heating set points determined by the occupants, internal gains (occupancy and use of equipment), solar radiation (and internal solar shading), and the apartments' placement on different floors.

Figure 1b presents the measured room temperatures, categorized by room type within each apartment throughout the year. It can be seen that the average room temperatures are generally consistent across apartments, except for Apartment 5 for the living room/kitchen. Located on the second floor, Apartment 5 lacks internal solar shading on its south-facing living room/kitchen facade, distinguishing its room temperature profile. Furthermore, the apartments on the left include an extra small bedroom/office, whereas those on the right do not.

In Figure 2, the floor plans of the apartments and associated heating systems are described. It can be seen that each story/floor consists of two apartments of different sizes, while the structural layout remains consistent across all floors. All the rooms in the apartments are fitted with floor heating. The red and blue lines indicate the circulation of the forwarding hot ($T_{forward}$) and returning colder water (T_{return}), respectively. The space heating system can be divided into the following main components:

- heat pump (controlled in 11 intensity levels) dimensioned power at 10 kW (see Section 4.2 for type of control),
- hot water buffer tank 200 liters,
- mixing loop (controlled in 3 levels),
- room thermostats and
- binary (open/closed) valves ($v_1, \cdots, v_6$).

The heat pump extracts heat from the ground to warm the water, which is then directed to the hot water buffer tank. The hot water buffer tank is an essential component for thermal storage. This tank efficiently stores heated water from the heat pump, ensuring a readily available supply for space heating. A dedicated buffer tank is used for Domestic Hot Water (DHW). This paper focuses on addressing the space heating issue, with the DHW aspect to be explored in future work. Acting as a buffer helps optimize the heat pump's performance, providing a consistent and responsive source of hot water when needed. The tank also receives colder water from the building, leading to produce temperature variations ($T_1, \cdots, T_n$) in different tank regions. The mixing loop blends a portion of colder water returning from the building with the hot water coming from the hot water buffer tank. This process helps achieve a controlled and balanced forward water temperature ($T_{forward}$). By carefully combining hot and cold water streams, the mixing loop ensures that

the supplied water maintains an appropriate temperature, optimizing comfort and energy efficiency in the heat pump system. Individual room temperatures are regulated through thermostats in each room, allowing occupants to control the climate independently.

Observing the historical data of the building, we noticed that the indoor temperature in most rooms of an apartment is in a close range. These indoor temperatures were produced when the heating system tried to achieve the relevant target temperature for each room. This similarity in room temperatures within an apartment makes it unnecessary to manage distinct set points for individual rooms, helping to avoid added complexity. Therefore, we simplify the problem by computing apartment-wise set points which is the weighted average of the set points assigned to individual rooms within each apartment. This implies larger rooms carry a greater weight in the calculated average than smaller rooms. Some apartments, such as apartments 3 and 5 (see Figure 1b), have bedrooms that are relatively colder than other rooms. Managing the heating system at the apartment level not only reduces system complexity, which benefits both model estimation and strategy synthesis but also provides an accurate estimation of the overall heat demand for the entire apartment. The system estimates heat demand based on the average setpoint temperature. However, since each room is equipped with a thermostat valve controlled by a traditional bang-bang (BB) controller, it ensures that colder rooms do not become overheated. In our model, individual room thermostats are not explicitly modelled. Instead, binary valves $(v_1, \cdots, v_6)$ are introduced, one for each apartment. When the temperature across all rooms in an apartment exceeds the average set point, indicating no heating demand, all room thermostats and their corresponding binary valve remain closed. Conversely, when the temperature in any room falls below the set point, signifying a need for heat, the thermostat opens, triggering the respective binary valve to allow hot water with a fixed mass flow for that apartment (denoted by $M_{ap1}, \cdots, M_{ap6}$). Recall that we modelled the floor-heating room thermostats as BB controllers. In this control setting, the deadband is implicit, with room temperature changes reported after a 15-minute control interval. The room thermostats are assumed to be simple mechanical devices, with no smart adjustment or explicit deadband estimation.

In this heating system, we control the heat pump and mixing loop with our intelligent UPPAAL STRATEGO [7] controller. The controller decides the heat pump intensity from 10 levels and selects the cold and hot water proportions for blending (via mixing valve v_m) to achieve a desired forward water temperature. The controller aims to minimize energy use costs while maintaining user comfort.

3. Methodologies

In Figure 3, we illustrate our methodological framework. In the first step, we formulate the thermal dynamics as ordinary differential equations (ODEs) that capture the building's physics. The historical sensor data and thermal dynamic equations are utilized in a grey-box modeling approach within CTSMR [6] to estimate constant heat transfer coefficients. This grey-box model refines the hidden coefficients by identifying the optimal fit, thereby providing constants that define how heat is exchanged between rooms and heaters.

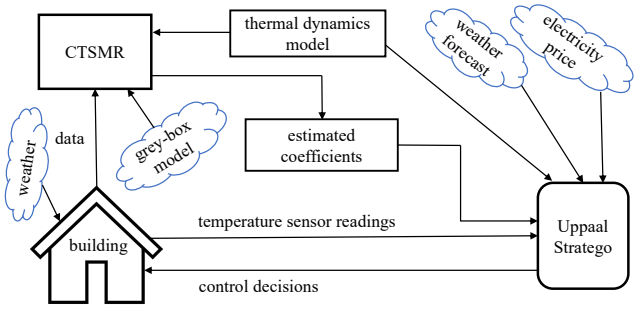


Figure 3: Methodological overview of our approach

In CTSMR, we perform estimation computations individually for each apartment to prevent computational challenges arising from a large number of variables. Subsequently, these individual models are reassembled to construct the entire system. Following the coefficient estimation, we design a Model Predictive Controller (MPC) in UPPAAL STRATEGO. This controller encapsulates the thermal dynamics model and instantiates an Euclidean Markov Decision Process (EMDP) [51]. The controller takes as input, current sensor readings, weather forecasts, and day-ahead electricity prices to learn near-optimal control strategies for efficiently operating the heat pump and mixing loop.

It is important to note that UPPAAL STRATEGO predictive model incorporates a degree of noise. For instance, while the house experiences actual weather conditions, the controller relies on weather predictions for decision-making.

3.1. An Overview of Modelling Tools

CTSMR: The CTSM software is designed for conducting continuous-time stochastic modelling [52]. It is developed at the Technical University of Denmark's (DTU) Informatics and Mathematical Modelling Department. It is available in R software as a package named CTSMR [6]. It can use grey-box as well as black-box models to describe dynamic systems using stochastic differential equations. We focus on grey-box models because they are more precise for buildings when mathematical equations and reasonable amount of historical data are available. It provides methods for estimating unknown parameters in both linear and nonlinear models by incorporating maximum likelihood (ML) and maximum a posteriori (MAP) methods for model identification and hidden parameter estimations of real-time stochastic models.

UPPAAL STRATEGO: The UPPAAL tool suite [53, 54] is a robust integrated environment rooted in formal methods specifically designed for the model checking, validation,

and formal verification of real-time systems modelled as concurrent timed automata that communicate.

UPPAAL STRATEGO [7] is a branch of the tool suite that incorporates built-in reinforcement-learning techniques for cost optimization and strategy synthesis of stochastic and hybrid game systems. UPPAAL STRATEGO has successfully demonstrated its practical use for formal verification, performance analysis, and strategy synthesis of cyber-physical systems, including applications like heating systems, cruise control, traffic control [55, 4, 56, 15]. Within STRATEGO tool, systems are modelled as networks of automata processes communicating through channels or shared variables. The tool provides a descriptive language, incorporating bounded integers, arrays, clocks, and functions to articulate the system behaviour. The real-time clock facility of the tool facilitates the representation of differential equations as continuous variables, capturing the crucial timing aspects inherent in real-time systems.

3.2. Data Cleaning and Processing

The data is cleaned by interpolating missing values (less than 5%) and removing duplicates. Outliers were identified and handled appropriately to ensure the data's accuracy (removed with Interquartile Range (IQR) analysis). Additionally, we converted all timestamps to UTC format for consistency across the dataset. The heating power is calculated to energy use and resampled where applicable, ensuring uniform data intervals. In this study, we utilized data with a 5-minute resolution to model the building's thermodynamics. We did not encounter missing data exceeding a few hours. However, if such gaps do occur, we will refrain from training the model during those periods. Instead, we will use data from older periods with complete information for training. Using data from an earlier period also remains equally useful, as retraining of the thermal dynamics is needed only every six weeks, not daily (see Section 3.6).

3.3. Mathematical Modelling of the Building

We express the thermal behaviour of the building components as dynamic thermal characteristics. Identifying these thermal characteristics is crucial to adequately capture their behaviour when interacting with variable conditions, such as varying heat flow and changing weather conditions. These thermal characteristics can be expressed as Ordinary Differential Equations (ODEs), combining constant heat transfer coefficients (representing resistance and capacitance) with thermodynamic variables. It means, these ODEs form a resistor-capacitor (RC) network [57], utilizing sensor data such as indoor air temperature, hot water mass flow and temperature, outdoor temperature, and solar radiation.

$$C_i^j \times \frac{d\tilde{T}_i^j}{dt} = \frac{(\tilde{T}_h^j - \tilde{T}_i^j)}{R_{i,h}^j} + \frac{(\tilde{T}_e^j - \tilde{T}_i^j)}{R_{i,e}^j} + \alpha^j \cdot \dot{S}^j \quad (1)$$

$$C_e^j \times \frac{d\tilde{T}_e^j}{dt} = \frac{(\tilde{T}_i^j - \tilde{T}_e^j)}{R_{i,e}^j} + \frac{(\dot{T}_a - \tilde{T}_e^j)}{R_{a,e}^j} \quad (2)$$

Table 2
Input data variables and heat transfer coefficients (where $j \in \{1,2,3,4,5,6\}$) represents apartment number

Variables	Description
$\tilde{T}_i^j$	Indoor air temperature (C)
$\tilde{T}_e^j$	Envelop temperature (C)
$\tilde{T}_h^j$	Floor pipes temperature (C)
$\overline{T}_{in}^j$	Inflow hot water temperature (C)
$\dot{T}_a$	Ambient temperature (C)
$\dot{S}^j$	Heating power of solar radiation (kW)
$\overline{M^j}$	Hot water mass flow (kg/s)
Constants	**Description**
C_i^j	Indoor air heat capacity (kWh/C)
C_e^j	Envelop heat capacity (kWh/C)
C_h^j	Floor pipes heat capacity (kWh/C)
$R_{i,h}^j$	Indoor-air-to-floor heat resistance (C/kW)
$R_{i,e}^j$	Indoor-air-to-envelope heat resistance (C/kW)
$R_{a,e}^j$	Ambient-to-envelop heat resistance (C/kW)
α^j	Solar-radiation-to-indoor-air heat exchange
β^j	Hot-water-to-floor pipes heat exchange

$$C_h^j \times \frac{d\tilde{T}_h^j}{dt} = \frac{(\tilde{T}_i^j - \tilde{T}_h^j)}{R_{i,h}^j} + \beta^j \cdot \overline{M}^j (\overline{T}_{in}^j - \tilde{T}_h^j) \quad (3)$$

Equations (1)-(3) provide the conventional mathematical representation of thermal dynamics for a single room [5]. Considering an apartment as a single-zone temperature system, we extend the use of these equations to characterize the thermodynamics of the entire apartment. Table 2 provides a description of data variables and constant coefficients used in these equations. We employ specific notations throughout these equations to differentiate between various thermal variables and constants:

- State variables are adorned with a tilde (e.g., $\tilde{T}_e^j$).
- Input variables directly influenced by the different control levels (i.e., heat pump, mixing loop, and binary valves) are denoted with an overline (e.g., $\overline{T}_{in}^j$).
- Weather effects are marked with a dot (e.g., $\dot{T}_a$).
- All constants remain undecorated (e.g., C_i^j, $R_{i,h}^j$).

As described in Section 2, we tackle our heating problem at the apartment level rather than on a room-by-room basis to simplify and minimize unnecessary complexity. Hence, we discuss data variables and constant coefficients in Equations (1)-(3) at the apartment level.

In Equations (1)-(3), $\tilde{T}_i^j$, $\tilde{T}_e^j$, and $\tilde{T}_h^j$ represent the evolution of indoor air temperature, envelope temperature, and floor (heater) temperature for apartment j, respectively. In

Equation (1), the floor temperature ($\tilde{T}_h^j$), envelope temperature ($\tilde{T}_e^j$), and solar radiation ($\dot{S}^j$) directly influence the indoor temperature ($\tilde{T}_i^j$). Equation (2) illustrates that the envelope temperature ($\tilde{T}_e^j$) is affected by both indoor temperature ($\tilde{T}_i^j$) and outdoor temperatures ($\dot{T}_a$). In Equation (3), the heater temperature ($\tilde{T}_h^j$) is determined by the indoor temperature ($\tilde{T}_i^j$) and the hot water entering the heater. Notably, the primary source of heat for the heater is derived from the hot water, which directly depends on both the mass flow (M^j) and temperature of the hot water ($\overline{T}_{in}^j$).

3.4. Model Identification Using CTSMR

Equations (1)-(3) depict the thermal dynamics of an apartment j. The grey-box modelling within CTSMR is run independently for each of the six apartments to determine the heat transfer coefficients. This approach reduces the overall model complexity and accelerates computations. Once the estimation for all apartments is completed, the decomposed models are reassembled to represent the entire system.

Experimental Setup: Prior to commencing the CTSMR simulations, we rearrange Equations (1)-(3) by combining heat capacities with resistances to reduce the number of constants to be estimated, resulting in modified equations (see Equations 4-6). This rearrangement maintains precision while simplifying the estimation process by reducing the number of constant coefficients. In the new equations, constants η and γ are to be predicted (instead of capacitance and resistance) through the model identification process. Our model uses a 3R3C configuration to represent indoor, envelope, and floor temperatures. It incorporates heat exchanges between the hot water and floor heaters (using forward water temperature and mass flow), floor heaters and indoor air, indoor air and the envelope, and the envelope and the outdoors (including solar radiation and ambient temperature) to capture the building's thermal dynamics. We do not consider other dimensions as the model captures all the heating effects reasonably well.

$$\frac{d\tilde{T}_i^j}{dt} = \gamma_h^j(\tilde{T}_h^j - \tilde{T}_i^j) + \gamma_e^j(\tilde{T}_e^j - \tilde{T}_i^j) + \alpha^j \cdot \dot{S}^j \tag{4}$$

$$\frac{d\tilde{T}_e^j}{dt} = \frac{\gamma_e^j}{\eta_e^j}(\tilde{T}_i^j - \tilde{T}_e^j) + \gamma_a^j(\dot{T}_a - \tilde{T}_e^j) \tag{5}$$

$$\frac{d\tilde{T}_h^j}{dt} = \frac{\gamma_h^j}{\eta_h^j}(\tilde{T}_i^j - \tilde{T}_h^j) + \beta^j \cdot \overline{M}^j(\overline{T}_{in}^j - \tilde{T}_h^j) \tag{6}$$

The historical data is available at a resolution of 5 minutes. The inaugural experiment executes a simulation over three weeks starting from March 01, 2022. Throughout the estimation phase, CTSMR employs historical data about solar radiation ($\dot{S}^j$), outdoor temperature ($\dot{T}_a$), hot water mass flow ($\overline{M}^j$), and inflow water temperature ($\overline{T}_{in}^j$). Before the experiment, we calculate a new set point for the overall indoor temperature of each apartment by computing a weighted average of the set points for each room within the apartment. The objective is to ascertain an approximation of the indoor temperature ($\tilde{T}_i^j$) that aligns closely with the new average indoor temperature. The estimation process for an individual apartment requires approximately 10 minutes.

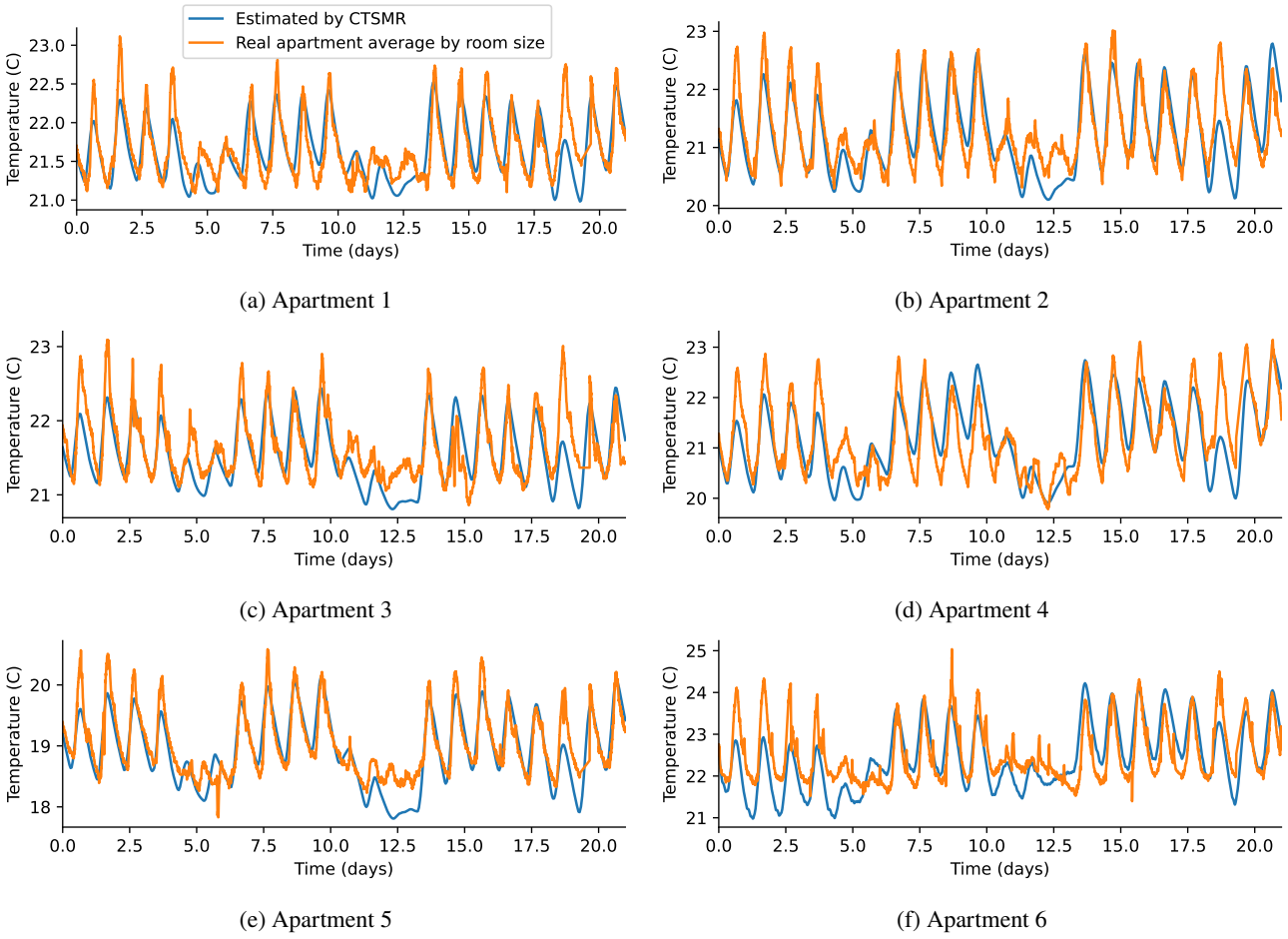
3.5. CTSMR Model Performance Evaluation

We shall now evaluate the estimation quality of our identified model. After the completion of the estimation process, the resultant outputs encompass the acquired heat transfer coefficients delineating the degree of alignment between the estimated apartment indoor temperature trajectory ($\tilde{T}_i^j$) and the real apartment average indoor temperature curve calculated (by room size) from historical data. The estimated coefficients for all six apartments are detailed in Table 3, while the visual representation of these estimations is illustrated in Figure 4. In this experiment, we assess the estimation quality by comparing the predicted indoor temperature with the real apartment average indoor temperature for the same three weeks used in the estimation (from March 01-21). However, we will later evaluate the estimation quality for some other periods. The results depicted in Figure 4 show that the estimated indoor temperatures generally align with the apartments' average indoor temperature curves indicating that CTSMR captures the overall temperature patterns well. However, the estimations exhibit some deviations and fail to accurately capture nearly half of the upper peaks in the average apartment temperature, indicating that CTSMR requires more detailed energy flow information for precise predictions.

We believe that these unaccounted peaks in apartment average indoor temperature are due to some factors not captured well in the estimation process. To investigate this, we first need to determine the time of occurrence and duration of the peaks. We use apartment 1 as an example. Other apartments show similar higher peak patterns; therefore, it is sufficient to analyze one apartment to understand the reasons behind these higher peaks. For the duration, we calculate the 3-hour average window of apartment indoor temperature over the same three-week period. For the occurrence time, we analyze historical data. We plotted the 3-hour window indoor temperatures (for apartment 1) along with its estimated apartment temperature (by CTSMR) in Figure 5a. It can be seen that the duration of these peaks is approximately 3 hours, except a case where it exceeds 3 hours (see highest peak between 17.5 and 20.0 on x-axis).

Historical data analysis reveals that these peaks occur during the daytime, between 15:00 and 18:00. Now, we need to determine the factors that cause these peaks to occur. In Figure 5b, we plot the ambient temperature and solar radiation along with the estimated apartment temperature (by CTSMR) and average apartment indoor temperature for apartment 1. The peaks, where estimations show deviations of approximately 0.5°C or higher, are highlighted in black dotted areas.

It can be seen that the temperature peaks occur when the outdoor temperature and solar radiation increase. In



(a) Apartment 1

(b) Apartment 2

(c) Apartment 3

(d) Apartment 4

(e) Apartment 5

(f) Apartment 6

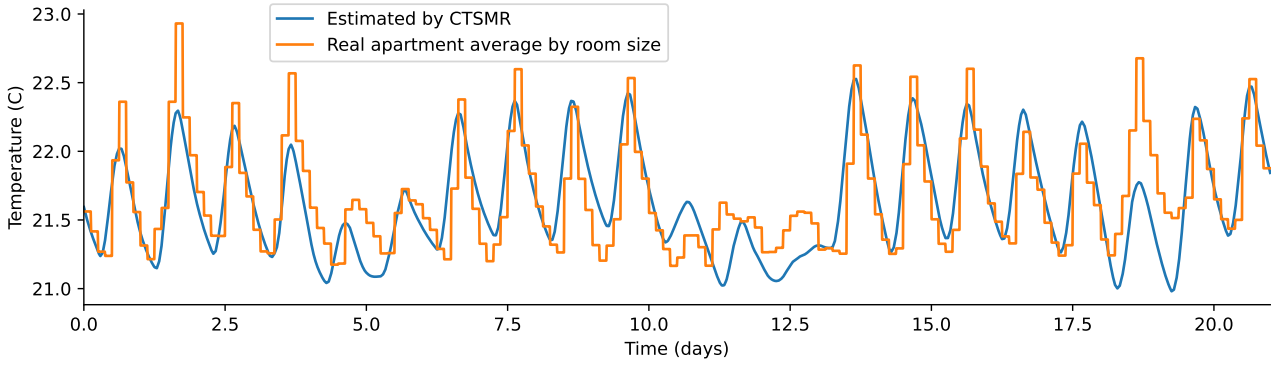
Figure 4: Estimated and average apartment indoor temperatures for the same training data and duration

Table 3
Estimated constant thermal coefficients for the six-apartment model

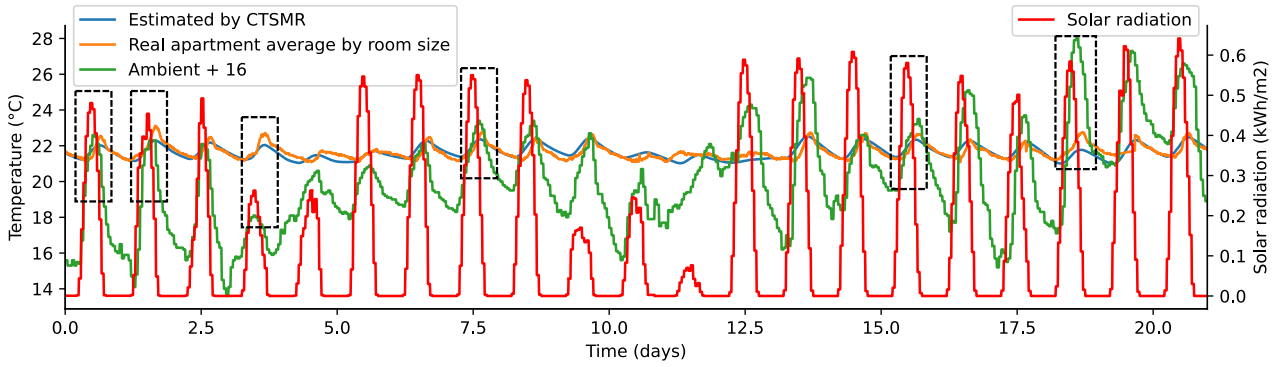
Coefficient	Apartment 1	Apartment 2	Apartment 3	Apartment 4	Apartment 5	Apartment 6
γ_h^j	5.4984e-01	3.9208e-01	4.2124e-01	3.0797e-01	2.3934e-01	1.3960e-01
γ_e^j	7.7262e-02	1.1844e-01	5.2489e-02	1.6988e-01	7.1642e-02	4.6054e-01
γ_a^j	4.6656e+00	4.0069e+00	3.6404e+00	3.1992e+00	6.9912e+00	7.7827e+00
η_e^j	2.7991e+01	3.3544e+01	1.9733e+01	2.0374e+02	2.7328e+01	1.5582e-02
η_h^j	4.5649e+01	9.7201e+01	6.6977e+01	9.5039e+01	3.8544e+01	4.1357e-01
α^j	2.1851e+00	3.6520e+00	2.5039e+00	3.1148e+00	2.6493e+00	2.7877e+00
β^j	1.4624e-02	4.1684e-03	2.9271e-07	6.0119e-01	7.0586e-22	1.2128e+01

our estimations, we use abstraction and perform apartment-wise estimations instead of room-wise estimations, which prevents the estimation model from accurately capturing the detailed information about the outdoor temperature and solar radiation concerning the walls and windows shared between individual rooms and the outdoors. This indicates that predictions can be improved by using room-wise and more detailed behaviour. However, this approach would significantly increase the system's complexity by adding numerous new variables, which could make it difficult for the controller to develop effective strategies. While room-wise estimations are expected to improve estimation quality to some extent, this improvement is unlikely to outweigh the additional computational effort required for the controller to learn optimal strategies, in terms of both CPU time and memory. Therefore, in this work, we will use apartment-wide estimations rather than room-wise estimations.

Beyond visual concordance, we further evaluate the estimation quality utilizing six conventional error metrics,



(a) Apartment 1 estimated and average (with 3-hour window) indoor temperatures



(b) Apartment 1 estimated and average indoor temperatures, along with ambient temperature and solar radiation

Figure 5: Analyzing the peaks in apartment 1 average indoor temperature

Table 4
Conventional error metrics for thermal dynamics estimation for the six-apartment model

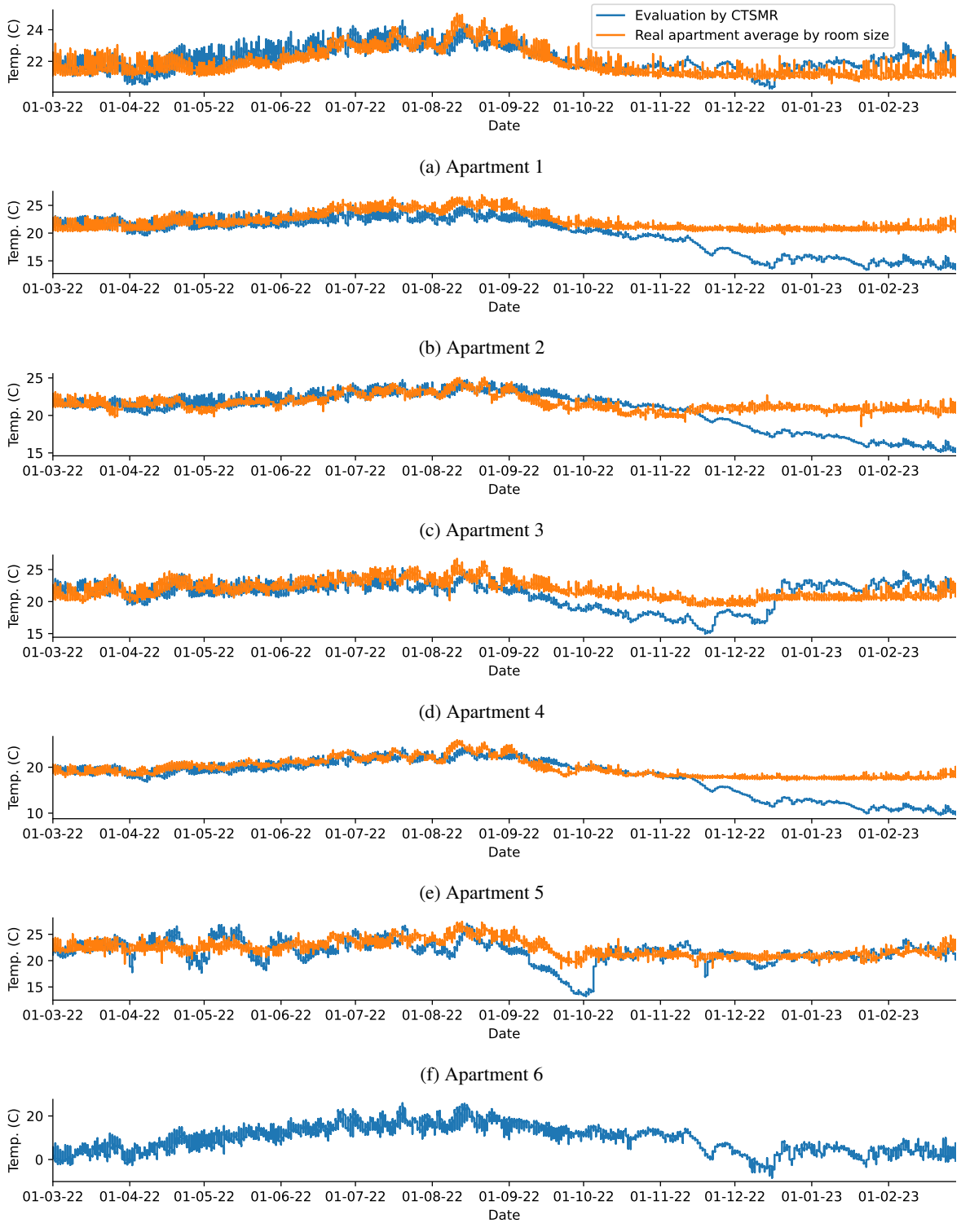
Indoor Temp.	MAE	MAPE (%)	MBE	NMBE (%)	RMSE	CVRMSE (%)
Apartment 1	0.2086	0.9607	0.0198	0.0914	0.2698	1.2460
Apartment 2	0.2929	1.3717	0.0618	0.2901	0.3725	1.7494
Apartment 3	0.2479	1.1411	0.0857	0.3958	0.3225	1.4903
Apartment 4	0.3982	1.8686	0.0481	0.2265	0.4893	2.3036
Apartment 5	0.2262	1.1900	0.0926	0.4866	0.2963	1.5568
Apartment 6	0.4769	2.1133	0.0884	0.3924	0.5591	2.4824
Average	0.3084	1.4409	0.0661	0.3138	0.3849	1.8048

namely Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Mean Bias Error (MBE), Normalized Mean Bias Error (NMBE), Root Mean-Square Error (RMSE), and Coefficient of Variation of The Root Mean-Square Error (CVRMSE) [58, 59, 60].

MAE, MBE, and RMSE metrics assess average errors in absolute values, providing valuable insights into the accuracy of predictions. In contrast, MAPE, NMBE, and CVRMSE offer a perspective on mean errors in percentages. Achieving lower values (close to zero) across these error metrics signifies higher accuracy and better performance for predicting thermal dynamics [58, 61].

Table 4 shows that MAE error is 0.3°C. This metric explicitly represents the average absolute difference between the predicted and actual values, offering a tangible measure of precision in the model's predictions. Similarly, MBE and RMSE error values, standing at averages 0.06°C and 0.38°C for all apartments, respectively, further highlight the consistent and acceptable performance of CTSMR in predicting thermal behaviours.

Delving into the percentage errors, MAPE, NMBE, and CVRMSE unveil values of 1.44 %, 0.3 %, and 1.8 %, respectively. These low percentage errors indicate the model's ability to identify the thermal model with reasonable accuracy,



(a) Apartment 1

(b) Apartment 2

(c) Apartment 3

(d) Apartment 4

(e) Apartment 5

(f) Apartment 6

(g) Outdoor Weather

Figure 6: Temporal validity of estimated coefficients in the Cfb climate zone (Koppen-Geiger classification)

showcasing good predictions of heat transfer coefficients. The small values across these conventional error metrics strengthens the confidence in capabilities of CTSMR, affirming its proficiency in capturing the dynamics of thermal systems. Recall that the building has five entrances, and we are specifically focusing on the rightmost staircase. Apartments 2, 4, and 6 have no neighbours on the right hand side, while apartments 1, 3, and 5 are semi-open (to the outside)

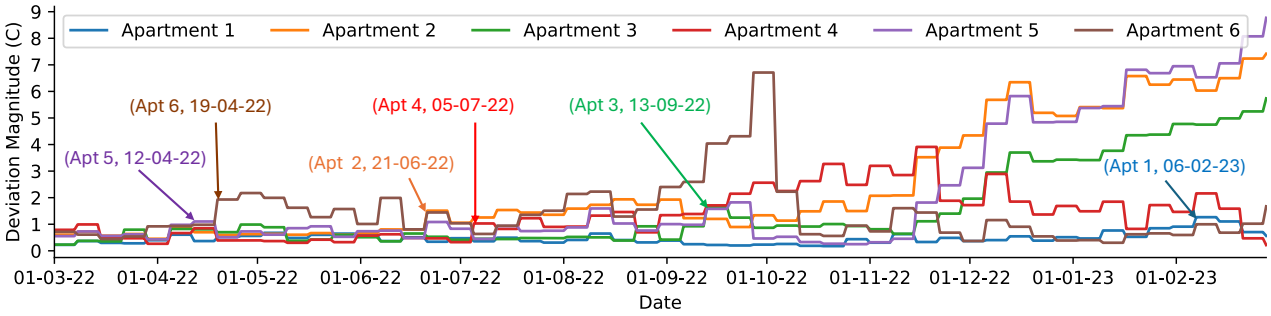


Figure 7: Deviation magnitudes during the temporal validity of estimated coefficients, where 'Apt' refers to an apartment

from the left due to neighbouring apartments from another staircase. Therefore, apartments 2, 4, and 6 face more wind effects than apartments 1, 3, and 5 due to their placement location in the building. Apartment 6 is the most exposed to wind due to its additional rooftop area open to outdoor. It is important to note that our models do not account for the effects of wind, which is why we observe larger but still acceptable discrepancies in errors for apartments 2, 4, and 6.

Apartment 6 exhibits the most significant errors but still falls within an acceptable range. However, in terms of NMBE, apartment 6 does not have the highest error. The NMBE considers the scale of observed values, suggesting that although the average absolute difference in predictions is more significant for apartment 6, this difference is proportionally smaller when considering the scale of its observed values.

Furthermore, we experiment by incorporating household electricity use information into the thermal dynamics model of the building. However, the observed outcomes did not demonstrate substantial improvement. One possible reason can be that the chosen electricity use data may lack a strong correlation with the thermal dynamics of the building. In other words, electrical energy consumption within households might not be directly influenced or significantly impacted by the thermal behaviour of the apartments. This lack of correlation can result in minimal/zero enhancements to the accuracy and predictive capabilities of the identified model. The other possible reason can be that the lack of improvement indicates that the current model effectively captures the thermal behaviour of the building, rendering the inclusion of extra variables such as household electricity use unnecessary.

3.6. Temporal Validity of Identified Model

We assess the temporal validity of the identified model by extending our evaluation beyond the prediction horizon. Utilizing the estimated coefficients, we extend our analysis over a longer period, surpassing the initial prediction window, to gauge their effectiveness in capturing the building's thermal dynamics. It is essential to note that our simulation initially spanned three weeks, and we now validate the model's performance over an entire year.

In Figure 6, we illustrate the temporal validity of the estimated heat transfer coefficients using historical data from March 2022 to February 2023. Notably, the coefficients demonstrate satisfactory performance for apartment 1. However, significant discrepancies emerge for apartments 2, 3, 4, 5, and 6 during the year's last five to six months. This observation suggests that the building's thermal dynamics change over time, rendering the initially estimated coefficients less reliable for an entire year. The thermodynamics of a building is a dynamic and evolving aspect that can change over time. Several factors, such as aging, climate/seasonal variations, and changes in usage, contribute to this temporal variation, influencing the building's thermal behaviour and energy dynamics.

The aging factor is concerned with building materials and systems. As a building ages, its components' thermal properties may change. Insulation materials may degrade, and the overall thermal performance may be affected. As the building has recently been renovated, aging is only a significant factor if underlying faults or leakages exist in the building envelope. The integrity of the envelope plays a crucial role in maintaining thermal performance. The discrepancies are likely not due to the aging of the building elements, as these typically do not degrade rapidly or uniformly across different apartments.

Secondly, the inherent variability of seasonal patterns and climate changes significantly affects the thermodynamics of buildings. Seasonality in certain identified parameters and factors not included in the model can impact performance in seasons that differ significantly from the training data. Using three weeks in March for training is insufficient for the entire year, not because of the aging of the building elements, but due to the limitations of the model structure itself—namely, its size, linearity, and time-invariance. These changes necessitate a periodic reassessment of the model parameters to account for the dynamic nature of external conditions.

Moreover, considering the building in question is a renovated low-energy structure, occupant behaviour emerges as a potent influencing factor. Activities such as window opening, adjustments to heating setpoints, and the presence of other internal loads cause a more pronounced impact on the overall thermal dynamics.

Lastly, it is crucial to acknowledge that, over time, any mathematical model is expected to gradually deviate from the real-world scenario due to dynamic factors that are challenging to capture comprehensively. Therefore, continuous refinement and validation of the model may be necessary to enhance its predictive accuracy over extended time frames.

In summary, the thermodynamics of a building is subject to change over time. Continuous monitoring and adaptation to evolving conditions are essential for optimizing buildings' thermal performance and energy efficiency throughout their life-cycle. Therefore, it becomes imperative to re-evaluate the constant coefficients periodically to enhance the precision of capturing the building thermal dynamics. For instance, a more practical approach might involve re-estimating these coefficients at regular intervals. This adaptive strategy acknowledges the evolving nature of the building thermal behaviour, ensuring that our model remains accurate and reliable over extended time frames.

Figure 7 illustrates the deviations of estimated temperatures from the average apartment temperatures over one year, measured using Mean Absolute Error (MAE) metrics on a 7-day window to demonstrate the magnitude of error over time. Recall that we calculate the average temperature for each apartment by taking the average of its room temperatures, and the estimations are based on these average values. Therefore, deviations of up to 1°C are considered acceptable. The arrows in different colors indicate the date at which each apartment first time exhibits a 7-day window with a deviation of more than 1°C in the plots. Generally, the error magnitude remains below 1°C for all apartments for more than three months, except for apartment 6, which experiences higher deviations as it is more exposed to wind and weather conditions than any other apartment due to its location in the building. Additionally, apartment 5 also shows a deviation of above 1°C for one of the weeks within these three months. As all other apartments begin to show deviations at various times after apartment 5, we consider apartment 5 as a baseline and suggest re-estimating the coefficients every six weeks. Given that the estimation process is lightweight, it can be performed more frequently, such as every other week. This more frequent updating would help monitor the system more effectively and allow for quicker detection of any problems.

3.7. Modelling in UPPAAL STRATEGO

We design both the evaluation and learning model in UPPAAL STRATEGO using the identified thermodynamics model and estimated heat transfer coefficients. The evaluation model represents the building's behaviour, while the learning model functions as the MPC controller.

Figure 8 illustrates the UPPAAL STRATEGO model and its overall structure. We represent different functionalities or sub-parts of the system as parameterized timed automata, referred to as templates. These templates are shown within the dashed areas, labelled as *Apartment*, *Optimization*, *DataReader*, *BufferTank*, *BufferLayer*, *BufferUpdate*,

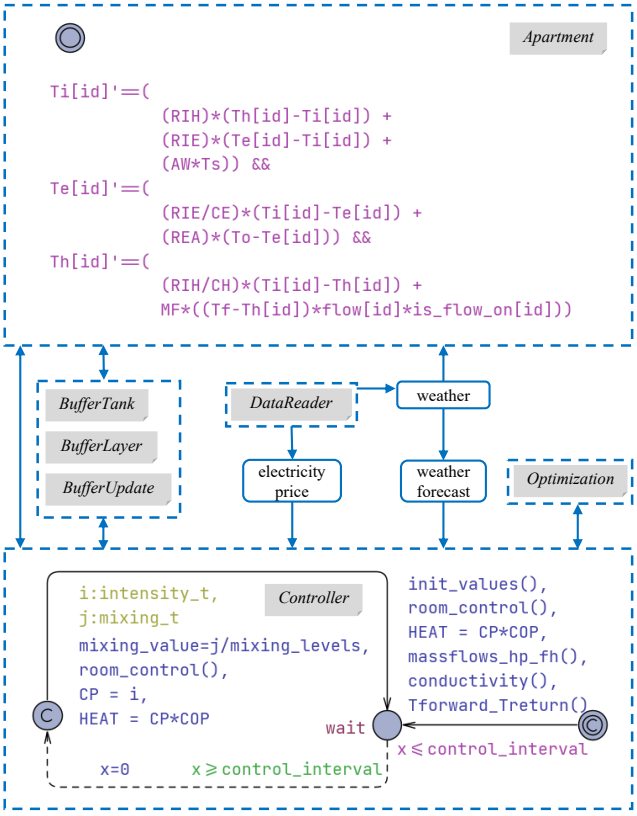


Figure 8: Overview of UPPAAL STRATEGO model

and *Controller*. In the templates, double circles represent initial locations, while single circles represent regular locations. The *Apartment* template models the behaviour of the three-state energy model represented by differential equations (4-6) and expresses the continuous variables $\tilde{T}_i^j$, $\tilde{T}_e^j$, and $\tilde{T}_h^j$ as real-time clocks. The *Optimization* template implements the fitness function (see Equation 8). The *DataReader* template supplies the model with weather and electricity data. In our modelling, we divide the water in the hot water buffer tank into several virtual layers to precisely calculate the water temperature within the tank. Templates related to buffer tank functionality are grouped together: the *BufferTank* template calculates the water temperature in the top and bottom layers, while the *BufferLayer* template handles the temperature of the intermediate layers. The *BufferUpdate* template updates the water flow direction and the heat conductivity values in the tank. The complete model, along with the details of all templates, is publicly available at [62].

In the *Controller* template, some locations are referred to as committed locations, labelled with the letter "c". The system does not allow delays in committed locations. The arrows between locations represent actions or transitions. Solid-line transitions indicate controllable actions, while dashed-line transitions represent uncontrollable actions (which are controlled by the environment). Several functions are designed to perform specific tasks. The function `init_values()` takes the apartment temperatures,

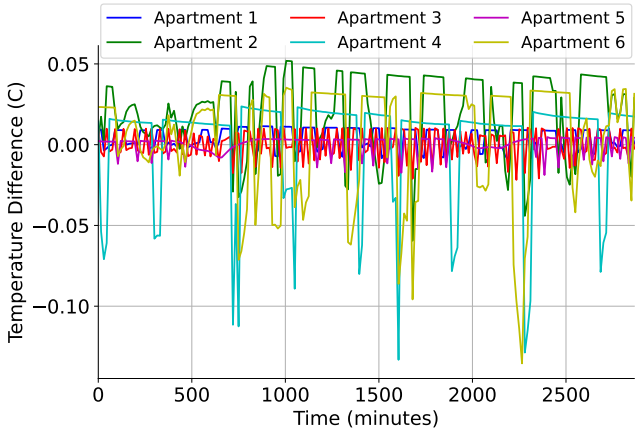
power consumption (CP), and heat pump coefficient of performance (COP) from the previous control interval (of the evaluation model), while also loading relevant data such as outdoor temperature, solar radiation, and electricity prices directly from data files into the UPPAAL STRATEGO controller. The `room_control()` function operates the room thermostats based on the indoor temperature, using a traditional bang-bang (BB) controller. The BB controller simply turns the thermostat *on* when the indoor temperature falls below the setpoint, and *off* when it rises above the setpoint. The `massflows_hp_fh()` function calculates the mass flow of hot water entering the buffer tank from the heat pump and the mass flow of cold water returning to the buffer tank from the floors. The `conductivity()` function computes the heat conductivity within the tank depending on the direction of the water flow, and the variable `HEAT` represents the amount of heat produced. The `Tforward_Treturn()` function determines the temperature of the water after mixing (forwarded water) and the temperature of the cold water returning from the floors. At the `wait` location, the invariant `x<=control_interval` allows the system to stay in this location for upto 15 minutes. At the following transition, the guard `x>=control_interval` ensures that the system can only take this transition when the clock `x` has reached a value of at least 15. This means the system must remain at the `wait` location for exactly 15 minutes. Afterwards, the clock `x` is reset with `x=0`. In the transition after the next committed location, the controller can choose any energy intensity level (out of ten) for operating the heat pump and any of the three mixing loop values for the next control interval. The select statements, `i:intensity_t` and `j:mixing_t`, assign values to the temporary variables `i` and `j` selected from the predefined sets `intensity_t` $= \{1, \ldots, 10\}$ and `mixing_t` $= \{1, \ldots, 3\}$. The learning model uses Q-learning [63] and engages in repeated sampling to acquire a near-optimal strategy for controlling the heat pump and mixing loop. Through repeated sampling, the controller effectively adapts to temporal changes, including fluctuations in energy prices and upcoming weather forecasts. The controller continues this repeated sampling for next 12 hours to synthesize a strategy.

The controller uses the formula presented in Equation 7 to synthesize a strategy. It processes a feature vector comprising the directly measurable variables of the original system and aims to minimize the fitness function F (as shown in Equation 8).

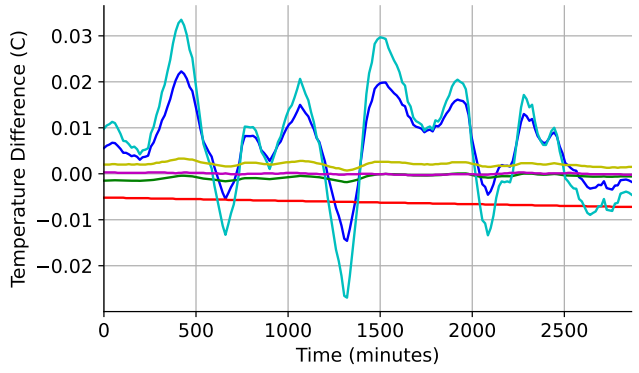
(7)

```
strategy S = minE(F) [t<=12*60)] {} ->
       {Ti[0], Ti[1], Ti[2], Ti[3], Ti[4],
               Ti[5], t, To, Tf}: <> t==(12*60)
```

We set a 15-minute control interval suitable for managing slow dynamics of floor heating systems. Every 15 minutes, the evaluation model adopts a new decision predicted by the learning model through an external library. A 15-minute control interval provides sufficient time for the controller to learn and compute strategies for future control intervals. A 15-minute control interval is feasible and provides more fine grained control of the system, as it allows for observing the system state information every 15 minutes. This enables more precise adjustments to the heat pump intensity to better meet heating demands. Given that we have only three mixing levels for the mixing loop, a 15-minute interval is adequate for re-estimating the mixing. A longer control interval would necessitate additional mixing levels, thereby increasing the complexity of the system.



(a) Deviation of $\tilde{T}_h^j$ estimated values (by STRATEGO controller) compared to original values from evaluation model



(b) Deviation of $\tilde{T}_e^j$ estimated values (by STRATEGO controller) compared to original values from evaluation model

Figure 9: Overview of the quality of $\tilde{T}_h^j$ and $\tilde{T}_e^j$ estimations by the STRATEGO controller

Our experiments show that a 12-hour horizon is sufficient for effective learning, meaning the controller looks only 12 hours ahead in the future while synthesizing a strategy. This approach effectively utilizes the 24-hour limited future energy price availability and the relatively higher accuracy of near-term weather forecasts.

Our heating system effectively addresses the challenge of unobservable variables, such as floor and envelope temperatures. As expressed in Equation (1), it is evident that changes in indoor temperature ($\tilde{T}_i^j$) directly depend on the variations in envelope temperature ($\tilde{T}_e^j$) and floor temperature ($\tilde{T}_h^j$). Therefore, good estimation of these unobservable variables

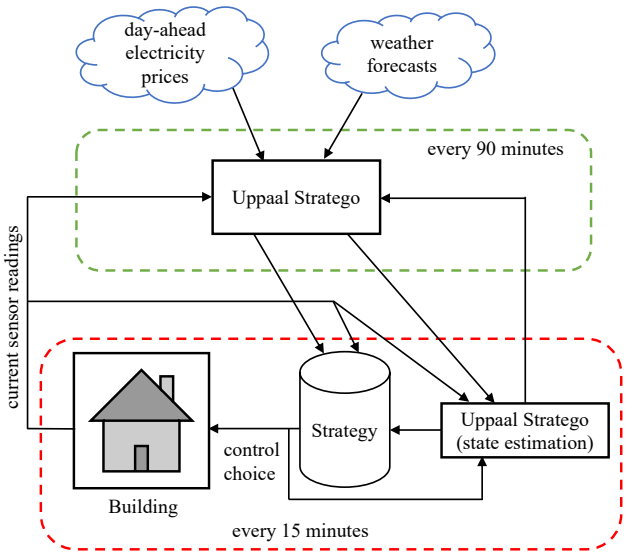


Figure 10: Overview of online strategy synthesis approach

is crucial for the effective operation of the system. At any control interval T, the variables $\tilde{T}_e^j$ and $\tilde{T}_h^j$ are predicted for the future time interval $T+1$ through forward simulation using the current state of the evaluation model and the implemented control decision.

As previously discussed, the floor temperature ($\tilde{T}_h^j$) and the envelope temperature ($\tilde{T}_e^j$) are unobservable variables. This implies that the controller lacks direct information regarding the states of these variables throughout the simulation period. Therefore, the controller must independently estimate these values. To assess the accuracy of the controller's estimations, we recorded the values of these variables from both the evaluation and the learning models. The learning model is a simplified version of the ground truth used by the STRATEGO controller for reinforcement learning. Figure 9 illustrates how the estimated $\tilde{T}_h^j$ and $\tilde{T}_e^j$ values in the learning model deviate the values in the evaluation model over a 48-hour simulation period. Experimental findings demonstrate that the controller precisely estimates these hidden variables, allowing only a marginal drift of less than 0.15°C between the values in the evaluation and learning models.

To address the aforementioned challenges, we also employ an online strategy synthesis approach, wherein the controller periodically learns a strategy to iteratively learn near-optimal control decisions. This approach enables the controller to oversee the system and rectify errors inherent in the initial state estimation of the evaluation model. In contrast, offline synthesis approaches generate only a single strategy, followed by the system throughout the simulation. Figure 10 provides an overview of our suggested online approach.

The UPPAAL STRATEGO controller, highlighted within the green-dashed area, takes inputs (every 90 minutes) such as electricity prices, weather forecasts, and current sensor readings from the building, including room temperatures, buffer tank water temperature, selection of power intensity, and the current mixing level of the mixing loop. Utilizing this information, the controller initiates its learning process from a relevant starting point and looks 12 hours into the future. We also tested synthesizing strategies at intervals of 15 minutes instead of 90 minutes. The results show that while a 15-minute strategy offers only marginal performance improvements, the additional computational effort required is not justified. Consequently, we have opted for a 90-minute strategy, which provides reasonable performance while requiring six times less computational effort. After a strategy is synthesized, we save it in an external file.

Note that the control frequency remains 15 minutes, regardless of whether the strategy is synthesized in 15 or 90 minutes. The activities within the red-dashed area occur at 15-minute control intervals. Over the following 90 minutes (spanning 6 control intervals), the evaluation model transmits current sensor readings to the saved strategy file and retrieves control decisions in return. Simultaneously, UPPAAL STRATEGO predicts unobservable state variables ($\tilde{T}_h^j$ and $\tilde{T}_e^j$) every 15 minutes. This state estimation process utilizes current sensor readings, the current control choice, electricity prices, and weather forecasts, executing a forward simulation in UPPAAL STRATEGO. As a result, when the controller synthesizes the next strategy after 90 minutes, it integrates the most recent estimated values of $\tilde{T}_h^j$ and $\tilde{T}_e^j$.

4. Simulation Experiments

This section details our experimental setup, providing an overview of the associated conditions and assumptions. Following this, we conduct a series of experiments to evaluate the performance of our proposed controller approach.

4.1. Experimental Setup

To assess the effectiveness of our approach, we simulate the heating control of the evaluation model under weather conditions matching the second week of January 2023, incorporating real-time energy prices from the Danish day-ahead electricity market. In reality, the house and the controller rely on different sources of weather information, with the house experiencing actual weather conditions and the controller depending on weather forecasts. To mimic this difference, we use the actual weather information in the evaluation model and introduce some noise factor to generate weather forecasts for the controller. With the introduced noise, the weather data for the controller may deviate by up to 1°C, either positively or negatively, compared to the evaluation model. We also conducted experiments with noise levels of 2°C, 3°C, and 4°C deviations. The results demonstrated that the controller's performance remained consistent, with only a marginal drop in energy cost savings ranging from 0.5 % to 1 %. However, all the experiments presented in this paper include a 1°C noise deviation. Additionally, as previously mentioned (in Section 3.7), the controller operates without receiving information about unobservable variables (floor and envelope temperatures) from the evaluation model. Instead, it autonomously estimates these variables,

emphasizing that the controller operates under conditions of partial observability.

4.2. Control Strategy Approaches

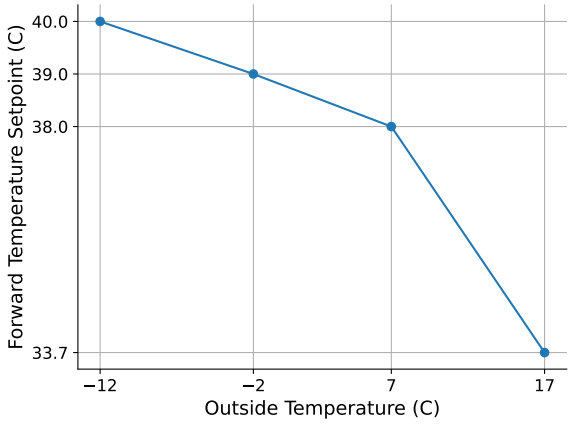


Figure 11: Weather compensated control (WCS)

In the following, we describe the already existing control strategy in the building and our proposed UPPAAL STRATEGO controller:

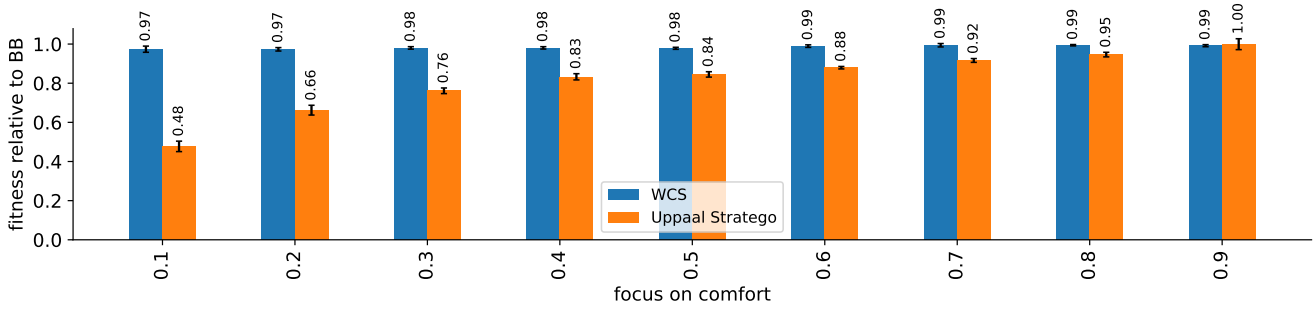
- ***Existing Control Strategy:*** Weather Compensated Control (WCS) is the strategy already implemented in our case building. The WCS adjusts the forward water temperature in response to variations in outdoor temperature. Figure 11 provides a visual representation of the WCS curve illustrating the relationship between the forward water temperature and outdoor temperature. It can be seen that WCS employs specific setpoints for the forward water temperature corresponding to particular outdoor temperatures. For instance, when the outdoor temperature is 7°C, the strategy maintains the forward temperature at 38°C. If the outdoor temperature exceeds 17°C, it adheres to a forward temperature of 33.7°C, while for temperatures below -12°C, it maintains a forward temperature of 40°C. As the outdoor temperature drops, the forward water temperature needs to increase to maintain indoor comfort, and when the outdoor temperature rises, the forward water temperature can be reduced to avoid overheating or unnecessary energy use. The WCS controller helps to automatically adjust the water temperature to achieve the setpoints, reacting to any differences between the desired and actual water temperature. The WCS calculates the error in the current difference between the desired forward water temperature and the actual temperature. If the forward water temperature is too low, the controller increases the heat supply, and vice versa. PID is only used for setting the operating frequency of the heat pump to keep the forward temperature close to setpoints.

  We can use fixed 10 intensity levels in the control choices of the WCS. This would entail implementing a PID-like controller to correctly adjust the intensity levels. We assume that this change to the WCS controller will have little to no impact on the results, the average runtime and periods of operation will converge, resulting in the same overall cost and consumption profiles (statistically). Note that modern frequency-regulated heat pumps have efficiency gains when operating at less-than-full capacity. However, we have not included these effects in the modelling of this study, and leave the modelling and estimation of these effects to further work.

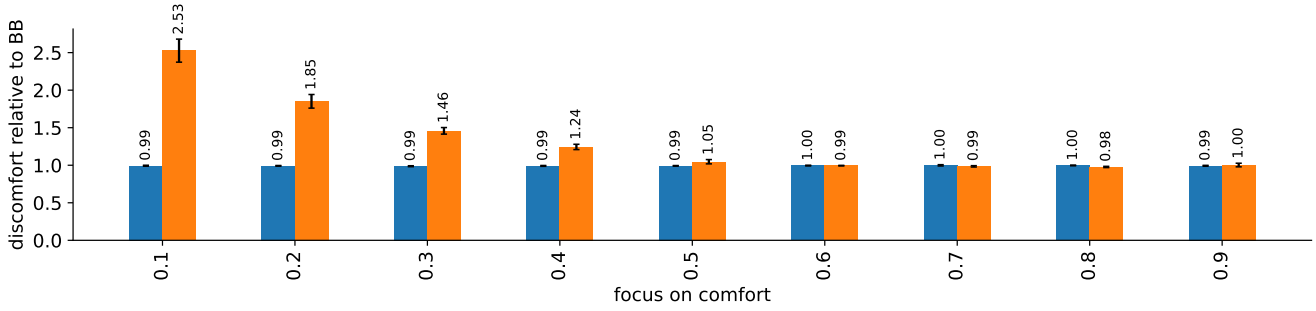
- ***Our Proposed Controller:*** This paper proposes an intelligent UPPAAL STRATEGO controller incorporating future energy prices and weather forecasts to optimize the heating system efficiently. The controller operates the heat pump across 10 intensity levels, evenly distributed between zero and the maximum power of 10 kW. This controller also controls the mixing loop valve, determining which proportion of cold water returning from the apartments to blend with the hot water from the buffer tank. This adjustment influences the forward water temperature, enhancing user comfort. To simplify the learning process and reduce complexity, the controller selects from three options for the mixing loop: circulating only hot water, circulating only cold water, or mixing them equally. We observed from the data that the forward water temperature typically ranges between 33°C and 40°C, while the returning water temperature varies from 26°C to 33°C. If we consider a forward water temperature of 40°C and a returning water temperature of 30°C, then by mixing them equally, the new forward water temperature will be 35°C. The controller's goal is to minimize energy use costs while ensuring user comfort remains uncompromised.

  The local control of the residential heat pump regulates the forward temperature according to WCS (i.e. the forward temperature is a function of the outdoor temperature), not PID. PID is only used for setting the operating frequency of the heat pump to keep the forward temperature (this is only for the recently introduced propane pumps). We are not aware of any residential heat pump model currently on the market using more advanced mechanisms of control. However, our proposed controller uses reinforcement learning to predict the future control choices for the operating frequency (from 10 predefined levels) of the heat pump instead of a traditional PID.

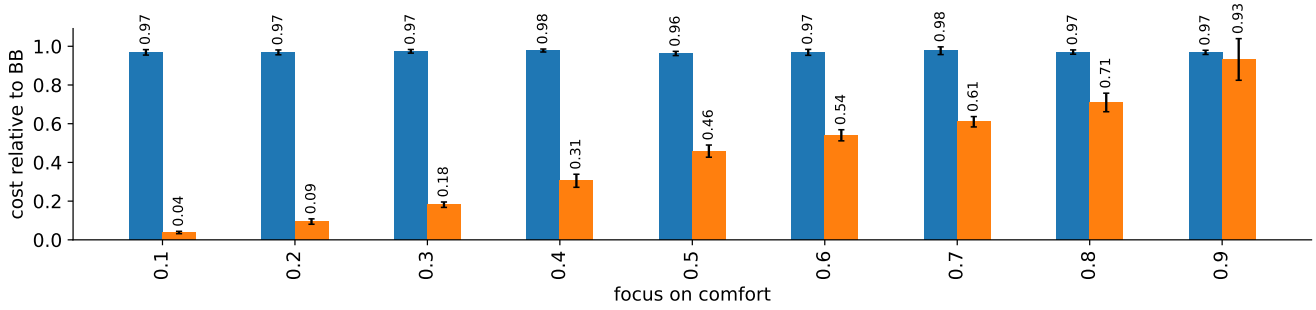
  Regarding the heat pump sizing, for this paper, we consider an existing installation (10 kW), i.e. we assume that the pump has been dimensioned for the building in accordance with industry practice. It is an interesting question whether intelligent control can change the dimensioning of the pump system; we



(a) Fitness function values measured against different settings of W_1



(b) User comfort measured against different settings of W_1



(c) energy use cost measured against different settings of W_1

Figure 12: Experimental results for a week under the control of the WCS and UPPAAL STRATEGO controller.

suspect not, as the heating system must be sized to meet demand in worst-case scenarios.

We assess the performance of both the WCS and UPPAAL STRATEGO controllers by comparing them with a traditional bang-bang (BB) controller. The BB controller operates deterministically, governing the heat pump based on forward water temperatures within the same range as WCS (i.e., 33.7 to 40°C). It activates the heat pump at full intensity when the forward water temperature falls below 33.7°C and turns it completely off when the forward temperature exceeds 40°C.

4.3. Experimental Results

Typically, consumers aim to strike a balance between comfort and energy cost, often opting for settings where they are willing to tolerate a certain level of discomfort to achieve significant savings in energy use costs. To address this trade off paradigm and optimize the controller mechanism, we introduce an objective/fitness function (F) similar to what was proposed in [32]. The fitness function (F) in Equation 8 effectively balances the trade off between energy cost and user comfort. We achieve this balance by adjusting the weights W_1 and W_2 in increments of 0.1, where $0 \leq W_1 \leq 1$ and $W_2 = 1 - W_1$.

$$F(\ell_0, \ell_n) = \int_{\ell_0}^{\ell_n} \left(W_1 \cdot NF \cdot cost(\ell) \right) + W_2 \cdot \sqrt{\sum_{j=1}^{k} (T_g^j - \tilde{T}_i^j(\ell))^2 \, d\ell} \tag{8}$$

Furthermore, to ensure proportional weighting between comfort (measured in squared degrees Celsius) and energy cost (DKK), we compute a normalization factor as $NF = \frac{comfort}{cost}$. The formula obtains the accumulated comfort and cost values from controlling the heat pump with BB controller a week before the experimental week. The primary objective of the controller is to reduce energy use costs

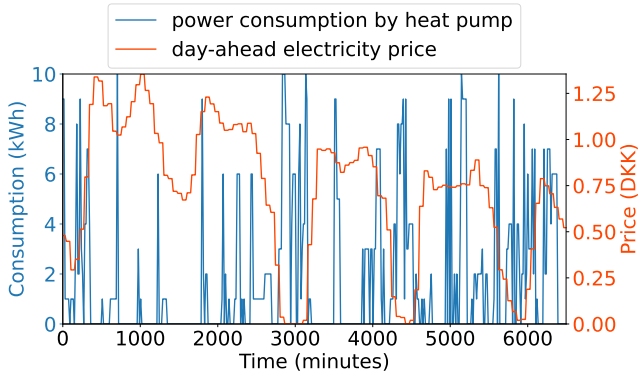
and minimize the gap between the indoor temperatures of apartments and their set points.

For each apartment, the set point is computed by taking the weighted average of its room set points. Consequently, the set points for apartments 1, 2, 3, 4, 5, and 6 are 20.7°C, 20.5°C, 20.2°C, 20.4°C, 17.11°C, and 20.0°C, respectively. The function is applied over a specific period ℓ_1 to ℓ_n, provided that the energy prices, heat pump settings and apartment temperatures $\tilde{T}_i^j$ (denoted by k apartments) are known.

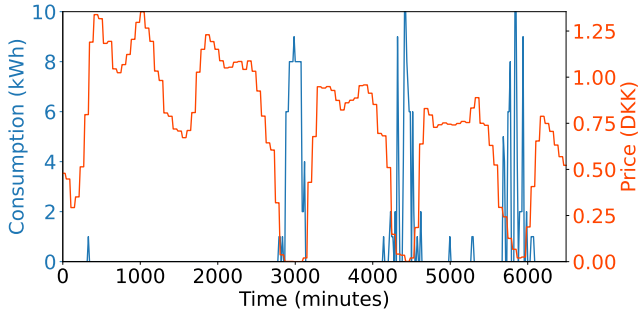
In Equation 8, the variable $cost(\ell)$ is a time-dependent function encompassing the energy use cost derived from the day-ahead energy price and the heat pump's energy use. The term $\tilde{T}_g^j$ represents the indoor set point for apartment j. The first part of the function aims to minimize energy use costs, while the second part seeks to minimize the gap between indoor temperatures and set points for each apartment. Squared differences imply a more significant penalty for larger temperature deviations.

In Figure 12, we take BB controller as our reference point and compare the WCS and UPPAAL STRATEGO controller performances. The x-axis portrays different settings of weight W_1, symbolizing a focus on comfort, with W_1 values of 0 and 1 omitted, as they respectively entail turning the heat pump off completely or entirely ignoring cost considerations. On the y-axis, the *fitness* and *cost* measures represent the fitness function values and energy use costs against different settings of W_1, respectively. However, the *discomfort* measure indirectly mirrors *comfort*, and a decrease in its value signifies an improvement in comfort and vice versa. These measures are derived by computing their values from WCS or UPPAAL STRATEGO with those from BB, expressed as $measure = \frac{controller}{BB}$. A value lower than 1 for a measure indicates the superior performance of the respective controller compared to BB.

Figure 12 illustrates the performance of WCS and UPPAAL STRATEGO in comparison to BB. Initially, focusing on WCS, it exhibits slightly better fitness and similar comfort levels while also achieving 2-4% energy cost savings compared to BB. Shifting the comparison to UPPAAL STRATEGO against WCS, Figure 12 demonstrates that UPPAAL STRATEGO consistently outperforms WCS across all W_1 configurations for fitness, comfort, and cost measures except at $W_1 = 0.9$, where it marginally underperforms WCS in fitness and comfort. Now, with the help of Figures 12b and 12c, let us explore the ability of STRATEGO controller to efficiently manage the balance between user comfort and power consumption cost. For instance, at $W_1 = 0.1$, STRATEGO achieves over 93% savings in energy use costs but introduces more than 150% discomfort, which is unrealistic and unacceptable. The setting $W_1 = 0.6$ is interesting where STRATEGO provides comparable comfort to WCS but saves 43% in energy costs. Users can achieve additional cost savings if they are open to a marginal decrease in comfort. For example, at $W_1 = 0.5$, with a minor increase in discomfort by 5%, users can save an additional 7% in power consumption costs.



(a) With $W_1 = 0.6$



(b) With $W_1 = 0.1$

Figure 13: Price of electricity and power consumed by the heat pump (operated under the STRATEGO controller).

Table 5 provides an overview of power consumption and associated costs with a setting $W_1 = 0.6$ under BB, WCS and UPPAAL STRATEGO controllers. The results presented in the table were generated by controlling the heat pump over one week, specifically matching the weather conditions from January 9-15, 2023. The table displays power consumption for the heat pump (P_{hp}) as well as for individual apartments (P_{apn}, where $n \in \{1, 2, \ldots, 6\}$). It can be observed that the total power consumed by all the apartments (P_{ap_total}) is lower than the power consumed by the heat pump (P_{hp}). We believe this discrepancy is due to energy losses from the hot water buffer tank and transmission lines.

Now we discuss P_{hp} and the related cost under BB, WCS and UPPAAL STRATEGO controllers. It can be seen that BB consumes approximately 5% more energy, leading to a 3% increase in energy costs compared to WCS. This variance in percentage energy use and cost arises from potential differences in the energy prices utilized by the controllers during different control intervals. The power consumed by STRATEGO is 335.5 kWh, which is 31.2% lower than the power consumed by WCS (487.5 kWh). STRATEGO's reduction in power consumption is a cumulative outcome of its ability to leverage future information about weather and electricity price through learning—an additional factor is controlling the heat pump across 10 intensity levels and effectively mixing cold and hot water by the mixing loop. With the reduced power consumption, STRATEGO can save 31.2% in energy use costs compared to WCS.

Table 5
One-week power consumption (in kWh) at apartment and heat pump levels, with corresponding costs (in Danish Krone, DKK), under various control methods with $W_1 = 0.6$ setting. Here, P_{hp} denotes heat pump power, P_{apn} denotes apartment n's power (where $n \in \{1, 2, \ldots, 6\}$), and P_{ap_total} represents total power consumption by all the apartments.

Controller	P_{ap1}	P_{ap2}	P_{ap3}	P_{ap4}	P_{ap5}	P_{ap6}	P_{ap_total}	P_{hp}	Cost (DKK)
BB	92.2	108.4	126.0	33.2	12.9	88.3	461.0	511.8	224.2
WCS	93.4	109.9	121.8	32.4	12.6	85.0	455.1	487.5	217.5
STRATEGO	53.2	99.1	67.5	17.3	6.2	76.5	319.8	335.5	124.0

However, a closer look at power consumption costs reveals that WCS incurs 217.5 DKK, while STRATEGO incurs 124.0 DKK. This depicts that STRATEGO achieves approximately 43% savings in energy costs. STRATEGO'S additional 11.8% cost savings result from its ability to use power when it is more economical. Figure 13a illustrates that STRATEGO consistently utilizes power during its periods of very cheap or zero-cost. It sometimes also adapts to less cheap periods because, with W_1 set to 0.6, there is a 60% emphasis on comfort and a 40% on cost. Therefore, STRATEGO also has to choose some periods of less cheap energy to maintain user comfort by turning the heat pump on.

On the other hand, when we set W_1 to 0.1 (Figure 13b), STRATEGO controller prioritizes cost over comfort (90% vs. 10%). Consequently, it disregards comfort by activating the heat pump only during periods of highly cheap energy. The significantly reduced power consumption cost introduces unacceptable discomfort in the apartments by keeping the heat pump on even during heat demand periods. Therefore, it is essential to balance power consumption cost and user comfort by choosing appropriate values of W_1.

Recently, one of our co-authors launched CEDAR [64], a company offering machine learning-based solutions to optimize heat pump systems. The company has implemented an enhanced version of the system outlined in this paper across multiple real homes. According to CEDAR, the setup can run on a mini-computer similar to a Raspberry Pi 4 (or Raspberry Pi 5 at a comparable price). The initial setup cost is around 2200 DKK (300 EUR), with an additional 1000 DKK (134 EUR) for sensors. The operational consumption of the Raspberry Pi 4 CM is at max 4 watts. Savings are projected to range from 2000-5000 DKK (270-670 EUR) per year for a moderately modern household. The hardware and setup cost will be recovered within 1-2 years.

5. Conclusion

In this study, we presented a comprehensive tool-chain for efficiently controlling the heating system in a floor heating based multi-apartment building. The tool-chain explores methods to address the complexities related to real sensor data integration, scalability, applicability, varied weather effects for each apartment, and diverse heating demands and habits of the occupants sharing the same resources in the apartments.

First of all, we modelled the thermal dynamics of each apartment as Ordinary Differential Equations (ODEs) to characterize the thermal behaviour of the building. We then utilized CTSMR software to estimate the heat exchange coefficients, including resistance and capacitance, among various heat variables. The findings demonstrated strong predictive capabilities of CTSMR, showing no divergent behaviour when evaluated against standard error metrics. Subsequently, we analyzed the temporal validity of the estimated heat exchange coefficients over a year-long horizon. The results highlighted the dynamic nature of building thermodynamics over time, emphasizing the necessity for regular re-estimations of the heat exchange coefficients to effectively capture the evolving thermal behaviour of the building.

Subsequently, we designed an intelligent and online UPPAAL STRATEGO controller to optimize the trade-off between energy cost and user comfort by adjusting heat use. We compared the performance of our proposed controller with the weather-compensated control (WCS) already implemented in the building. The results revealed that the proposed UPPAAL STRATEGO Controller achieved a 31.2% reduction in energy use and 43% reduction in energy costs while maintaining the same level of comfort as the WCS. The higher reduction in energy costs compared to energy use is due to the ability of the controller to utilize electricity when it is cheaper. Moreover, UPPAAL STRATEGO offered an additional 7% savings if users were willing to tolerate a slight increase in discomfort.

Our approach presents what can be expected from simulations for intelligent heating solutions for larger and more complex multi-apartment buildings in real-world scenarios. Additionally, it suggests methods for reducing the complexity of such systems without sacrificing crucial information. In our approach, we simplify the learning process by reducing the number of variables, focusing on the average behaviour of an apartment rather than treating each room as a separate temperature zone. This approach works well even when some rooms in an apartment have relatively lower setpoints. The thermostat in each room employs a traditional bang-bang controller, which prevents overheating. Exploring a completely explicit room-wise control remains an area that requires further investigation.

6. Data Availability Statement

The data that support the findings of this study are openly available in the heat_pump_control repository at https://github.com/aauphd2024/heat_pump_control. Complete reproducibility package of the study publicly available at [62].

7. Acknowledgments

We would like to thank Hisham Johra for his assistance in understanding the temporal changes in the building's thermodynamics. We gratefully acknowledge the partial funding of this research by the ERC Advanced Grant LASSO, the Villum Investigator Grant S4OS, and DIREC: Digital Research Centre Denmark. We also thank the European Union's Horizon 2020 research and innovation program and various partners for their support through the research project "Self-Assessment Towards Optimization of Building Energy" (SATO), grant agreement No. 957128.

References

[1] Eurostat (statistics explianed). https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Energy_consumption_in_households, 2024.

[2] GLOBAL STATUS REPORT for BUILDINGS and CONSTRUCTION, towards a Zero-Emissions, Efficient and Resilient Buildings and Construction Sector. https://globalabc.org/sites/default/files/2021-10/GABC_Buildings-GSR-2021_BOOK.pdf, 2021.

[3] M. Adolph, N. Kopmann, B. Lupulescu, and D. Müller. Adaptive control strategies for single room heating. *Energy and Buildings*, 68:771–778, 2014.

[4] Kim G. Larsen, Marius Mikučionis, Marco Muñiz, Jiří Srba, and Jakob Haahr Taankvist. Online and compositional learning of controllers with application to floor heating. In Marsha Chechik and Jean-François Raskin, editors, *Tools and Algorithms for the Construction and Analysis of Systems*, pages 244–259, Berlin, Heidelberg, 2016. Springer Berlin Heidelberg.

[5] Hessam Golmohamadi, Kim Guldstrand Larsen, Peter Gjøl Jensen, and Imran Riaz Hasrat. Optimization of power-to-heat flexibility for residential buildings in response to day-ahead electricity price. *Energy and Buildings*, 232:110665, 2021.

[6] Rune Juhl, Jan Kloppenborg Møller, and Henrik Madsen. Ctsmr-continuous time stochastic modeling in r, 2016. https://ctsm.info/ctsmr-reference.pdf.

[7] Alexandre David, Peter Gjøl Jensen, Kim Guldstrand Larsen, Marius Mikučionis, and Jakob Haahr Taankvist. Uppaal stratego. In *Tools and Algorithms for the Construction and Analysis of Systems: 21st International Conference, TACAS 2015, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2015, London, UK, April 11-18, 2015, Proceedings 21*, pages 206–211. Springer, 2015.

[8] Samuel Prívara, Jiří Cigler, Zdeněk Váňa, Frauke Oldewurtel, Carina Sagerschnig, and Eva Žáčeková. Building modeling as a crucial part for building predictive control. *Energy and Buildings*, 56:8–22, 2013.

[9] P.J.C. Vogler-Finck, R. Wisniewski, and P. Popovski. Reducing the carbon footprint of house heating through model predictive control – a simulation study in danish conditions. *Sustainable Cities and Society*, 42:558–573, 2018.

[10] Francesco Ferracuti, Alessandro Fonti, Lucio Ciabattoni, Stefano Pizzuti, Alessia Arteconi, Lieve Helsen, and Gabriele Comodi. Data-driven models for short-term thermal behaviour prediction in real buildings. *Applied Energy*, 204:1375–1387, 2017.

[11] Alessandro Fonti, Gabriele Comodi, Stefano Pizzuti, Alessia Arteconi, and Lieve Helsen. Low order grey-box models for short-term thermal behavior prediction in buildings. *Energy Procedia*, 105:2107–2112, 2017. 8th International Conference on Applied Energy, ICAE2016, 8-11 October 2016, Beijing, China.

[12] G. Reynders, J. Diriken, and D. Saelens. Quality of grey-box models and identified parameters as function of the accuracy of input and observation signals. *Energy and Buildings*, 82:263–274, 2014.

[13] Nivine Attoue, Isam Shahrour, Hussein Mroueh, and Rafic Younes. Determination of the optimal order of grey-box models for short-time prediction of buildings' thermal behavior. *Buildings*, 9(9), 2019.

[14] Lorenzo Nespoli, Vasco Medici, and Roman Rudel. Grey-box system identification of building thermal dynamics using only smart meter and air temperature data. In *Building Simulation Conference*, 2015.

[15] Imran Riaz Hasrat, Peter Gjøl Jensen, Kim Guldstrand Larsen, and Jiří Srba. End-to-end heat-pump control using continuous time stochastic modelling and uppaal stratego. In Yamine Aït-Ameur and Florin Crăciun, editors, *Theoretical Aspects of Software Engineering*, pages 363–380, Cham, 2022. Springer International Publishing.

[16] Imran Riaz Hasrat, Peter Gjøl Jensen, Kim Guldstrand Larsen, and Jiří Srba. A toolchain for domestic heat-pump control using uppaal stratego. *Science of Computer Programming*, 230:102987, 2023.

[17] GJ Ríos-Moreno, M Trejo-Perea, R Castañeda-Miranda, VM Hernández-Guzmán, and G Herrera-Ruiz. Modelling temperature in intelligent buildings by means of autoregressive models. *Automation in construction*, 16(5):713–722, 2007.

[18] B. Sohlberg and E.W. Jacobsen. Grey box modelling – branches and experiences. *IFAC Proceedings Volumes*, 41(2):11415–11420, 2008. 17th IFAC World Congress.

[19] Lennart Ljung. *System identification toolbox 7: Getting started guide*. The MathWorks, 2008.

[20] Hong Liu, Yuxin Wu, Baizhan Li, Yong Cheng, and Runming Yao. Seasonal variation of thermal sensations in residential buildings in the hot summer and cold winter zone of china. *Energy and Buildings*, 140:9–18, 2017.

[21] Virginia Gori, Phillip Biddulph, and Clifford A. Elwell. A bayesian dynamic method to estimate the thermophysical properties of building elements in all seasons, orientations and with reduced error. *Energies*, 11(4), 2018.

[22] Virginia Gori and Clifford A. Elwell. Estimation of thermophysical properties from in-situ measurements in all seasons: Quantifying and reducing errors using dynamic grey-box methods. *Energy and Buildings*, 167:290–300, 2018.

[23] Arnaldo Sepulveda, Liam Paull, Walid G Morsi, Howard Li, CP Diduch, and Liuchen Chang. A novel demand side management program using water heaters and particle swarm optimization. In *2010 IEEE Electrical Power & Energy Conference*, pages 1–5. IEEE, 2010.

[24] Liam Paull, Derek MacKay, Howard Li, and Liuchen Chang. Awater heater model for increased power system efficiency. In *2009 Canadian Conference on Electrical and Computer Engineering*, pages 731–734. IEEE, 2009.

[25] Shuai Lu, Nader Samaan, Ruisheng Diao, Marcelo Elizondo, Chunlian Jin, Ebony Mayhorn, Yu Zhang, and Harold Kirkham. Centralized and decentralized control for demand response. In *ISGT 2011*, pages 1–8. Ieee, 2011.

[26] Liam Paull, Howard Li, and Liuchen Chang. A novel domestic electric water heater model for a multi-objective demand side management program. *Electric Power Systems Research*, 80(12):1446–1451, 2010.

[27] Ruisheng Diao, Shuai Lu, Marcelo Elizondo, Ebony Mayhorn, Yu Zhang, and Nader Samaan. Electric water heater modeling and control strategies for demand response. In *2012 IEEE power and energy society general meeting*, pages 1–8. IEEE, 2012.

[28] Junji Kondoh, Ning Lu, and Donald J Hammerstrom. An evaluation of the water heater load potential for providing regulation service. In *2011 IEEE Power and Energy Society General Meeting*, pages 1–8. IEEE, 2011.

[29] Ruisheng Diao, Shuai Lu, Marcelo Elizondo, Ebony Mayhorn, Yu Zhang, and Nader Samaan. Electric water heater modeling and

control strategies for demand response. In *2012 IEEE power and energy society general meeting*, pages 1–8. IEEE, 2012.

[30] Abdul Atisam Farooq, Abdul Afram, Nicola Schulz, and Farrokh Janabi-Sharifi. Grey-box modeling of a low pressure electric boiler for domestic hot water system. *Applied Thermal Engineering*, 84:257–267, 2015.

[31] Simon Furbo. Heat storage for solar heating systems. *Educational Note, BYG.DTU U-071, ISSN 1396-4046*, 2005.

[32] Imran Riaz Hasrat, Peter Gjøl Jensen, Kim Guldstrand Larsen, and Jiří Srba. Modelling of hot water buffer tank and mixing loop for an intelligent heat pump control. In Alessandro Cimatti and Laura Titolo, editors, *Formal Methods for Industrial Critical Systems*, pages 113–130, Cham, 2023. Springer Nature Switzerland.

[33] X.Q. Zhai, X.L. Wang, H.T. Pei, Y. Yang, and R.Z. Wang. Experimental investigation and optimization of a ground source heat pump system under different indoor set temperatures. *Applied Thermal Engineering*, 48:105–116, 2012.

[34] Lei Xia, Zhenjun Ma, Craig McLauchlan, and Shugang Wang. Experimental investigation and control optimization of a ground source heat pump system. *Applied Thermal Engineering*, 127:70–80, 2017.

[35] P. Potočnik, B. Vidrih, A. Kitanovski, and E. Govekar. Analysis and optimization of thermal comfort in residential buildings by means of a weather-controlled air-to-water heat pump. *Building and Environment*, 140:68–79, 2018.

[36] Qiong Chen, Nan Li, and Wei Feng. Model predictive control optimization for rapid response and energy efficiency based on the state-space model of a radiant floor heating system. *Energy and Buildings*, 238:110832, 2021.

[37] Minsu Park, Yong Hwan Eom, and Min Soo Kim. Predictive optimization of the air flow rate for a residential heat pump in seasonal performance conditions. *International Journal of Refrigeration*, 131:51–60, 2021.

[38] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

[39] Zhe Wang and Tianzhen Hong. Reinforcement learning for building controls: The opportunities and challenges. *Applied Energy*, 269:115036, 2020.

[40] Liyuan Zhao, Ting Yang, Wei Li, and Albert Y. Zomaya. Deep reinforcement learning-based joint load scheduling for household multi-energy system. *Applied Energy*, 324:119346, 2022.

[41] Lissy Langer and Thomas Volling. A reinforcement learning approach to home energy management for modulating heat pumps and photovoltaic systems. *Applied Energy*, 327:120020, 2022.

[42] Paulo Lissa, Conor Deane, Michael Schukat, Federico Seri, Marcus Keane, and Enda Barrett. Deep reinforcement learning for home energy management system control. *Energy and AI*, 3:100043, 2021.

[43] Javier Arroyo, Carlo Manna, Fred Spiessens, and Lieve Helsen. Reinforced model predictive control (rl-mpc) for building energy management. *Applied Energy*, 309:118346, 2022.

[44] Christian Blad, Simon Bøgh, and Carsten Skovmose Kallesøe. Data-driven offline reinforcement learning for hvac-systems. *Energy*, 261:125290, 2022.

[45] C. Blad, S. Bøgh, C. Kallesøe, and Paul Raftery. A laboratory test of an offline-trained multi-agent reinforcement learning algorithm for heating systems. *Applied Energy*, 337:120807, 2023.

[46] Liangliang Chen, Fei Meng, and Ying Zhang. Fast human-in-the-loop control for hvac systems via meta-learning and model-based offline reinforcement learning. *IEEE Transactions on Sustainable Computing*, 8(3):504–521, 2023.

[47] A. Mugnini, F. Ferracuti, M. Lorenzetti, G. Comodi, and A. Arteconi. Advanced control techniques for chp-dh systems: A critical comparison of model predictive control and reinforcement learning. *Energy Conversion and Management: X*, 15:100264, 2022.

[48] Liangliang Chen, Fei Meng, and Ying Zhang. Mbrl-mc: An hvac control approach via combining model-based deep reinforcement learning and model predictive control. *IEEE Internet of Things Journal*, 9(19):19160–19173, 2022.

[49] Kamilla Heimar Andersen, Anna Marszal-Pomianowska, Benas Jokubauskis, and Per Kvols Heiselberg. Influence of temporal- and spatial resolutions on building performance simulation models: A danish residential building case study. In *Journal of Physics: Conference Series*, volume 2600, page 132007. IOP Publishing, 2023.

[50] Kamilla Heimar Andersen, Anna Marszal-Pomianowska, Henrik N Knudsen, Hicham Johra, Simon Pommerencke Melgaard, Marc Zein Dahl, Patrick Andersen Hundevad, and Per Kvols Heiselberg. Room-based indoor environment measurements and occupancy ground truth datasets from five residential apartments in a nordic climate. 2023.

[51] Manfred Jaeger, Peter Gjøl Jensen, Kim Guldstrand Larsen, Axel Legay, Sean Sedwards, and Jakob Haahr Taankvist. Teaching stratego to play ball: Optimal synthesis for continuous space mdps. In *Automated Technology for Verification and Analysis: 17th International Symposium, ATVA 2019, Taipei, Taiwan, October 28–31, 2019, Proceedings 17*, pages 81–97. Springer, 2019.

[52] Henrik Madsen Niels Rode Kristensen. Ctsm-continuous time stochastic modeling, 2003. http://www.imm.dtu.dk/~ctsm/UsersGuide.pdf.

[53] Gerd Behrmann, Alexandre David, Kim Guldstrand Larsen, John Håkansson, Paul Pettersson, Wang Yi, and Martijn Hendriks. Uppaal 4.0. 2006.

[54] Gerd Behrmann, Alexandre David, and Kim G Larsen. A tutorial on uppaal. *Formal methods for the design of real-time systems*, pages 200–236, 2004.

[55] Andreas Berre Eriksen, Harry Lahrmann, Kim Guldstrand Larsen, and Jakob Haahr Taankvist. Controlling signalized intersections using machine learning. *Transportation Research Procedia*, 48:987–997, 2020.

[56] Kim Guldstrand Larsen, Marius Mikučionis, and Jakob Haahr Taankvist. Safe and optimal adaptive cruise control. In *Correct System Design: Symposium in Honor of Ernst-Rüdiger Olderog on the Occasion of His 60th Birthday, Oldenburg, Germany, September 8-9, 2015, Proceedings*, pages 260–277. Springer, 2015.

[57] Filip Belić, Dražen Slišković, and Željko Hocenski. Detailed thermodynamic modeling of multi-zone buildings with resistive-capacitive method. *Energies*, 14(21), 2021.

[58] Nicola Cibin, Alessandro Tibo, Hessam Golmohamadi, Arne Skou, and Michele Albano. Machine learning-based algorithms to estimate thermal dynamics of residential buildings with energy flexibility. *Journal of Building Engineering*, 65:105683, 2023.

[59] Hassan Harb, Neven Boyanov, Luis Hernandez, Rita Streblow, and Dirk Müller. Development and validation of grey-box models for forecasting the thermal response of occupied buildings. *Energy and Buildings*, 117:199–207, 2016.

[60] Lucile Sarran, Kevin M. Smith, Christian A. Hviid, and Carsten Rode. Grey-box modelling and virtual sensors enabling continuous commissioning of hydronic floor heating. *Energy*, 261:125282, 2022.

[61] Argyris Oraiopoulos and Bianca Howard. One-step ahead modelling of building thermal dynamics: A comparison of backward selection and lasso approaches. International Building Performance Simulation Association: IBPSA, 2020.

[62] Imran R. Hasrat, Kamilla H. Andersen, Peter G. Jensen, Rasmus L. Jensen, Kim G. Larsen, and J. Srba. Artifact for the paper "towards model-driven heat pump control in a multi-story building". https://doi.org/10.5281/zenodo.14178322, 2024.

[63] Jesse Clifton and Eric Laber. Q-learning: Theory and applications. *Annual Review of Statistics and Its Application*, 7:279–301, 2020.

[64] Modern and safe machine learning solutions to improve the operation of heat pumps. https://www.cedar-heat.dk/, 2024.