Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Federal Department of the Environment,
Transport, Energy and Communications DETEC

**Swiss Federal Office of Energy**
Energy Research and Cleantech

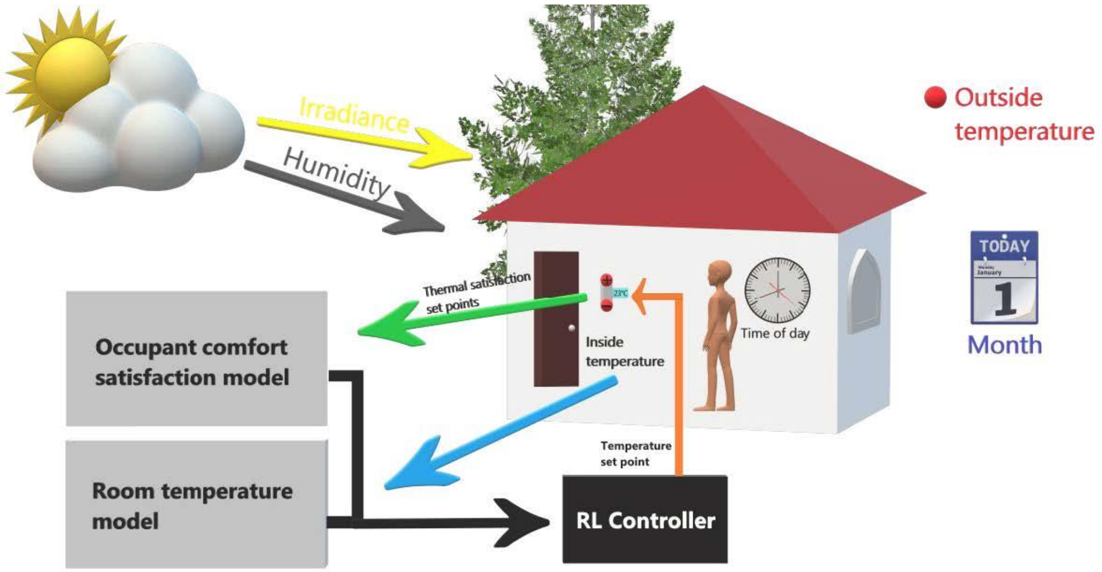
**Final report from 20 May 2025**

---

# UC-DPC

## Energy and comfort optimization in living and working environments through a user-centred data-driven predictive control

---



**Publisher:**
Swiss Federal Office of Energy SFOE
Energy Research and Cleantech
CH-3003 Berne
www.energy-research.ch

**Subsidy recipients:**
Empa
Überlandstrasse 129
8600 CH-Dübendorf
https://www.empa.ch/

**Authors:**
Philipp Heer, Empa, philipp.heer@empa.ch
Dr. Bratislav Svetozarevic, Empa, Bratislav.svetozarevic@empa.ch
Dr. Arnab Chatterjee, Empa, arnab.chatterjee@empa.ch
Dr. Jannis Wernery, Empa, jannis.wernery@empa.ch

**SFOE project coordinators:**
Andreas Eckmanns, andreas.eckmanns@bfe.admin.ch
Dr. Anne-Kathrin Faust, anne-kathrin.faust@bfe.admin.ch
Nadege Vetterli, nadege.vetterli@anex.ch
Martin Ménard, Lowtechlab, menard@lowtechlab.ch

**SFOE contract number:** SI/502188-01

# Summary

Detecting the user needs for thermal comfort and addressing them within the building automation systems is one of the most critical barriers for achieving energy-efficient living and working indoor environments. Current building automation systems are designed to either provide a “general” comfort defined by the standards [1,2] or simply “respond” to users' setpoint temperature changes, for example, as adjusted on a wall-mounted thermostat. In either of these approaches, the users’ needs for comfort are not profoundly considered within the building controllers. Thus, the opportunity to provide a user-specific comfort response is missed. More precisely, the opportunity to learn user-specific comfort needs and provide such comfort at minimal energy consumption is missed.

In this project, new control algorithms were developed that leverage user data on thermal comfort needs to provide user-specific comfort throughout the day automatically. The project consisted of three main parts: i) decision about what data (from which sensors and/or which survey questions) to take into consideration to determine user needs for comfort, ii) developing a comfort satisfaction model, and iii) developing the control algorithm and testing it in simulation and the real environment at the NEST building at Empa in Dübendorf.

To this objective the project was divided into two primary phases. Phase 1 consisted of data collection from two single office units (Rooms 183 and 184) from Nest Sprint unit. During the data collection which took place in the winter months, the spaces were occupied by two test subjects and Thermal Preference (TP) votes were gathered periodically using Survey app deployed in desk-mounted tablet. In addition, environmental data consisting of outdoor variables and indoor variables and physiological data were recorded. Using the data gathered, Support Vector Machine (SVM) based Personal Thermal Comfort Models (PTCM) were developed for both the subjects. These PTCMs provided 89% and 83% accurate TP predictions for the two participants respectively.

The developed PTCM was then embedded in a DRL-based controller framework modulating the heating valve states independently in each of the rooms with the objective to provide personalized thermal comfort and minimize heating energy usage. Thus, phase 2 of the UC-DPC project involved deploying the PTCM embedded Deep Reinforcement Learning (DRL) controller framework in the Nest Sprint unit and validating its performance through periodic collection of TP votes as in phase 1. The PTCM-DRL controller was quite instrumental in providing personalized indoor environmental conditions for both the subjects and helped in reducing the heating energy consumption of both the rooms in the Nest Sprint unit. It underscores the need for large-scale studies with diverse user profiles to realize the full potential of PTCM-embedded DRL controllers to provide energy-efficient and personalized indoor conditions in Swiss buildings.

# Zusammenfassung

Die Erfassung der Bedürfnisse der Nutzer nach thermischem Komfort und deren Berücksichtigung in Gebäudeautomationssystemen ist eine der kritischsten Hürden für die Schaffung energieeffizienter Wohn- und Arbeitsumgebungen. Aktuelle Gebäudeautomationssysteme sind darauf ausgelegt, entweder einen durch Standards [1,2] definierten "allgemeinen" Komfort bereitzustellen, oder einfach auf Änderungen der Solltemperatur durch die Nutzer zu "reagieren", beispielsweise über einen wandmontierten Thermostat. Bei keinem dieser Ansätze werden die Komfortbedürfnisse der Nutzer tiefgreifend in die Gebäudesteuerung einbezogen. Somit wird die Möglichkeit verpasst, einen nutzerspezifischen Komfort zu bieten. Genauer gesagt wird die Chance verpasst, nutzerspezifische Komfortbedürfnisse zu erlernen und diesen Komfort bei minimalem Energieverbrauch zu gewährleisten.

In diesem Projekt wurden neue Steuerungsalgorithmen entwickelt, die Nutzerdaten zu thermischen Komfortbedürfnissen nutzen, um automatisch nutzerspezifischen Komfort über den Tag hinweg zu bieten. Das Projekt bestand aus drei Hauptteilen: i) Entscheidung darüber, welche Daten (von welchen Sensoren und/oder welchen Umfragefragen) zur Ermittlung der Komfortbedürfnisse der Nutzer berücksichtigt werden sollen, ii) Entwicklung eines Komfortzufriedenheitsmodells und iii) Entwicklung des Steuerungsalgorithmus und dessen Erprobung in Simulation und realer Umgebung im NEST-Gebäude der Empa in Dübendorf.

Zu diesem Zweck wurde das Projekt in zwei Hauptphasen unterteilt. Phase 1 bestand aus der Datenerhebung in zwei Einzelbüroeinheiten (Räume 183 und 184) der Nest Sprint-Einheit. Während der Datenerhebung, die in den Wintermonaten stattfand, wurden die Räume von zwei Testpersonen besetzt, und TP-Bewertungen (Thermische Präferenz) wurden periodisch über eine auf einem Tisch-Tablet installierte Umfrage-App erfasst. Zusätzlich wurden Umweltdaten bestehend aus Außen- und Innenvariablen sowie physiologische Daten aufgezeichnet. Mit den gesammelten Daten wurden Support Vector Machine (SVM) basierte PTCMs (Personalisierte Thermische Komfortmodelle) für beide Testpersonen entwickelt. Diese PTCMs lieferten für die beiden Teilnehmer TP-Vorhersagen mit einer Genauigkeit von 89 % bzw. 83 %.

Das entwickelte PTCM wurde dann in ein Deep Reinforcement Learning (DRL) basierte Controller Framework eingebettet, das die Heizungsventilzustände in jedem der Räume unabhängig moduliert, mit dem Ziel, personalisierten thermischen Komfort zu bieten und den Heizenergieverbrauch zu minimieren. Phase 2 des UC-DPC-Projekts beinhaltete somit die Implementierung des PTCM-eingebetteten DRL-Controller-Frameworks in der Nest Sprint-Einheit und die Validierung seiner Leistung durch periodische Sammlung von TP-Bewertungen wie in Phase 1. Der PTCM-DRL-Controller war maßgeblich an der Bereitstellung personalisierter Raumklimabedingungen für beide Testpersonen beteiligt und half, den Heizenergieverbrauch beider Räume in der Nest Sprint-Einheit zu reduzieren. Dies unterstreicht die Notwendigkeit von großangelegten Studien mit vielfältigen Nutzerprofilen, um das volle Potenzial von PTCM-eingebetteten DRL-Controllern zur Bereitstellung energieeffizienter und personalisierter Innenraumbedingungen in Schweizer Gebäuden zu realisieren.

# Résumé
La détection des besoins des utilisateurs en matière de confort thermique et leur prise en compte dans les systèmes d'automatisation des bâtiments constituent l'un des obstacles les plus critiques à la réalisation d'environnements intérieurs de vie et de travail économes en énergie. Les systèmes d'automatisation des bâtiments actuels sont conçus soit pour fournir un confort "général" défini par les normes [1,2], soit simplement pour "répondre" aux changements de température de consigne des utilisateurs, par exemple, tels qu'ajustés sur un thermostat mural. Dans aucune de ces approches, les besoins des utilisateurs en matière de confort ne sont profondément pris en compte dans les contrôleurs de bâtiments. Ainsi, l'opportunité de fournir une réponse de confort spécifique à l'utilisateur est manquée. Plus précisément, l'opportunité d'apprendre les besoins de confort spécifiques à l'utilisateur et de fournir ce confort avec une consommation d'énergie minimale est manquée.

Dans ce projet, de nouveaux algorithmes de contrôle ont été développés qui exploitent les données des utilisateurs sur les besoins de confort thermique pour fournir automatiquement un confort spécifique à l'utilisateur tout au long de la journée. Le projet se composait de trois parties principales : i) décision sur quelles données (provenant de quels capteurs et/ou quelles questions d'enquête) prendre en considération pour déterminer les besoins des utilisateurs en matière de confort, ii) développement d'un modèle de satisfaction du confort, et iii) développement de l'algorithme de contrôle et test dans la simulation et l'environnement réel au bâtiment NEST de l'Empa à Dübendorf.

Pour atteindre cet objectif, le projet a été divisé en deux phases principales. La phase 1 consistait en la collecte de données provenant de deux bureaux individuels (Salles 183 et 184) de l'unité Nest Sprint. Pendant la collecte de données qui a eu lieu durant les mois d'hiver, les espaces étaient occupés par deux sujets tests et les votes de TP (Préférence Thermique) étaient recueillis périodiquement à l'aide d'une application de sondage déployée sur une tablette fixée au bureau. En outre, des données environnementales comprenant des variables extérieures et intérieures ainsi que des données physiologiques ont été enregistrées. À l'aide des données recueillies, des PTCM (Modèles de Confort Thermique Personnalisés) basés sur SVM ont été développés pour les deux sujets. Ces PTCM ont fourni des prédictions de TP précises à 89% et 83% pour ID1 et ID2 respectivement.

Le PTCM développé a ensuite été intégré dans un cadre de contrôleur basé sur l'apprentissage par renforcement profond (DRL) modulant les états des vannes de chauffage indépendamment dans chacune des salles avec l'objectif de fournir un confort thermique personnalisé et de minimiser l'utilisation d'énergie de chauffage. Ainsi, la phase 2 du projet UC-DPC impliquait le déploiement du cadre de contrôleur DRL intégrant le PTCM dans l'unité Nest Sprint et la validation de ses performances par la collecte périodique de votes de TP comme dans la phase 1. Le contrôleur PTCM-DRL a été très efficace pour fournir des conditions environnementales intérieures personnalisées pour les deux sujets et a aidé à réduire la consommation d'énergie de chauffage des deux salles dans l'unité Nest Sprint. Cela souligne la nécessité d'études à grande échelle avec divers profils d'utilisateurs pour réaliser le plein potentiel des contrôleurs DRL intégrant des PTCM afin de fournir des conditions intérieures économes en énergie et personnalisées dans les bâtiments suisses.

# Main findings

- Outdoor environmental variables (air temperature, relative humidity), indoor environmental variables (air temperature, relative humidity) and physiological variables (heart rate variability) are crucial parameters for PTCMs
- SVM-based PTCMs provided good prediction accuracy (89% and 83% for ID1 and ID2) of the TP responses of the participants. The prediction accuracy was better (>90%) on majority classes and slightly worse (<70%) on minority classes
- PTCM embedded DRL controller ensured personalization in the variation of the heating valve states for ID1 and ID2 (Room 184 and Room 183 respectively). Additionally, from the energy usage perspective, it was not only effective in lowering the baseline energy requirements but also it ensured reduced heating energy usage at higher HDD values (for both rooms).
- Future research should incorporate more varied indoor environmental conditions in association with diverse user profiles as test subjects to train more robust PTCMs capable of accurately predicting responses across the full spectrum of thermal preferences.

# Contents

# List of figures

# List of tables

# List of abbreviations

| | |
|---|---|
| BAC | Building automation controller |
| BAS | Building automation system |
| BMS | Building management system |
| DDPG | Deep deterministic policy gradient |
| DRL | Deep reinforcement learning |
| EV | Electric Vehicle |
| H | Relative Humidity |
| HDD | Heating Degree Days |
| HRV | Heart Rate Variability |
| HVAC | Heating, ventilation, and air-conditioning |
| $G_h$ | Global Horizontal Irradiation |
| MIMO | Multiple Input Multiple Output |
| ML | Machine Learning |
| MPC | Model predictive controller |
| NN | Neural networks |
| PMV | Predicted Mean Vote |
| PPD | Predicted Percentage of Dissatisfied |
| PPO | Proximal policy optimization |
| PTCM | Personal Thermal Comfort Model |
| RL | Reinforcement Learning |
| RB | Rule-based |
| RF | Random Forest |
| SFOE | Swiss Federal Office of Energy |
| SP | Set-point |
| SVM | Support Vector Machine |
| T | Temperature |
| TP | Thermal Preference |
| WP | Work package |
| VS | Valve State |

# 1 Introduction

This chapter establishes the context for developing user-centered thermal comfort control systems, highlighting that current building automation systems fail to consider individual occupant needs and instead rely on standardized comfort bounds. The project aims to develop data-driven DRL algorithms that can learn personal thermal preferences and provide customized comfort while minimizing energy consumption, addressing the critical gap between occupant satisfaction and energy efficiency in modern buildings.

## 1.1 Context and motivation

Indoor thermal comfort is one of the most important aspects of modern building control. First, people spend almost 90% of their time indoors [3], and their thermal comfort is associated with health [4], working productivity [5,6] and well-being [7]. Therefore, it is essential to ensure occupant thermal comfort at all times of the day. Second, the comfort shall be satisfied with the least possible energy consumption in buildings due to the strong contribution of their operation to climate change [8]. It follows that, occupant comfort and building energy consumption are strongly correlated. Furthermore, both are determined by the control algorithms implemented in building automation systems (BAS). So, the structure, type, and function of the control algorithm strongly impact the achieved comfort level and energy consumption in buildings [9].

Classical building automation systems (BAS), and in particular room temperature controllers as one of their most important parts related to indoor thermal comfort, are focusing on regulation of heating, ventilation, and air-conditioning (HVAC) systems while keeping the indoor temperature within the bounds defined by thermal comfort standards [1,2]. These controllers are typically rule-based (RB), defined by a set of if-then rules, and are used in the majority (more than 90%) of today's buildings [10]. However, there are several drawbacks and issues related to them when it comes to achieving optimal, user-tailored comfort at minimal energy requirements.

Firstly, regarding user comfort, as indicated above, RB controllers are designed to satisfy the comfort bounds defined by standards. However, these bounds are defined by HVAC standards and do not take the individual needs of occupants for comfort. The results of laboratory studies and realistic investigations were summarized in the ASHRAE 55 standard (last update 2017) [2]. The standard contains a variety of methods for determining the indoor climate using measurable parameters. Interestingly, however, laboratory tests and investigations contradict each other clearly in some cases [11]. The standard recommended designs are based on long-term laboratory studies under steady-state indoor conditions. ASHRAE 55-2002 stipulates that the variation in the operative temperature cannot exceed 2.2 °C during any 1 hour. The same recommendations as per ISO 7730 stands at 2 °C during any 1 hour. However, it also specifies that this fluctuation requirement shall only apply to the situation where the indoor environment is not under the direct control of the occupant. Consequently, the target of HVAC control has been intended to maintain the indoor climate relatively constant. So, it turns out to be difficult to reduce the feeling of comfort to just a few physical factors. Individual factors such as experience, expectation, health or mood are neglected with this approach. In summary, it can be said that the individual occupant need for comfort must be considered to sustainably reduce the energy consumption of buildings.

Secondly, RB controllers do not typically guarantee the optimality of performance in their classical use in the building automation industry. Even though RB controllers could provide optimal performance, obtaining such controllers would require complex engineering [10]. Thus, the effort, defined as expert personal cost and time, typically outweighs the benefits for this type of controllers. Hence, these classical building automation controllers (BAC) deliver typically only sub-optimal performance, which means sub-optimal satisfaction of comfort bounds and sub-optimal energy performance. Advanced building controllers, such as model predictive controllers (MPC), typically require building models of a certain quality, which is also expensive in terms of expert personal time and costs. Hence, MPC, as one of the most popular representatives of advanced controllers and a mature technology, has not been widely adopted into building automation yet [12].

Due to the recent increase in the availability of sensors data in buildings (the number of sensors in buildings connected to a centralized database) and due to advances in affordable high-performance computational power and machine learning (ML), in particular deep learning, data-driven ML-based building controllers came to the forefront of interest of academia and industry [13]. Several works showed that it is possible to create building models directly from the data, for example, by using neural networks, and then use these models as simulation environments for reinforcement learning algorithms to obtain (learn) the control policy [14]–[17]. Compared to RB and MPC controller designs, the advantage of this approach is that it could achieve a close-to-optimal building control with a considerably lower engineering effort. Also, Svetozarevic et al. [14] have demonstrated the successful application of deep reinforcement learning (DRL) to room temperature control of an actual building (the DFAB HOUSE at the NEST building [18]) at Empa in Dübendorf. The DRL controller was able to achieve up to 30% energy saving while maintaining similar comfort levels to a conventional RB controller. Similarly, Di Natale et al. [15] managed to achieve both comfort improvement and energy savings at the UMAR unit at the NEST building [19]. Therefore, this approach demonstrated potential for obtaining (near-)optimal building control policies and addressing both comfort satisfaction and energy minimization. Some of the challenges to this approach are related to the size of the datasets required to build good building models and an issue of generalization of neural networks, which are also the core questions in deep learning [20]. However, these issues are not dominating the performance of DRL-based building controllers, and DRL-based control is a very active field of research [16]. Thus, further advances in DRL-based control are expected during the duration of this project.

Overall, as discussed above, data-driven ML and DRL methods are promising approaches to room temperature control at minimal energy use, which is at the core of this project (WP4 and WP5). However, the other part of the problem addressed in this project, and discussed above, is achieving individual, user-tailored, or user-centred comfort control. This understands two things: i) decision about what data (from which sensors and/or which survey questions) to take into consideration to determine user needs for comfort (WP1 and WP2) and ii) building the comfort satisfaction model (WP3). Personal comfort models have become a subject of study in the last few years, with overviews given by Kim and coworkers and André and colleagues [21,22]. Personal comfort can be queried by different metrics through user surveys and physiological data which will be outlined later. However, there is no clear consensus on which user comfort metrics and physiological parameters to use and how to automate the acquisition of such data. Also, many studies on physiological data work with setups that cannot be integrated in a non-invasive way (i.e., setups that need to be attached to the user and require large and expensive recording equipment). Hence, a suitable method to do so needs to be evaluated if physiological data shall be used.

Regarding the comfort satisfaction modelling, there are two major directions to this: physics-based (also physics-inspired) thermal comfort models and data-driven models. The first type of models relies on modelling the effects related to user comfort, such as indoor air temperature, mean radiant temperature and relative humidity, but also metabolic rate and clothing insulation. Typical examples are Predicted Mean Vote (PMV) and Predicted Percentage of Dissatisfied (PPD), which are based on ISO and ASHRAE standards [1,2]. The second type of models are directly derived from user data using some of the ML methods. As the focus of this project is on data-driven methods while avoiding complex physics-based modelling, we will focus on data-driven methods for occupant comfort modelling, such as random forests [23] and neural networks [24]. Multiple studies have experimented with physiological measurements for thermal comfort assessment. Nkurikiyeyezu et al. [25, 26] concluded that Heart rate variability (HRV) can accurately predict thermal comfort with up to 93.7% accuracy, suggesting it could be used in automatic real-time thermal comfort controllers. In a separate study by Yao et al. [27], Skin temperature, ECG ratio, and EEG power were found to be sensitive to ambient temperatures and thermal sensations, suggesting their integration in thermal comfort studies. Based on a review paper from Mansi et al. [28], the conclusion was that the wearable sensors can effectively measure human physiological indices and can potentially be correlated with thermal comfort. Thus, in terms of input data, we will consider two models: one based on a standard set of sensor data (outdoor temperature, outdoor humidity, solar irradiation, and indoor temperature) and another based on the extended data set, which will involve measuring the skin temperature and/or heart rate based on recommendations from literature [25-28].

## 1.2 Project purpose

The project aims to investigate, develop, test, and analyse a learnt user-centred room temperature control algorithm containing a personalized room temperature comfort model.

**Preliminary Groundwork for UC-DPC Project Development (WP1)**
In the initial phase workshops are to be conducted by UES focused on collecting essential framework conditions and system requirements through stakeholder collaboration. These workshops, involving industry experts are centred on sharing experiences and expertise to inform the development of both a user survey and system requirements guide. The outcomes of these collaborative sessions would be condensed into comprehensive guidelines pivotal for the user survey execution and the UC-DPC's subsequent implementation and field testing.

**How to detect user comfort (WP2)**
The project's initial milestones ensures that the development aligns with industry standards and user-centric design principles. The purpose of the user-centred aspect of the project is to reliably detect the comfort of building users in a minimally invasive way. The goal is to automate the process to get user feedback as much as possible to integrate it easily with the personal user comfort model. The process of acquiring user feedback needs to be in line with the requirements of the different stakeholders: the users, the facility managers and buildings owners.

**Development of a user comfort satisfaction model (WP3)**
In this part of the project, we use the data collected from the users at NEST to create the comfort satisfaction models. In addition to these models, we defined the artificial users to be able to generate larger amounts of synthetic data. Overall, the comfort satisfaction models are an essential part of the project. First, they capture the user need for comfort in the form of desired indoor room temperature. Second, this knowledge is integrated and used within the control algorithm to predict the preferred indoor temperature throughout the day for each user. Therefore, the user comfort satisfaction model links the user data (input) to the control policy (output).

**Development of a user-centred room temperature control algorithm (WP4)**
The control algorithm (or several of them) is the central outcome of this project. Both user comfort and building energy consumption depend on the control algorithm. At the end of the project, the successful control algorithm shall demonstrate the user-tailored (i.e. user-centred) comfort control at minimal energy used. The control algorithm is supposed to be implemented within BAS at NEST within WP4. The algorithm will be based on DRL, and due to some already successful demonstrations of DRL-based building control at the NEST building at Empa [14,15], there is potential for the success also of the algorithm derived in this project.

**Dissemination of UC-DPC Project (WP5)**
Throughout the dissemination phase of the UC-DPC project, efforts led by UES would be focused on sharing the research findings and technological advancements. This involved publishing scientific papers in reputable journals and presenting the work at key conferences related to applied machine learning and building automation. Public engagement would be facilitated through press releases, updates in the NEST newsletter, and active communication on various social media channels. In the context of WP5, final meetings with the SFOE would be conducted to inform and incorporate feedback into project planning and culminated the WP5 with a comprehensive final report. This final report when successfully accepted by the SFOE would mark the project's readiness for technology transfer and practical application.

## 1.3 Overview and timeline

The Gantt chart below in Table 1 shows the timeline of the different phases of the project.

Table 1: Project timeline with the different measurement phases

| Year | 2023 | | | | 2024 | | | | | | | | | | | | 2025 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Month | 9 | 10 | 11 | 12 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 1 | 2 | 3 | 4 |
| *Phase 1 - Calibration* | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | | | | | | | | | |
| 1.1 Preparation | ■ | | | | | | | | | | | | | | | | | | | |
| 1.2 Data collection (measurements) | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | | | | | | | | | | | |
| 1.3 Testing of SVM classification | | | | | | | | ■ | | ■ | | | | | | | | | | |
| 1.4 Data analysis and conclusions | | | | | | | | ■ | ■ | ■ | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |
| *Phase 2 - Intervention* | | | | | | | | | | | | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ | ■ |
| 2.1 PTCM model calibration and controller preparation in simulation | | | | | | | | | | | | ■ | ■ | ■ | | | | | | |
| 2.2 Controller implementation and testing of BMS | | | | | | | | | | | | | | | ■ | | | | | |
| 2.3 Data collection (measurements) | | | | | | | | | | | | | | | ■ | ■ | ■ | ■ | | |
| 2.4 Data analysis | | | | | | | | | | | | | | | | | | ■ | ■ | |
| 2.5 Conclusions and dissemination | | | | | | | | | | | | | | | | | | | ■ | ■ |

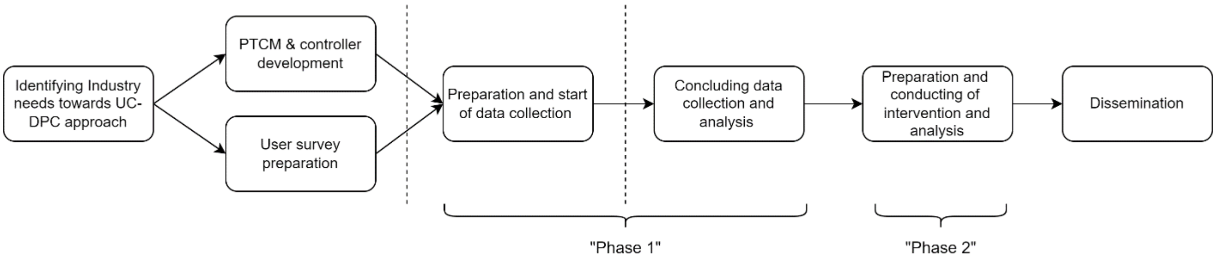



Figure 1: Overview of the project steps

## 1.4 Objectives

**1) Preliminary Groundwork for UC-DPC Project Development (WP1)**

In this initial phase of the project the following outcomes are to be expected:

- Organization of workshops aimed at outlining critical framework conditions necessary for developing a user survey
- Deliver guidelines for implementing user surveys

**2) How to detect user comfort (WP2)**

This part has the following objectives:

- Building of the user survey prototype
- Evaluation and adaptation of user survey prototype.

Overall, this should lead to a process that can automatically determine personal user comfort and integrate it with the user comfort satisfaction model.

**3) Development of a user comfort satisfaction model (WP3)**

This part has a general objective of creating user comfort satisfaction models based on data. The following aspects of the models shall be explored:

- Source of data
    - synthetic data from the artificial agents
    - data from the users of NEST

- Datasets:
    - Standard: outdoor temperature, outdoor humidity, solar irradiation, and indoor temperature
    - Extended: standard + skin temperature and/or heart rate.
- Dataset size: What is the minimum dataset size needed to successfully create the comfort satisfaction model?
- Output: Classification and regression.

**4) Development of a user-centered room temperature control algorithm (WP4)**
The following control performance indicators shall be examined:
- learning performance indicators (number of epochs needed, best reward, etc.)
- control algorithm performance in simulation (stability, robustness)
- control algorithm performance on a real building with the users (stability, robustness).

We will also compare the performance of the DRL-based algorithm to conventional RB algorithms. The comfort satisfaction will be measured as the amount of time multiplied with the indoor temperature deviation from the comfort bounds or set point in Kelvin-hours (Kh). The difference between algorithms will be measured in energy used in [kWh] or energy costs [CHF]. We expect that DRL-control bits RB controller in terms of less comfort dissatisfaction and less energy used.

**5) Dissemination of UC-DPC Project (WP5)**
In the final phase of the project the following are the expected outcomes:
- Publication of the results of the project in scientific journals and presentations at relevant conferences
- Final meeting with SFOE and submission of the final project report

# 2 Previous Approaches and Learnings

This chapter reviews foundational work in data-driven DRL control for thermal comfort, demonstrating that previous studies achieved significant energy savings (17-42%) and comfort improvements (19%) through personalized control systems. The chapter outlines the development of artificial agents, user comfort satisfaction models using random forests and neural networks and establishes the control system framework that integrates comfort satisfaction models with room temperature prediction and DRL-based controllers.

In this chapter, we delve into the foundational work that underpins our current project, reflecting on the earlier methodologies that were employed. This retrospective analysis serves not only to document our journey but also to provide valuable insights into the evolution of our approach and the lessons learned along the way. The previous work, conducted with alternative strategies, laid the groundwork for our current direction. It is essential to acknowledge these former efforts as they have significantly influenced the project's trajectory, contributing to a deeper understanding of the challenges and complexities inherent in creating user-specific thermal comfort through building automation systems. This chapter will present a detailed account of the initial methodologies, including the rationale behind their selection, the execution of related experiments, and the analysis of the outcomes. We will also discuss the pivot points, and the critical decisions made that steered our project towards the adoption of the data-driven Deep Reinforcement Learning approach that is now at the forefront of our research.

## 2.1 Demonstration Case: Data-Driven DRL Control for User Specific Thermal Comfort

In Svetozarevic et al. [14], the author presents a compelling case for the use of data-driven Deep Reinforcement Learning (DRL) to optimize building automation systems for energy efficiency and user comfort. By focusing on a fully data-driven approach, the study shows the potential of DRL-based methods in handling Multiple Input Multiple Output (MIMO) problems, such as the joint control of room temperature and bidirectional Electric Vehicle (EV) charging. The room temperature model was a pivotal component, utilizing advanced predictive algorithms to intelligently regulate the indoor climate. The model's performance was quantitatively assessed, revealing a commendable 20% reduction in energy usage for heating, ventilation, and air conditioning (HVAC) compared to the baseline. Furthermore, the model achieved a 90% accuracy rate in predicting room temperature within a comfortable range for occupants, a testament to its precision. This was especially notable during peak load conditions where the model sustained comfort levels with less than a 2% deviation from the setpoint. The key takeaways from the study that are directly applicable to our project include:

1) *Energy and Cost Savings:* The DRL agents demonstrated significant reductions in energy consumption and costs, achieving an average of 17% energy savings and up to 42% cost savings.
2) *Enhanced Comfort:* The approach resulted in a 19% improvement in comfort satisfaction, by learning and adapting to the specific requirements of the building and its occupants.
3) *Transferability:* The DRL policy trained in simulation was successfully transferred to a real-world scenario, indicating the approach's robustness and adaptability.
4) *Model-Free Potential:* The study suggests that the DRL-based method could be model-free, relying on data-driven policies that are potentially transferable to other buildings or rooms.

The goals of our project align closely with the findings of the reference study. We aim to develop control algorithms that not only meet the general standards of thermal comfort but also cater to the specific needs of individual users within their living and working environments. To integrate the insights from the reference study into our demonstration case, we can follow several steps:

1) *Data Acquisition:* Like the reference study, we will leverage user-specific data through surveys and physiological monitoring to inform our model. By using the SurveyApp and devices such as the Fitbit Charge and Greenteg Core, we can collect high-quality, user-centric data.
2) *Model Development:* We will develop a comfort satisfaction model informed by the collected data. This model will serve as the basis for training our agents. The comfort model will be designed to understand and predict individual preferences for thermal comfort.
3) *Control Algorithm:* The algorithm will be trained using the comfort satisfaction model. We will employ a similar architectural structure for the Neural Networks (NN) and DRL as the reference study, ensuring that our control policies are founded on proven methodologies.
4) *Simulation and Real Environment Testing:* Initially, our DRL-based control algorithms will be tested in a simulated environment to ensure that they are capable of learning effective policies. Following successful simulation results, we will implement the control algorithm in the real environment of the NEST building at Empa in Dübendorf.
5) *Energy Conservation and Comfort Optimization:* By considering the user-specific data, our control algorithm will be fine-tuned to minimize energy consumption while providing personalized thermal comfort. This will directly address the user's needs and potentially exceed the performance of general BAS.

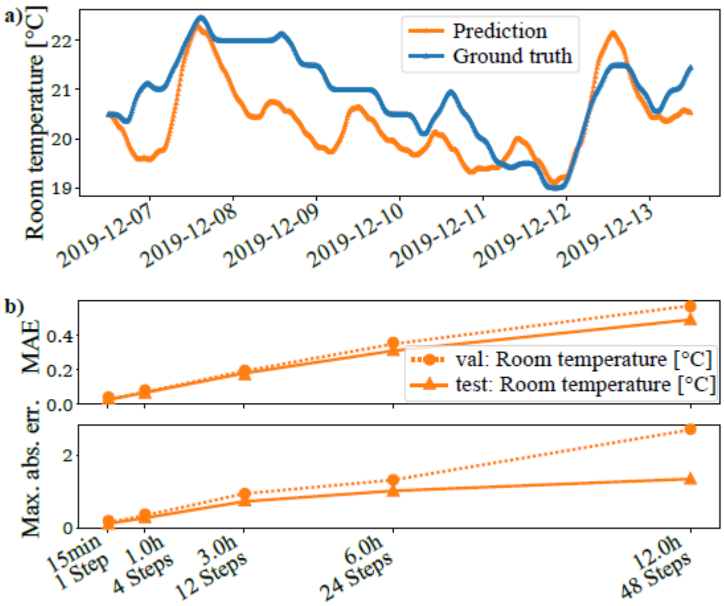



Figure 2: Room temperature prediction models: Comparison of the piece-wise linear model and the RNN as trained for the weather model. a) Qualitative comparison - example. b) Quantitative evaluation [14]

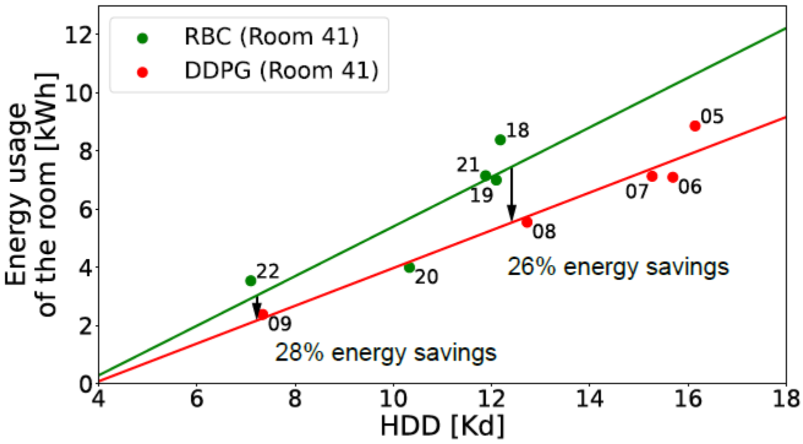



Figure 3: Experimental comparison of the DDPG and RB controller performances in term of required energy over the heating degree days. The DDPG provides in average 27% energy saving and better comfort at the same time. The numbers in the plot correspond to days in February 2020 [14]

The demonstration case in our project stands to benefit greatly from the methods and results presented in the reference study. By adopting a similar data-driven DRL approach, we aim to push the boundaries of building automation to create a system that is not only energy-efficient but also responsive to the nuanced comfort needs of individual users.

## 2.2 Artificial agents

In the first iterations of the UC-DPC project, we developed two artificial agents that are represented by a set of rules. One of them is shown in Table 3. These agents, based on the defined rules, give a satisfaction set point for a given outside conditions. With the current room temperature, we can get an answer about the current agents' thermal comfort satisfaction (cool, satisfied or warm). Summer and winter periods are divided because it also depends on the season, whether it is a period of cooling or heating in buildings. This is important because it is impossible to turn on heating and cooling at the same time, and thus it must be clearly defined which period is in question. This data, combined with the previously mentioned standard data set, is used to train personal comfort satisfaction models.

An overview of the variables used for artificial agents is in Table 3. Other assumptions and rules are:

*Assumptions:*
- Summer/Spring period is from April to September
- Winter/Autumn period is from October to March

*Agent intra-day rules:*
1. If I increased for > 300 and stayed for 30 min, then lower the SP for 0.5 °C - 1 °C.
2. If the morning Tout was 10 °C lower than the noon Tout, then lower set point for 1 °C.
3. If the morning Tout was 10 °C higher than the noon Tout, then increase set point for 1 °C.

Table 2: Overview of variables used in the model and their corresponding mathematical spaces

| Variable name, Unit | Symbol | Interval |
|---|---|---|
| Outside temperature [°C] | $T_{out}$ | [-15.0,40.0] |
| Room temperature [°C] | $T_{in}$ | [10.0,40.0] |
| Solar Irradiance [$Wm^{-2}$] | $I$ | [0.0,1300.0] |
| Outside Humidity [%] | $H$ | [0.0,100.0] [1ex] |

Table 3: Set of rules of the artificial agents

| Season | Solar Irradiation | Humidity | Outside Temperature | Set Point |
|---|---|---|---|---|
| Summer/Spring | Clear Sky e.g $I > 1000$ | Low $H < 60\%$ | $T_{out} > 33\,°C$ | 22 °C |
| Summer/Spring | Clear Sky e.g $I > 1000$ | Low $H < 60\%$ | $18\,°C < T_{out} < 33\,°C$ | 21.5 °C |
| Summer/Spring | Clear Sky e.g $I > 1000$ | Low $H < 60\%$ | $T_{out} < 18\,°C$ | 22.5 °C |
| Summer/Spring | Cloudy e.g $I < 1000$ | High $H > 60\%$ | $T_{out} > 33\,°C$ | 21 °C |
| Summer/Spring | Cloudy e.g $I < 1000$ | High $H > 60\%$ | $18\,°C < T_{out} < 33\,°C$ | 21.5 °C |
| Summer/Spring | Cloudy e.g $I < 1000$ | High $H > 60\%$ | $T_{out} < 18\,°C$ | 22.5 °C |
| Summer/Spring | Clear Sky e.g $I > 1000$ | High $H > 60\%$ | $T_{out} > 33\,°C$ | 21 °C |
| Summer/Spring | Clear Sky e.g $I > 1000$ | High $H > 60\%$ | $18\,°C < T_{out} < 33\,°C$ | 22 °C |
| Summer/Spring | Clear Sky e.g $I > 1000$ | High $H > 60\%$ | $T_{out} < 18\,°C$ | 22.5 °C |
| Summer/Spring | Cloudy e.g $I < 1000$ | Low $H < 60\%$ | $T_{out} > 33\,°C$ | 22.5 °C |
| Summer/Spring | Cloudy e.g $I < 1000$ | Low $H < 60\%$ | $18\,°C < T_{out} < 33\,°C$ | 23 °C |
| Summer/Spring | Cloudy e.g $I < 1000$ | Low $H < 60\%$ | $T_{out} < 18\,°C$ | 23.5 °C |
| Winter/Autumn | Clear Sky e.g $I > 500$ | Low $H < 70\%$ | $T_{out} > 15\,°C$ | 23 °C |
| Winter/Autumn | Clear Sky e.g $I > 500$ | Low $H < 70\%$ | $8\,°C < T_{out} < 15\,°C$ | 23.5 °C |
| Winter/Autumn | Clear Sky e.g $I > 500$ | Low $H < 70\%$ | $T_{out} < 8\,°C$ | 24 °C |
| Winter/Autumn | Cloudy e.g $I < 500$ | High $H > 70\%$ | $T_{out} > 15\,°C$ | 22.5 °C |
| Winter/Autumn | Cloudy e.g $I > 500$ | High $H > 70\%$ | $8\,°C < T_{out} < 15\,°C$ | 23 °C |
| Winter/Autumn | Cloudy e.g $I > 500$ | High $H > 70\%$ | $T_{out} < 8\,°C$ | 23.5 °C |
| Winter/Autumn | Clear Sky e.g $I > 500$ | High $H > 70\%$ | $T_{out} > 15\,°C$ | 22 °C |
| Winter/Autumn | Clear Sky e.g $I > 500$ | High $H > 70\%$ | $8\,°C < T_{out} < 15\,°C$ | 22.5 °C |
| Winter/Autumn | Clear Sky e.g $I > 500$ | High $H > 70\%$ | $T_{out} < 8\,°C$ | 23.5 °C |
| Winter/Autumn | Cloudy e.g $I < 500$ | Low $H < 70\%$ | $T_{out} > 15\,°C$ | 23 °C |
| Winter/Autumn | Cloudy e.g $I > 500$ | Low $H < 70\%$ | $8\,°C < T_{out} < 15\,°C$ | 23.5 °C |
| Winter/Autumn | Cloudy e.g $I > 500$ | Low $H < 70\%$ | $T_{out} < 8\,°C$ | 24 °C |

## 2.3 User comfort satisfaction model

In this iteration, random forest (RF) and neural networks (NN) were used to create the occupant comfort satisfaction models. In the following text, we give a brief explanation of how these methods are applied. The overview of the model are as follows:

We use four different machine learning methods. Two methods solve multi-class classification problems of an occupant's thermal comfort satisfaction ('cool'/'satisfied'/'warm'). The other two methods solve regression problems of an occupant's desirable set point. We used thermal comfort satisfaction and desirable set point as the dependent variables. They inform about how to improve current comfort conditions by describing the occupant's current comfort state and desired ambient temperature. Therefore, they can be used, especially the desired set point, to control the room temperature to improve occupant's comfort satisfaction.

Primarily, as this part was done in advance and user data is not yet available, we used data we received from our artificial agents to train our models. As mentioned before, we have approximately three years of data available from sensors at Nest for the UMAR unit, so our artificial agents gave us feedback for all this data. This is quite enough input to train the models (around 100'000 inputs). Given that we will not have as much input from users in a short period of time, we do a couple of data splitting experiments:

1. Division of data into a 60:40 ratio for the training and test set, respectively
2. Data split so that test set contains a year of consecutive data, which is around 65:35 ratio for the training and test set, respectively
3. Data division so that training set only contains one month of summer and one month of winter, which is around 5% of total data.

The general point of testing these algorithms on artificial data is to see how they behave in different cases and to determine which will perform better.

RF is an ensemble learning method for classification, regression and other tasks that operate by constructing a multitude of decision trees at training time. For classification tasks, the output of the random forest is the class selected by most trees. For regression tasks, the mean or average prediction of the individual trees is returned [9]. We implemented RF using the scikit-learn package in python [10]. With grid search, we selected the best number of estimators, their depth and the size of the feature set considered at each split.

NNs are another learning method for classification and regression. Hence, we use them for both. We implemented ANN using the Keras in python [11]. With a gird search we selected the best number of layers and neurons, activation functions, optimizer, learning rate and the dropout rate for the selected layer.

## 2.4 Control System Overview

The schematic diagram of the control diagram with all the elements is shown in Figure 4. In the following text, the individual elements of the control diagram are described.

The *room temperature model* estimates the current indoor temperature based on the previous indoor temperature, available weather and time information, and the given action. The action is a change in temperature for ΔT (desired temperature change by the artificial agent / human) that is converted to the opening / state of the valve. Weather information includes outside temperature Tout and solar irradiance I. The model with its input and output parameters is represented in Figure 5.

To verify that the DRL agents for controlling temperature setpoints make contributions, primarily in terms of improving occupants' thermal comfort and secondary in energy savings, we will compare obtained OCC policy with baseline methods. The following baseline control methods are implemented.
*Rule based control* - control that uses predefined rules to make the temperature change. These rules will be based on weather conditions, e.g. "IF the outside temperature is below 20°C, THEN the temperature will be set to 24°C". This control tries to keep the temperature within certain limits but does not pay attention to personal comfort and energy savings.

*Control based on occupant comfort satisfaction model* – the models described previously give the comfort satisfaction of the occupant and her/his desired set point as an output. As we have occupant's preferred indoor temperature, these models can also be used for control. These models specify the action to change the temperature as the difference between the current room temperature and the desired temperature of the occupant obtained from the model. This control addresses personal comfort but not energy consumption.

*DRL occupant-centered control policy* - Since our temperature controller should set the appropriate temperature setpoint, which is a continuous space variable, DRL agents will be trained with DRL algorithms that support continuous actions. We use Deep deterministic policy gradient (DDPG) [12] or Proximal policy optimization (PPO) [13] algorithm. The policy network will map the input state into the temperature change ΔT to maximize the reward, which will include occupant comfort satisfaction and energy consumption. The novelty in our approach compared to [14] is that we would have a comfort satisfaction model (see Fig. 4) and use it in simulation to obtain a DRL policy.

New input data that is not considered already in the implemented simulation model will be added to the state of the environment. The output of the occupant comfort satisfaction model will be used in the reward function for the DRL algorithms. This control pays attention to both personal comfort and energy consumption. In addition to this, we also integrate a presence signal into the DRL pipeline. The presence signal is available from a sensor in the Nest offices where we performed the experiment and in which we will experimentally test obtained control policy. When the occupants are present in the room, the comfort model will be used to predict their thermal satisfaction. However, when the room is empty, the policy will focus on keeping the temperature within predefined comfort limits and reducing energy consumption as much as possible.

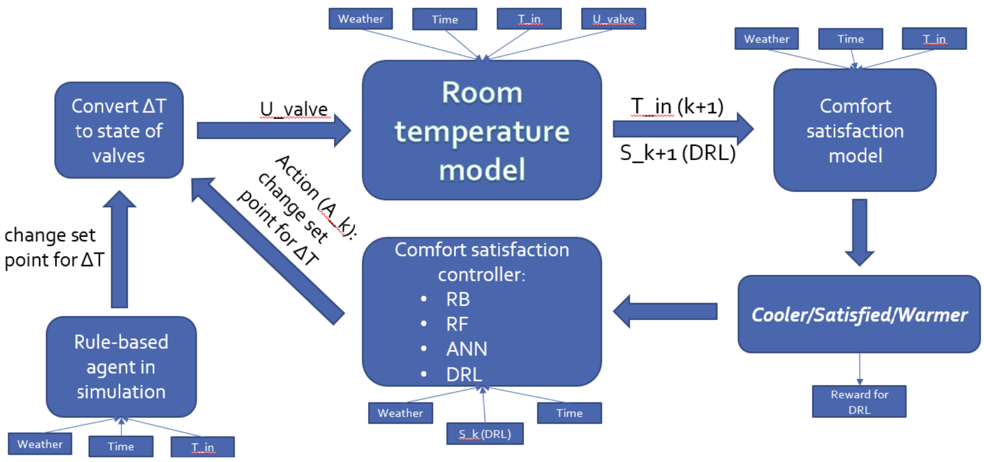



Figure 4: Overview of the control system: comfort satisfaction model, room temperature model, controller, human in the control loop

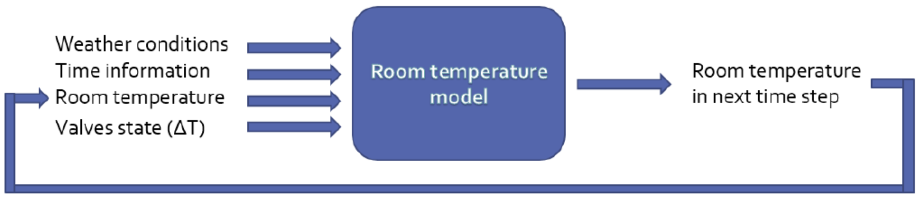



Figure 5: Overview of the room temperature model, measurements used and model output

At the end, since the temperature change is not set in degrees but in the form of valve opening, it is necessary to convert the ΔT into the state of the valve. For that we used a simple bang-bang controller. Also, as the heating season and the cooling season are separate, we must be careful about that. These are the cases for control depending on the season:

1. *Heating season*

   a. When the ΔT is greater than 0.5 °C then the valve is open
   b. When the ΔT is less than 0.5 °C then the valve is closed

2. *Cooling season*

   a. When the ΔT is less than -0.5 °C then the valve is open
   b. When the ΔT is greater than -0.5 °C then the valve is closed.

## 2.5 Guidelines for user survey

The guidelines for the user surveys are based both on the existing literature and the workshop with Mehr als Wohnen.

Personal comfort models for buildings are a relatively young concept. Its development has been spurred by technological advances both with respect to building technology control mechanisms and with respect to user feedback. Hence, there are no standardized procedures yet for how user surveys for personal comfort models should be implemented. There are several metrics to address comfort, namely thermal sensation, thermal acceptability, thermal preference and thermal satisfaction [21]. Also, there are different scales with which these can be implemented [22]. Apart from the question of which metric and which scale to use, one needs to decide how and how often user feedback should be queried. The frequency of surveys should be large enough that the data is sufficient for the implemented comfort model but not so high as to interfere with users' daily tasks [21]. A typical frequency is a few times per day.

The workshop with Mehr als Wohnen showed further aspects that should be considered for the design of user surveys. They fall into the following categories: (i) energy optimization, (ii) facility management and (iii) tenants/users.

Regarding category (i), the goal is to reduce energy consumption using personal comfort models while maintaining or improving individual comfort – which is, of course, the motivation for the whole project. However, this goal is not identical with the expectations of the stakeholders in facility management and, finally, and most importantly, of the users.

In a rental property, facility management (ii) aims at minimizing the interactions with the tenants due to dissatisfaction with thermal comfort. Facility management is the first point of contact when temperatures are not as expected. This can lead to frequent interactions and the need to explain the correct procedures for setting heating setpoints to the tenants. Increasing thermal comfort in an automatic way would reduce these interactions and lead to financial savings in operation as the workload for facility management is decreased. For facility management, energy savings are not the primary goal. From the viewpoint of facility management, it is important that a predictive control system can handle changes of tenants in a convenient way – e.g. by a reset of the control system. Also, a modified approach should be found for shared flats, where the personal rooms are controlled based on a personal model and the shared areas, e.g. living and cooking area, follow an average model.

The main expectation of the tenants (iii) is that thermal comfort is established automatically, without the need for adjustments on their side. Since a predictive control system needs to gather information on building occupancy, privacy is an issue that must be addressed carefully so that behaviour patterns cannot be extracted from the system and potentially be misused.

Even though our starting point in this project is the goal of energy optimization, i.e. category (i), categories (i)-(iii) need to be satisfied in inverse order for the approach to be successful. This is because the tenants and their satisfaction with the system are central. As a second condition, facility managers need to be convinced by an added benefit for their work (e.g. fewer complaints from the tenants about discomfort). Only then can the system help to save operational energy.

For the practical implementation of a predictive control system, the following guidelines can be deduced from the considerations above:
1. The survey questions need to be simple not to burden the users (both cognitively and in terms of their time).
2. Survey frequency should not be more than a few times per day in order not to interfere with users' daily tasks.
3. The user surveys must be minimally invasive, easy to access and easy to answer.
4. Users shall be asked for their feedback on the survey and survey process.

These guidelines will be used for the creation of the user survey prototype. Additionally, Kim et al. [21] described several metrics that have been used in the published research so far. In terms of thermal comfort perception, the following metrics are used: thermal sensation, preference, acceptability, satisfaction. Of the personal factors, there are the physiological ones – skin temperature, heart rate, metabolic rate, clothing insulation, sex, age, body mass index and health status – and the behavioural ones – switching on and off fans or heaters, adjustments of thermostat and opening and closing of windows. In a later review, André and co-workers [22] list similar metrics and additionally also electrodermal activity, electrocardiography, wrist accelerometry, heart rate variability and body surface area. Also, other measures have been proposed in the literature for quantifying perception of comfort more generally, notably theta waves in electro-encephalograms (EEG) [33]. Hence, there is an array of metrics that can be used to parametrize thermal comfort. However, to satisfy the goal of creating a seamless and non-invasive user feedback system, this array needed to be boiled down to a minimal set sufficient as input for the personal comfort model. To do so, we first considered the feedback on thermal comfort perception queried by survey questions and second a set of physiological parameters that can be determined non-invasively.

## 2.6 Survey Questions

Since it is not clear from the literature how many survey questions could be asked to conform with our guidelines on user surveys, we studied how many questions people were ready to answer in the context of meeting rooms. For that, we set up tablets to present “mock-up” user surveys in two identical meeting rooms at NEST at Empa. The surveys were “mock-up” in the sense that it was not the answers that we studied but rather the numbers of complete survey replies given.

Six different conditions were considered, as listed as follows:

1. *3 questions, classic:* Three questions on thermal comfort implemented with LimeSurvey, which has a “classic” look (Figure 6) and provides a lot of features for different question types. The three questions were displayed on one page on a tablet on the main table of the meeting room. A detailed information text was placed next to the tablet.
2. *3 questions, classic, QR:* The same questions were implemented on three separate pages in LimeSurvey. They were available via QR code which was placed with a piece of short information on the main table of the meeting room. From the QR code, participants would be directed to the LimeSurvey website.
3. *3 questions, graphic:* The same questions as above were implemented with SurveyApp, an app that features a more “graphic” look, with little text, large fonts, different colours and even symbolic answers in the form of smileys. The three questions were displayed sequentially, and the tablet was placed on the main table with a short, printed info text.
4. *3 questions, graphic + reward:* As above but participants could take a sweet (small chocolate etc.) as reward for filling out the survey.
5. *1 question, graphic & simple:* Implemented as above with SurveyApp. One simple question without any information text, displayed on a tablet mounted on the wall next to the meeting room exit (Figure 7).
6. *3 questions, graphic & simple:* Implemented as above with SurveyApp. Three sequential simple questions without any information text, displayed on a tablet mounted on the wall next to the meeting room exit.

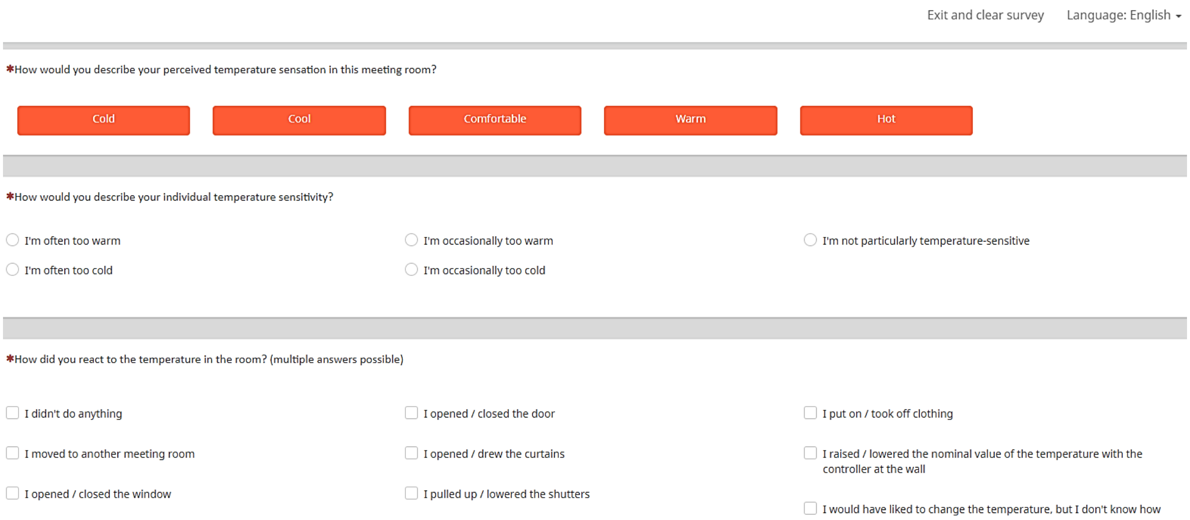



Figure 6: LimeSurvey with three questions: "classic" survey look, as used in condition 1, "3 questions, classic"


Figure 7: Easily accessible survey question, with just one question implemented in the "colourful" SurveyApp as in condition 5, "1 question, graphic & simple"

In the tablet implementation participants could just touch the tablet at the location of the preferred answer and the survey would complete, showing a short thank you note and then returning to the displayed question. The different surveys were displayed over the course of one and a half to five weeks and the number of meetings were recorded from the booking system and put into relation with the complete answers that were received. The number of meetings ranged from 12 to 49 for the different conditions. The results of this short pilot study are summarised in Figure 8. Of course, there is several confounding variables that were not controlled for – such as a number of participants per meeting, sequence effects (maybe some participants in the meetings had seen a previous version of the survey and were biased), change of survey language – the feedback shows a clear trend. Surveys with more than one question, with descriptive texts, small fonts and few colours, had a clearly lower feedback rate than a simple, colourful one-question survey. Interestingly, the QR code version was not used at all.

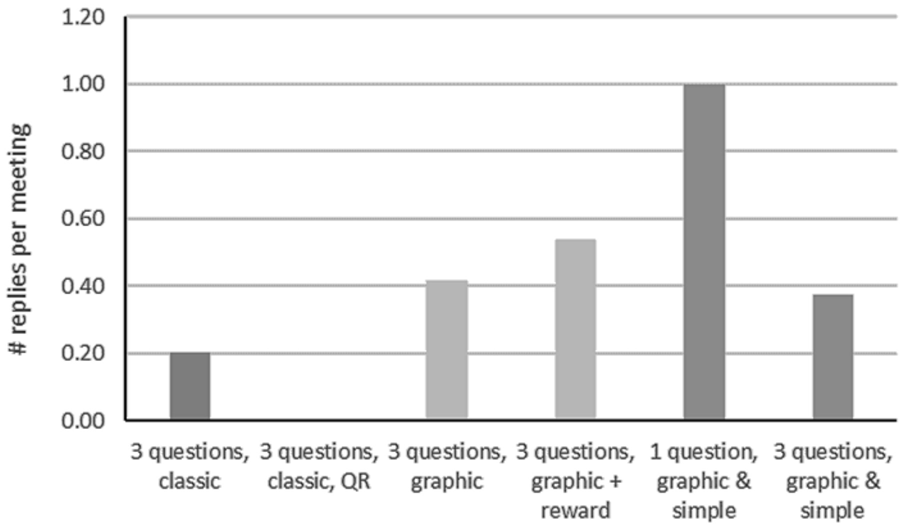



Figure 8: Numbers of replies given for the different types of setups

Our interpretation of this finding is that there is a barrier to answering any survey, which is rather high. For example, the QR code condition would require further action (getting one's phone or computer out, loading the site and only then filling in the survey); hence the lack of any answers is consistent with this explanation. Simple incentives such as chocolate (condition 4) did help to overcome this barrier, as seen in the larger number of answers given in condition 4 compared to condition 3. However, lowering the barrier by making the survey playful and attractive (simple, colourful, placement next to the exit, answer with touch on a virtual button) had a much stronger effect on the number of answers given. This pilot study gives a clear indication of how many survey questions can be asked in the most general case, i.e. without further incentives, namely only one. Also, making it as easy as possible to answer the question, i.e. with a simple clear question and just a finger-press on the tablet, helps a lot with getting more feedback.

Thus, from the possible commonly used metrics – thermal sensation, thermal acceptability, thermal preference and thermal satisfaction – one had to be selected as input for the thermal comfort model. Thermal satisfaction is typically used in the evaluation of buildings during post occupancy evaluations [21]. The thermal sensation is linked to thermal comfort via the assumption that the "neutral" state of sensation is equal to comfort, but different users might understand this differently. Also, for acceptability, the assumption is made that acceptable is comfortable. The closest proxy to ideal conditions, hence, is a thermal preference which is typically implemented via a 3-point scale ("warmer/no change/cooler") where the middle value is considered comfortable, and which can also be linked to users' changing of thermal setpoints [34]. Hence, we selected thermal preference as a metric for thermal comfort.

## 2.7 Selection of physiological variables

On top of the direct survey described above, the potential physiological metrics were screened based on the existing literature and the survey guidelines. We decided against surveying parameters that would result in the creation of subgroups for the personal comfort model, such as clothing insulation, sex, age, body mass index and health status. Given a significantly large population, these would likely make a comfort model more precise. However, in our case, the number of test users is rather limited, so that breaking them down into subgroups would make it impossible to get a significant number of data points for each group.

The metrics of skin temperature and heart rate can be used to assess thermal comfort, in combination with machine learning techniques, as demonstrated in several studies in the last years. Nkurikiyeyezu et al. [25, 26] concluded that Heart rate variability (HRV) can accurately predict thermal comfort with up to 93.7% accuracy, suggesting it could be used in automatic real-time thermal comfort controllers. In a separate study by Yao et al. [27], Skin temperature, ECG ratio, and EEG power were found to be sensitive to ambient temperatures and thermal sensations, suggesting their integration in thermal comfort studies. Based on a review paper from Mansi et al. [28], the conclusion was that the wearable sensors can effectively measure human physiological indices and can potentially be correlated with thermal comfort. Chaudhuri and colleagues [35] could determine thermal state correctly in 87% of cases in a normalized skin temperature model. Choi and Yeom [36] studied skin temperature in several locations and concluded that the back of the wrist resulted in best predictions. Subsequently, Chaudhuri and co-workers [37] improved on their previous results, achieving 97% accuracy for the prediction of thermal state index using machine learning methods and measurement of skin temperature and heart rate. Another study could also show high correlations of skin temperature in combination with heart rate to thermal sensation [38]. Furthermore, heart rate variability can be used to predict thermal comfort with up to 94% accuracy [25,26,39]. The thermal sensation was estimated also using wristband measurements of skin temperature and heart rate in two studies [40], [41]. Unfortunately, the used wristband was not specified in these studies. It is interesting to note that skin temperature can be approximated by analysis of skin colour on the back of the hand, using a visual camera and a machine learning model [42], [43]. While this is currently still in the research state, skin temperature could be measured in the future in a completely non-invasive and easy way.

Based on this analysis, the least invasive way to include physiological parameters in the personal comfort model seems to be the use of a wristband that is worn over the course of the experiment. Other measures, such as electric scalp activity (EEG) are considerably too invasive and too expensive to implement in this project. In NEST, the user can put on a wristband upon arrival in the office and take it off when leaving. A low-cost commercial wristband can be used if the data can be exported in real time. This is not a typical feature for commercial fitness wristbands. However, there are workarounds using API's (application programming interface) of wristband vendors. Fitbit Charge 4 device, that determines both skin temperature and heart rate (among other measurements), was chosen. Through the API the data shall be collected from the wristband and imported into the NEST system. Due to above-average ambient temperatures at the beginning of the measurement period, data collection was prolonged to collect sufficient expressive data.

## 2.8 Guidelines for UC-DPC technical implementation

The UC-DPC (the control algorithm) technical implementation guidelines are based on the workshop with Bouygues Energies & Services InTec Switzerland.

Control of the room temperature is one of the basic control loops in building automation and a very important one. Besides measuring temperature, it would be worthwhile investigating which other variables we could measure in the room and integrate some of them in the control methodology, such as indoor humidity, surface temperatures, and colours in the room. Even though these other variables are important for the occupant comfort, most of them, such as humidity, are not controllable in most of the buildings. Hence, room temperature control remains the basic comfort control problem.

Bouygues is interested in the results of this project as software that they could integrate into their products. In terms of system specifications, Bouygues does not yet have a centralized database. They use different PLCs. Databases remain at customers. They are starting now to work on some cloud-based solutions that would have centralized databases.

# 3 Procedures and methodology

This chapter describes the experimental setup at NEST building's Sprint units, detailing a two-phase approach where Phase 1 involves data collection from two office occupants using surveys, physiological sensors (Fitbit, Greenteg), and environmental measurements to develop SVM-based PTCMs while Phase 2 deploys these PTCMs within a DRL controller framework to validate performance through continued thermal preference monitoring and energy consumption analysis.

## 3.1 Facility Description

The experimental environment at the NEST building at Empa Dübendorf is shown in Figure 9. At Sprint units, the heating/cooling is provided through the water-based radiant panel system on the ceilings. The water flow through the radiant panels in each room is regulated by an electro-actuated valve from Belimo. The indoor room temperature ($T_{in}$) sensors are mounted either on the wall or on the ceiling. The room temperature is controlled by an industrial room temperature controller that receives information from the indoor room temperature sensor and controls the opening of the valve, i.e. the valve state (VS). The information about the outside weather conditions was obtained from a weather station mounted on the roof of the NEST building. The following measurements are available from the weather station: the outside temperature ($T_{out}$), humidity ($H_{out}$), and solar irradiance ($G_h$). The temperature in a room is controlled by setting the setpoint indoor temperature (SP) either through the touch panel on the wall or through a web interface.

Table 4: Measured variables in both phases of UC-DPC

| Variable | Measuring / Querying method | Frequency | Stored in | Phases |
|---|---|---|---|---|
| Thermal preference (TP) | Queried via a custom app via the following question: "What would you prefer at this moment? – Cooler, no change, warmer" | 4 times/day | NEST database | both |
| Heart rate (HR) | Measured with a 1st wristband device | once / min | Locally stored | both |
| Skin temperature (ST) | Measured with a skin sensor on the 2nd wristband device | once / 5 min | Locally stored | both |
| Indoor temperature ($T_{in}$) | Measured with a temperature sensor in the room | once / min | NEST database | both |
| Indoor humidity ($H_{in}$) | Measured with a humidity sensor in the room | once / min | NEST database | both |
| Solar irradiance ($G_h$) | Measured with an outside weather station mounted at NEST | once / min | NEST database | both |
| Indoor temperature setpoint (SP) | Measures within the NEST Building Management System. The SP can be changed either by pressing on a touchscreen in the room or through a web-based application | once / min | NEST database | both |
| Outdoor temperature ($T_{out}$) | Measured with an outside weather station mounted at NEST | once / min | NEST database | both |
| Outdoor humidity ($H_{out}$) | Measured with an outside weather station mounted at NEST | once / min | NEST database | both |
| Valve state (VS) | Measured within the NEST Building Management System | once / min | NEST database | both |

Regarding the building automation part, the Belimo valve, the SP touchscreen and the indoor temperature sensors are connected to the NEST data collection and the Building Management System (BMS). The occupant-centered algorithm was implemented in Python and ran on a local server at NEST, that has access to the BMS of NEST and can change the states of the heating valve. There is less than 1 min delay from when the algorithm in Python requests the change of VS until the change occurs in the valve. The local temperature controller (industrial standard) in the rooms at Sprint units autonomously regulate the room temperature.

Two office users (Empa staff members) consisting of (i) Young male adult present in the sprint unit 2-4 days per week between 9 am to 6 pm and, (ii) Young female adult present in the sprint unit 2-4 days per week between 9 am to 6 pm, participated in the data collection and validation of the study. In all the above cases participation was voluntary and only healthy participants were selected. Figure 10 illustrates the workplace of the test subjects and the location of the various sensors in the Nest Sprint unit. The windows in these units are thermally separated aluminum profiles with triple glazing achieving U-values of about 1.4 W/m²K. The internal wall consisted of either (a) carpet partition walls designed for optimal acoustic insulation or (b) recycled bricks with clay as a binder.

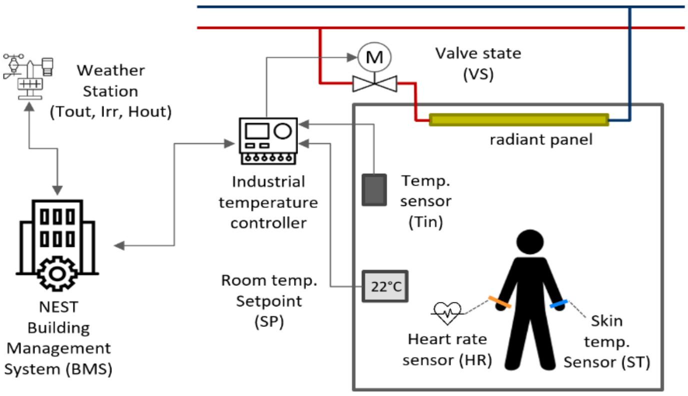



Figure 9: Experimental environment in phase 1 and phase 2 of UC-DPC project

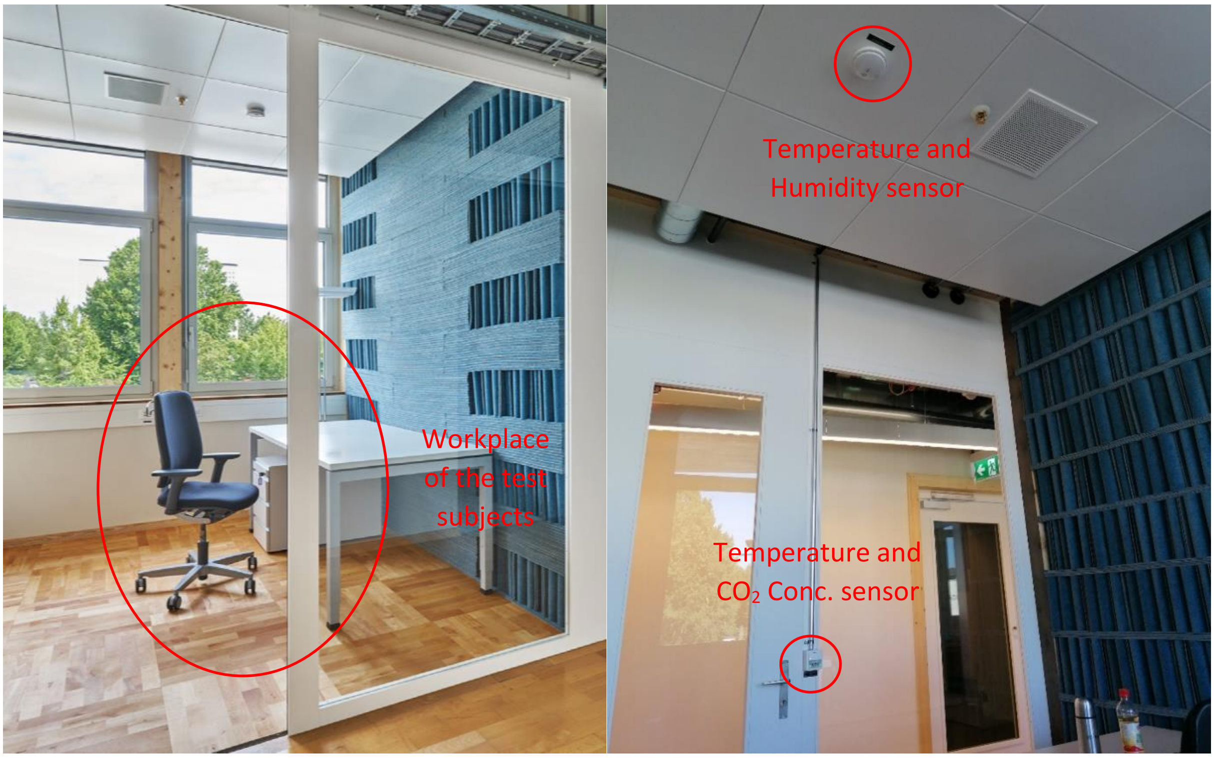



Figure 10: Illustrations from the Sprint office units showing the workplace of the test subjects (left) and, the location of the sensors in the room (right)

## 3.2 Measurement procedure

The study was performed in two single offices (Sprint) in NEST. It is split into two phases: phase 1 and phase 2, both lasting approximately six months; however, they overlap, as can be seen in Table 1. The actual measurement period will be two months for each of the phases.

### Phase 1 – Calibration

Before the start of the measurements the participants are informed about the experimental protocol via personal meeting. The participants are explained about the sensors that they would be equipped with and the daily procedure that should be followed during the measurements, i.e. what the sensors measure, the position where they should be equipped and how to charge the devices in between two measurement sessions. The implementation of the user survey and the recording of the physiological data was prepared. For the survey, the app "SurveyApp" was selected, and the survey was set up. The survey is presented using a tablet on the desk of the office users in Sprint. For the heart rate, the data is recorded via "Fitbit Charge" devices, which allow easy data extraction and provide a sampling of at least once per minute. The skin temperature is recorded using Greenteg "Core" devices. During each measurement session, the test subject arrives in the designated office space and equips them with the Fitbit charge device on the left wrist and the Greenteg core device on the right wrist. The sensors and the equipped positions were chosen to minimize the intrusiveness of the measurements. The tablet is placed on the desk close to the subject and participants are surveyed for their thermal preference (TP) using a custom app displayed on a tablet installed in the room. The tablet displays a visual signal and produces an acoustic cue to alert the participants to provide feedback on their thermal comfort preference. The feedback will be queried four times per day at varied times. Participants will enter their feedback by pressing a button on the tablet. In the final report the results from the data collection during this phase would also be included. Furthermore, HR and ST are recorded at the above-specified frequencies using two devices worn by the participants. Figure 10 shows the signals that go into the standard industrial temperature controller, and Figure 11 depicts the phase 1 calibration pipeline in detail.

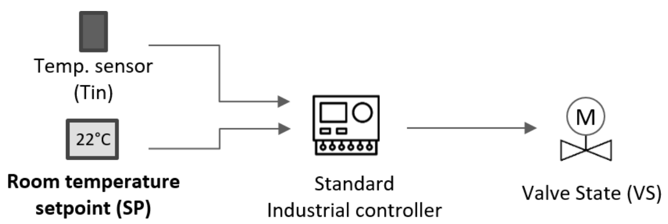



Figure 11: Phase 1 room temperature control setup

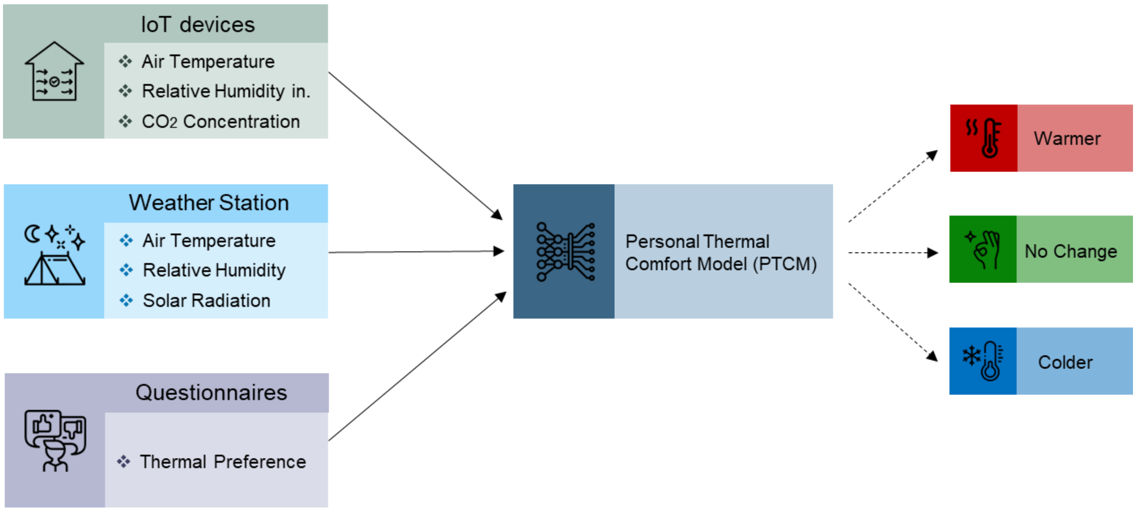


Figure 12: Phase 1 calibration pipeline - from measurements and data to personal thermal comfort model

For the prediction of the TP responses of the participants two different versions were developed. The more accurate version involved data from outdoor environment, indoor environment, physiological measurements and TP responses as gathered in phase 1. The other version that is also represented in Figure 8 excluded the physiological measurements in TP prediction due to the complexities in integrating the real-time data from wearables for the validation of PTCM in phase 2.

To facilitate a comprehensive thermal comfort analysis, PMV values were calculated for each timestep of the input data. These calculations were performed in accordance with the standard methodology outlined in ISO 7730, with several necessary assumptions applied where direct measurements were unavailable. Specifically, the air velocity was set at representative 0.15 m/s, the clothing insulation value was standardized at 0.6 clo (appropriate for typical office attire), and the metabolic activity was assigned a value of 1.5 met (corresponding to light office work). The mean radiant temperature was calculated using the Equation 1 below,

$$MRT = T1F1 + T2F2 + \ldots + TnFn \qquad \text{Equation 1}$$

Where:
T is surface temperature
F is view Factor between person and surface

For the above calculation 6 different surfaces (4 walls, floor and ceiling) and their respective temperatures were considered. The view factors between the 6 surfaces and the person were calculated considering the occupant was in the center of the room at a height of 1.1 m. For the ceiling temperature, the temperature measured from the sensor located in the ceiling was used and for all the other surfaces the temperature recorded from the sensor mounted on the wall was used. Additionally, the indoor environment was categorized based on the European standard EN 16798-1[21], which provides a framework for assessing indoor environmental quality. This standard establishes four categories of environmental quality: Category I (High) for occupants with special needs such as children, elderly, or handicapped individuals; Category II (Medium) representing the normal level typically used for design and operation; Category III (Moderate) which can still provide an acceptable environment despite some risk of reduced occupant performance; and Category IV (Low) which should only be implemented for short durations or in spaces with very limited occupancy periods. This categorization enabled a structured evaluation of the quality of the thermal environment experienced by the participants throughout the study period. The ranges of the indoor environmental parameters in various categories are shown below in Table 5.

Table 5: Indoor environment categorization as per EN 16798-1

| Parameter | Category I | Category II | Category III | Category IV |
|---|---|---|---|---|
| Operative Temp. ($T_{op}$) | 21-23 °C | 20-24 °C | 19-25 °C | 17-26 °C |
| Rel. Humidity (RH) | 30-50 % | 25-60 % | 20-70 % | <20, >70 % |
| Air Speed ($V_{air}$) | <0.1 m/s | <0.16 m/s | <0.21 m/s | >0.21 m/s |
| Vertical Temp. Diff. ($\Delta T_v$) | <2 °C | <3 °C | <4 °C | >4 °C |
| PMV | -0.2–0.2 | -0.5–0.5 | -0.7–0.7 | -1.0–1.0 |
| PPD | <6 % | <10 % | <25 % | >25 % |

For the development of the PTCM, a SVM classifier was implemented with rigorous parameter optimization. The dataset was partitioned into training and testing sets using a 70:30 ratio. To optimize the SVM classifier's performance, a grid search was conducted across multiple hyperparameters. The parameter grid encompassed regularization parameters (C) ranging from 0.01 to 100, gamma values including both 'scale', 'auto', and specific values (0.01, 0.1, 1), four distinct kernel types (linear, polynomial, radial basis function, and sigmoid), and polynomial degrees from 2 to 6 for the polynomial kernel. This exhaustive hyperparameter optimization approach was designed to identify the optimal configuration for each individual subject's thermal comfort prediction model. Following the grid search, the best-performing parameter combination was selected based on cross-validation performance metrics, and this optimized configuration was subsequently employed to train the final SVM-based PTCM. This individualized approach ensures that the models are specifically tailored to capture the unique thermal comfort preferences of each subject. Since all the classes were not equally represented in the collected data, it was crucial in model training to have balanced class weights to avoid modelling bias and overfitting.

### Phase 2 – Validation

In phase 2 of the project the PTCM was deployed in a Deep Reinforcement Learning (DRL) based controller framework. The primary idea behind this phase was to obtain real-life validation of the PTCM for both the participants in the Nest Sprint unit. TP was recorded using the surveys as in phase 1. Based on the created (calibrated) PTCM models, the occupant-centered control strategy was updated using the pipeline as shown in Figures 12 and 14. This DRL-based occupant-centered control updated the VS at each timestep (5 minutes) based on the real time data measurements. This real time data stream included outdoor environmental, indoor environment and TP predictions from PTCM. These automatically generated VS values were the interventions introduced by the developed occupant-centered thermal comfort control algorithm. Occupants found those interventions comfortable or not. If comfortable, they tend not to change (override) the proposed VS. However, if they felt uncomfortable, they could indicate either "cooler" or "warmer" as their TP at that point in time.

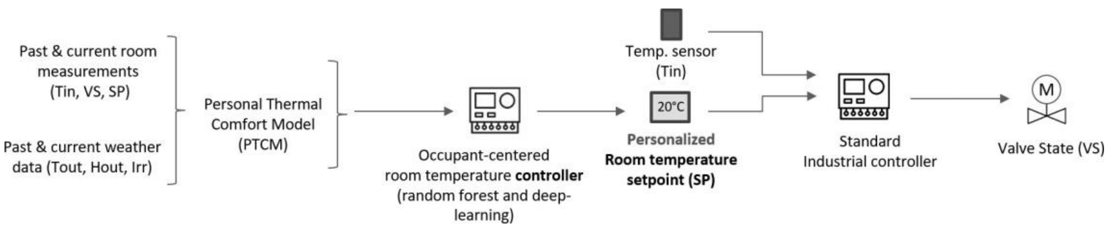



Figure 13: Phase 2 pipeline - From learned personal thermal comfort model to automatically generated room temperature setpoint and comfort validation

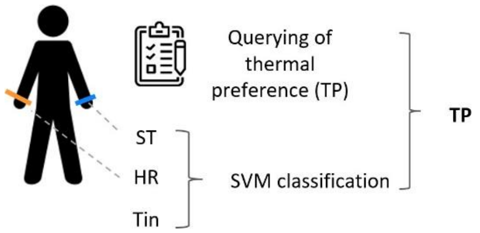


Figure 14: Phase 2 ongoing TP measurement for validation purposes

The occupancy centered control framework incorporating the PTCM has been shown below in Figure 14. As in any typical DRL framework, it consists of the following main elements:

*State:* The state of the Markov Decision Process (MDP) which is the underlying principle of reinforcement learning (RL) was the current indoor and outdoor thermal conditions at each time slot, represented as,

States ($\mathbf{S_t}$) = {$T_{out,t}$, $H_{out,t}$, $G_{h,t}$, $T_{in,t}$, $H_{in,t}$, $C_{co2,t}$, $VS_t$, TP} ………………..... Equation 2

Where,
$T_{out,t}$ is outdoor air temperature at time t,
$H_{out,t}$ is outdoor relative humidity at time t,
$G_{h,t}$ is global horizontal irradiation at time t,
$T_{in,t}$ is indoor air temperature at time t,
$H_{in,t}$ is indoor air relative humidity at time t,
$C_{co2,t}$ is indoor Carbon dioxide concentration at time t,
$VS_t$ is heating valve state in the room at time t,
TP is predicted thermal preference as per the PTCM at time t

*Action:* The action of the MDP is modifying the VS of the hot water supply to the room, represented as,

Action ($A_t$) = {$a_t$} …………………………..………………..... Equation 3

Where,
$a_t$ is the real number representing heating valve state between 0-100% at time t

*Reward*: The reward of the MDP consists of two major and two minor parts as shown in Figure 11. The major reward contributions came from (a) $P_t$ for maximizing timesteps where the predicted TP would be "no change" and (b) $E_t$ for minimizing the heating energy utilization which was calculated as being directly proportional to the percentage of valve opening. The heating energy supplied to the room is directly proportional to the heating valve opening. Higher the valve opening, more heating energy was consumed and vice versa. And the two minor penalty contributions, $V_d$ and $V_s$ were aiming to minimize high amplitude oscillation in valve states and ensure smooth transitions from one to next timesteps. The transition dynamics from the current state to the next state was obtained directly from the Nest BMS which collected all the indoor and outdoor environmental variables every 5 seconds. In addition, the calculated action values (heating valve states) were updated every 5 minutes autonomously for both rooms 183 and 184 in Nest Sprint.

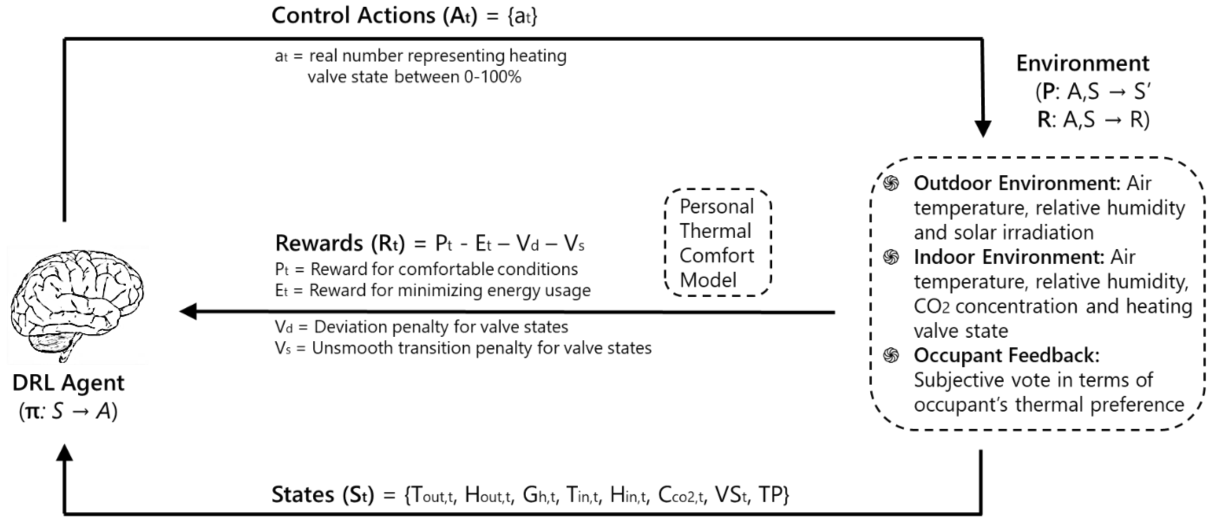


Figure 15: Occupant centric control framework deployed for PTCM validation in phase 2

## 3.3 Data Analysis

### Phase 1 – Calibration

HR and ST data, derivatives of these variables as well as room temperature (Tin) will be used as input for a support vector machine classifier (SVM) for TP (classification into three states: cooler, no change, warmer). The SVM classification will be performed for the first half of the measurement of phase 1, i.e., for a maximum of 80 TP survey feedback. The remaining data of phase 1 (80 TP feedback) will be used to calculate the Pearson correlation coefficient, r, between the TP feedback from the survey and the TP derived from the physiological parameters as a measure of the linear correlation between the two data sets. The Null hypothesis is that the two data sets are unrelated. A permutation test will be used to calculate the p-value for the found correlation coefficient[1]. This means that random permutations from the two data sets will be created, and their Pearson correlation coefficient determined. The position of the original correlation coefficient in the distribution of these artificial correlation coefficients determines the p-value. A significance level of 0.05 is defined as typical. As the permutation test is a nonparametric test, there are no conditions that need to be fulfilled for its application. If the found p-value is smaller than 0.05, HR and ST data will be useable as proxy for TP with the found correlation.

Support Vector Machine (SVM) algorithms have proven to be highly effective tools in the field of thermal comfort classification, as evidenced by multiple studies. For instance, Jiang et al. [29] have demonstrated that the C-SVC variant of SVM can effectively model individual thermal sensations, facilitating the design of personalized conditioning systems that optimize for both energy efficiency and occupant comfort. Similarly, research by Zhou et al. [30] supports the notion that SVM-based thermal comfort models significantly enhance the accuracy of predictions regarding occupants' thermal comfort levels, leading to reductions in predictive errors and a better overall fit. Liu et al. [31] further reinforce the utility of SVM by showing precise predictions of outdoor thermal comfort using parameters like local skin temperature and thermal load, noting particularly high accuracy when it comes to exposed body parts. Additionally, Megri et al. [32] have reported improvements in SVM's predictive ability for thermal comfort indices through the utilization of nonlinear kernels and pertinent experimental factors, conclusively outperforming traditional models such as the Fanger and 2-Node models. Collectively, these studies underscore the superiority of SVM in thermal comfort classification, highlighting its potential for revolutionizing environmental control in built environments.

Given the accuracies >90% reported in the literature [29-32] for the SVM prediction, it is likely to assume a large effect size (Cohen) for the correlation between TP data and HR/ST data (i.e., $r \geq 0.5$). A power analysis (R package "pwr") for a correlation test with the desired power of 0.8 and a significance level of 0.05 yields a minimum sample size of 28.2. Hence, the remaining half of the data of phase 1 for testing the SVM performance will be sufficient.

### Phase 2 – Validation

To determine if the thermal comfort is comparable between the two phases (i.e., standard control vs. PTCM) the distribution of TP replies will be compared with a Kolmogorov-Smirnov test at a p-value of 0.05. There are no special conditions for applying for this test, as it is nonparametric. The Null hypothesis will be that the two variables (TP of phase 1 and TP of phase 2) have the same statistical distribution. If the Null hypothesis is not rejected, the two phases will be considered to provide equally good thermal comfort. The null hypothesis in this case posits that there is no significant difference between the distribution of TP replies from phase 1 (standard control) and phase 2 (PTCM). In other words, it assumes the two phases result in a similar distribution of thermal comfort responses from the subjects. If the null hypothesis is not rejected, it suggests that the thermal comfort experienced by occupants under standard control is statistically like that experienced under the PTCM.

---

[1] The p-value is understood in statistics as the probability of getting results at least as extreme as the one found, given the Null hypothesis is correct.

Finally, a comprehensive analysis was performed comparing the daily heating energy consumption between phase 1 and phase 2 for both rooms, utilizing Heating Degree Days (HDD) as a standardized metric with a baseline temperature of 16 °C. The sprint unit is characterized by the installation of Belimo sensors that measure the heating and cooling valve positions and flow rates (with accuracy of ±2%) to each room. Also, these sensors measure the aggregate heating and cooling energy used by the Sprint unit for space conditioning. To accurately determine the heating energy usage, the total heating energy consumption was distributed to each of the 12 rooms based on their respective valve position. The equation below illustrates the calculation of heating energy consumption at every timestep in room 183,

$$Q_{heat,183} = \frac{vs_{183} * \dot{m}_{183}}{\sum vs_{i,sprint} * \dot{m}_{i,sprint}} * Q_{heat,tot} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad \text{Equation 4}$$

Where,

$Q_{heat,183}$ is the heating energy supplied by the heating panels to room 183

$vs_{183}$ is the instantaneous valve position in the heating panels in room 183

$\dot{m}_{183}$ is the maximum design flow rate of hot water to room 183

$vs_{i,sprint}$ is the instantaneous valve position in the heating panels in room i of Sprint unit

$\dot{m}_{i,sprint}$ is the maximum design flow rate of hot water to room i of Sprint unit

$Q_{heat,tot}$ is the total heating energy supplied to all the rooms in the Sprint unit

# 4 Results and discussion

This chapter presents the key findings showing that SVM-based PTCMs achieved 89% and 83% prediction accuracy for the two participants respectively, with the deployed PTCM-DRL controller successfully increasing occupant satisfaction by 18-19% (more "no change" thermal preference votes) while reducing heating energy consumption, particularly at higher Heating Degree Days, demonstrating the effectiveness of personalized thermal comfort control in balancing occupant needs with energy efficiency.

## 4.1 Phase 1 - Calibration

This section presents the findings from our personalized thermal comfort modelling approach using SVM classifiers in phase 1 of the project. The results are organized to examine the environmental and physiological conditions experienced by the participants along with an analysis of their thermal preference votes in relation to these conditions. The performance of the SVM-based PTCMs is then evaluated, comparing their predictive accuracy against traditional PMV-based approaches. Figure 16 presents a comparative analysis of environmental and physiological parameters across three TP categories (-1: colder, 0: no change, 1: warmer) for two subjects (ID 1 and ID 2). For Subject ID 1, outdoor temperature exhibits a notable correlation with TP, with median temperatures of 10.3°C, 9.4°C, and 8.5°C for preferences of colder, no change, and warmer, respectively. Outdoor relative humidity also demonstrates a good relationship with TP, with median values of approximately 58%, 66% and 86% respectively across categories. In terms of indoor temperature, there is almost no difference in its median value for no change and warmer TP. ID 1 exhibits slightly elevated heart rates (median ~ 72bpm) when preferring cooler conditions compared to neutral and warmer preference (median ~62 bpm).

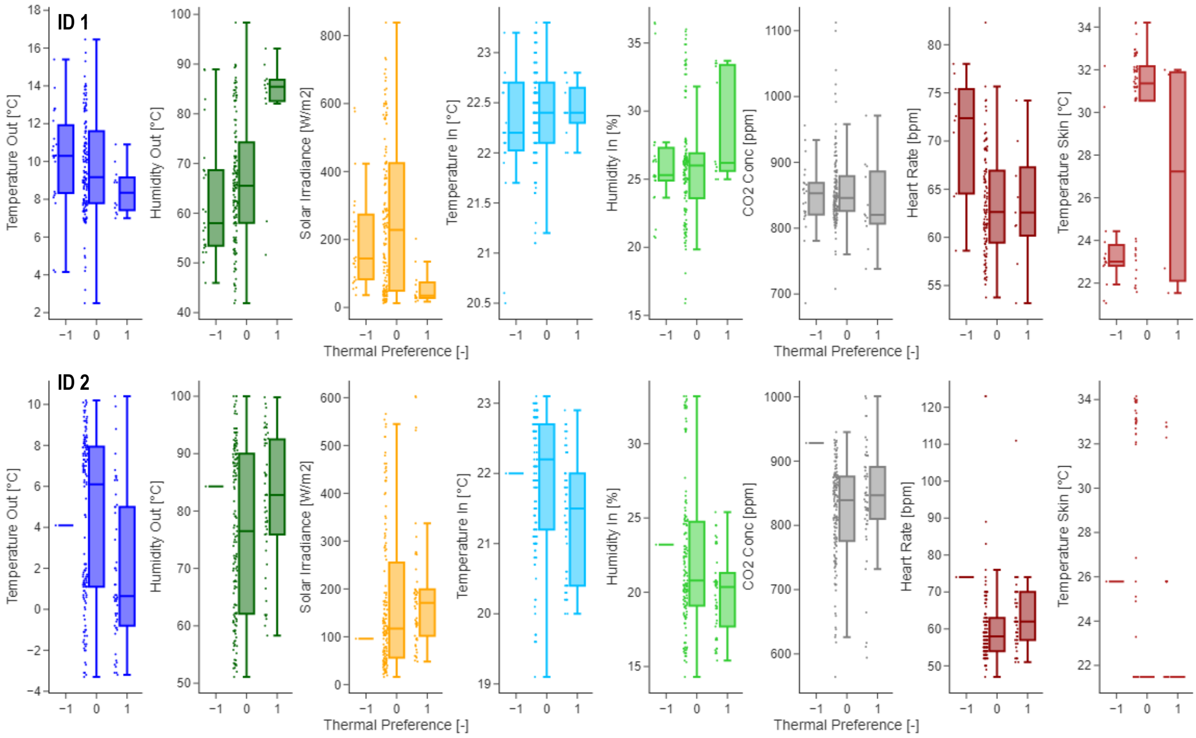


Figure 16: Comparative analysis of environmental and physiological parameters across TP categories in phase 1

For Subject ID 2, both the indoor and outdoor air temperature shows notable correlation with TP. Among all the data points captured, there was only one cooler TP vote expressed by ID2 despite the indoor temperature generally being warmer than ID1. The median temperatures for no change and warmer TP votes are 6.1°C and 0.6 °C for the outdoors and 22.2°C and 21.5°C for the outdoors respectively. Quite like ID1, in case of ID2 as well the outdoor humidity demonstrates a good relationship with TP, with median values of approximately 77% and 83% respectively across the latter two categories. In general, outdoor environmental parameters, particularly temperature and humidity demonstrate consistent relationships with thermal preferences. In terms of the correlation with the indoor environmental and physiological variables the observed inter-subject variations in TP votes, underscore the importance of personalized thermal comfort modelling approaches rather than population-based standards.

Figure 17 represents the TP indicated by both the participants with respect to the indoor environmental conditions (air temperature vs humidity ratio). The coloured zones (green, yellow and orange) represent the indoor environment categorization as per Table 5. Here again, it was interesting to note the variability in TP responses in quite similar indoor air temperature and humidity conditions. ID1 is satisfied for about 80% of the timesteps although for about 55% of the experimental time the environmental conditions lie in Category 3 and 4. For ID2, almost for 96% of the experimental time the environmental conditions are classified to be worse than Category 2. As expected, the subject was not always satisfied with the conditions, and the preference was for even warmer conditions. It is however interesting to note that for both subjects the no change TP is quite uniformly distributed across the indoor air temperature variation range of 20 to 23 °C. In case of ID2, the warmer TP response increased significantly when the indoor relative humidity was below 25% which was not the case for ID1.

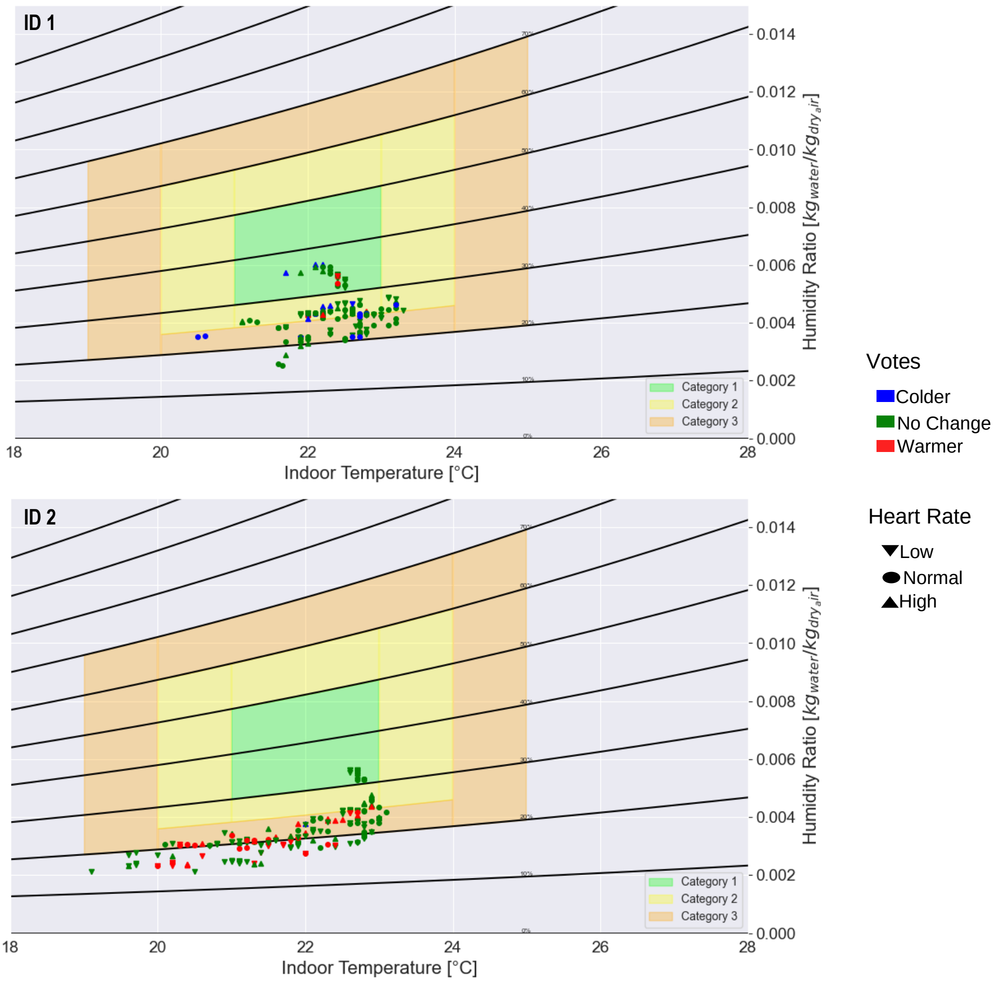



Figure 17: TP vs indoor environmental conditions for participants, ID1 and ID2 (top to bottom)

For the TP predictions using the above-mentioned environmental and physiological variables, SVM classifier was used. Figure 18 shows the confusion matrix for the SVM classifier in TP testing for both the participants. The SVM classifier achieved a high accuracy of 89% and 83% in predicting the TP responses for ID1 and ID2 respectively. This high accuracy suggests that the model is highly effective at capturing the relationship between the input variables (Tin, RHin, Tout and RHout, HR) and the TP classes. It underscores the strength of using the combination of environmental and physiological variables and SVM classifiers for predicting TP responses. However, it is important to remind the fact that class distribution of the data used for training and testing was not well distributed. This could lead to modelling bias and overfitting towards the classes with higher representation. Indeed, the SVM classifier has better accuracy in predicting "no change" votes compared to "warmer" and "cooler" votes for both participants. For ID1, the prediction accuracies for the "cooler", "no change" and "warmer" classes were 67%, 97% and 0% respectively. In case of ID2 the accuracies were 95% and 47% for "no change" and "warmer" votes. These numbers demonstrate overfitting to "no change" TP class for the above PTCMs. It calls for a more controlled indoor environment conditions for better representation of different TP classes in data collection. The analysis of PMV values in relation to the TP votes reveals significant individual variations that underscore the necessity for personalized thermal comfort models.

Figure 19 shows the variation in PMV values with the different categories of TP votes for both the participants. For Subject ID 1, the median PMV value corresponding to "no change" TP votes was -0.06, indicating neutral conditions according to PMV calculations. In case of Subject ID 2, the median PMV value corresponding to "no change" TP votes was -0.16. Although the PMV value was slightly lower in case of ID2 it still indicated almost neutral sensation as per the PMV scale. The median PMV value did not change much in case of ID1 for warmer TP, whereas it was at a slightly lower value of -0.32 in case of ID2. There was a clear correlation between the PMV values and the TP responses for the subject ID2. However, it could not be said the same for ID1. The above analysis illustrates the personalized nature of thermal perception in human subjects. Even with limited number of participants, there was quite some contrast in the correlations between TP responses and PMV of indoor conditions.

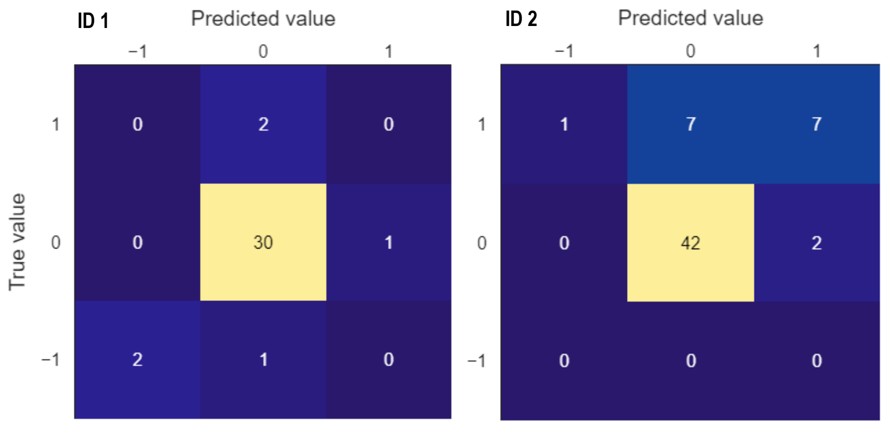



Figure 18: SVM Classifier performance in TP testing for ID1 and ID2

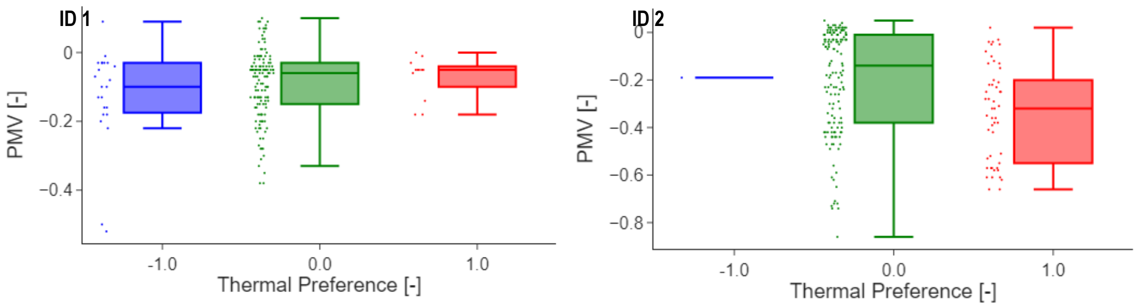



Figure 19: PMV vs TP for ID1 and ID2 (top to bottom) in phase 1

## 4.2 Phase 2 - Validation

This section presents the findings from our personalized thermal comfort modelling approach using SVM classifiers in the phase 2 of the project. The Figure 20 illustrates the variation in the environmental and physiological parameters across different thermal preference categories in the validation phase of the UC-DPC project. Across both individuals, there are clear associations between thermal preference and environmental variables. As seen in phase 1, TP votes are inversely and directly correlated with the variations in outdoor temperature and relative humidity respectively. Also, both the subjects exhibited preferences for warmer conditions when the indoor temperature conditions were lower and vice versa. The range of indoor air temperatures for ID1 and ID2 were quite similar, approximately ranging between 21.5-24 °C for both participants. The median values for no change TP for ID1 and ID2 was observed at indoor air temperatures slightly below and above 23 °C respectively.

Figure 21 below represents the TP indicated by both the participants with respect to the indoor environmental conditions in phase 2 of the project. As seen in the previous plots, the indoor environmental conditions are quite similar for the two subjects in this phase with slightly cooler exposure for ID2. The conditions for ID1 mostly lied between categories 1-3 with 5%, 28%, 61% and 9% of the timesteps in categories 1 to 4 respectively. Similarly, for ID2, the indoor conditions lied mostly in categories 2 and 3, with 33%, 46% and 21% of the timesteps in categories 2 to 4 respectively. In case of ID1, the subject was mostly satisfied for about 98% timestep, which is an improvement by about 18% as compared to phase 1. Also, ID2 was mostly satisfied for about 93% timestep showing an increment by 19% in no change TP votes. Comparing Figures 17 and 21, it can be observed that the indoor air temperature ranges were much more comfortably lying between 21.5-24 °C in phase 2 as compared to phase 1 especially in case of ID2.

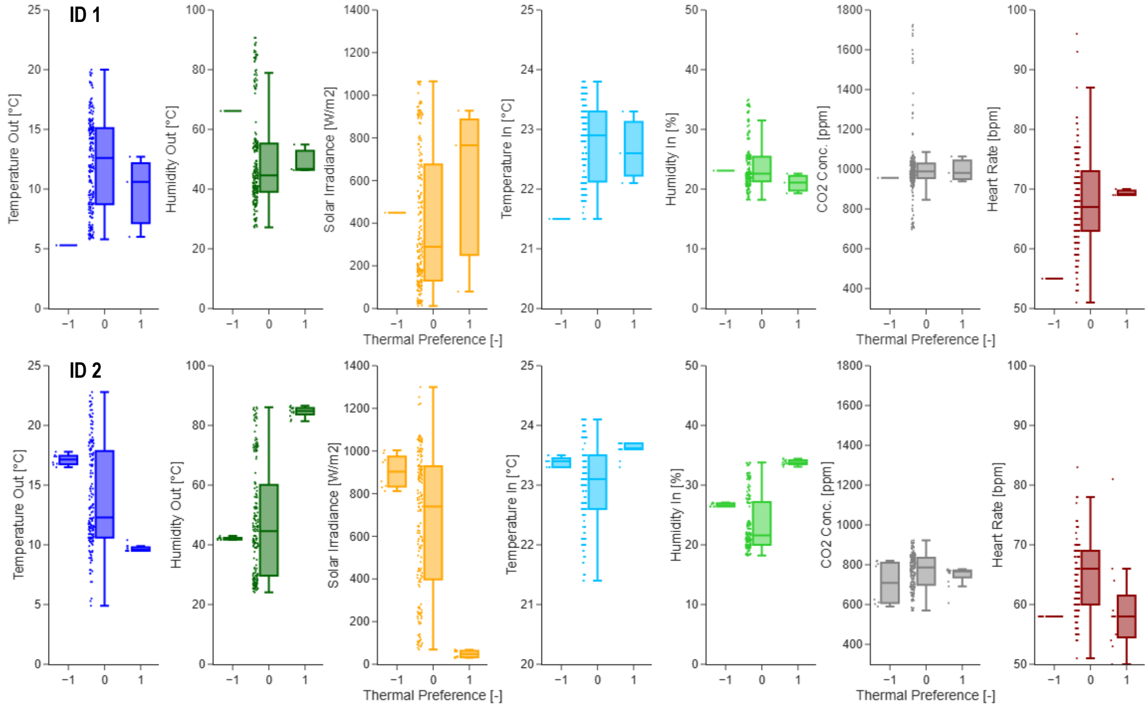



Figure 20: Comparative analysis of environmental and physiological parameters across thermal preference categories in phase 2

The comparison of calculated PMV values with the recorded TP responses in phase 2 is shown in Figure 22. PMV is a widely used thermal comfort index that predicts the mean thermal sensation of a group of occupants on a seven-point scale, with 0 representing thermal neutrality. For ID 1 (left), the PMV values fall in the range between -0.3 to 0.2, with the median PMV value for no change TP responses being 0. When this person expressed a preference for cooler conditions (TP = -1.0), their PMV values cluster tightly around -0.35. However, this inference cannot be generalized given very few "cooler" TP responses. For neutral preference, there was a much wider distribution of PMV values (-0.3 to 0.2), suggesting greater variability in their thermal sensation during "no change" votes. Similarly, for ID2 the range of PMV variations was between -0.3 and 0.3 with median PMV value of 0.03 for no change TP response. Comparing performances with phase 1, PTCM-DRL controller was effective in maintain the comfortable conditions with median PMV of 0 for ID1 in phase 2. For ID2 it reduced the PMV values of the exposed indoor conditions by 0.13 on average towards neutral exposure in phase 2.

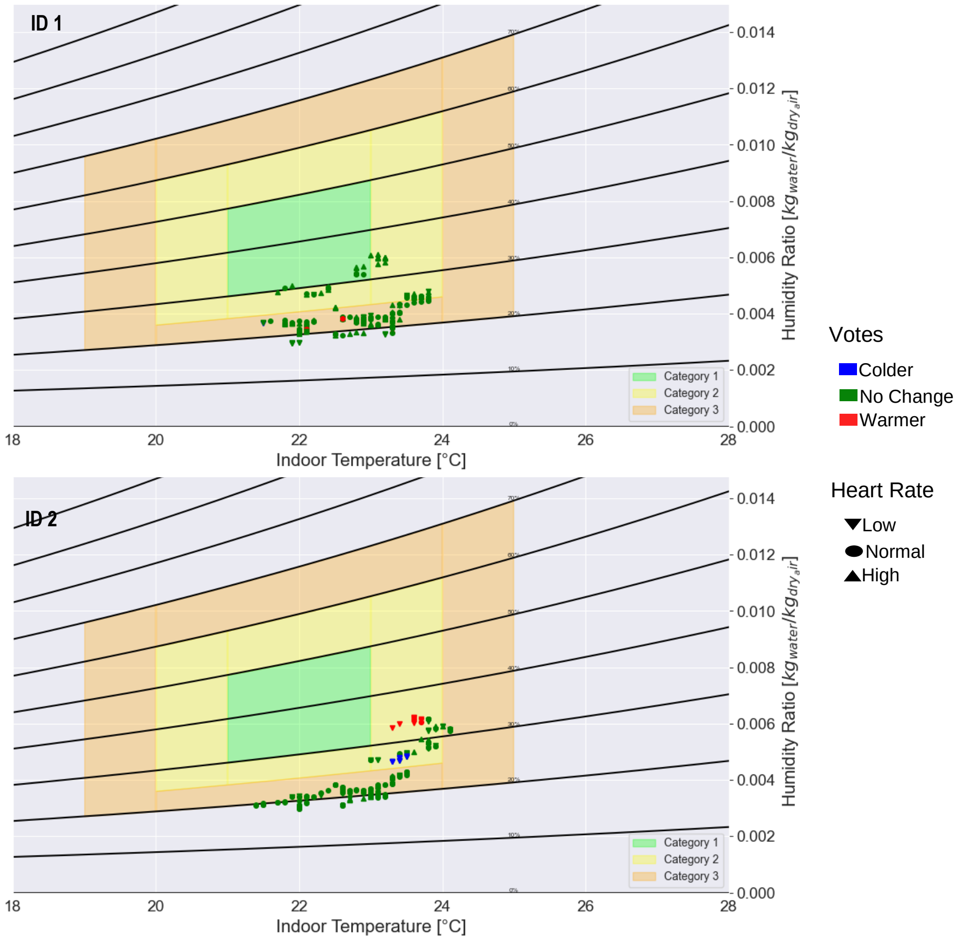



Figure 21: TP vs indoor environmental conditions for participants ID1 and ID2 in phase 2

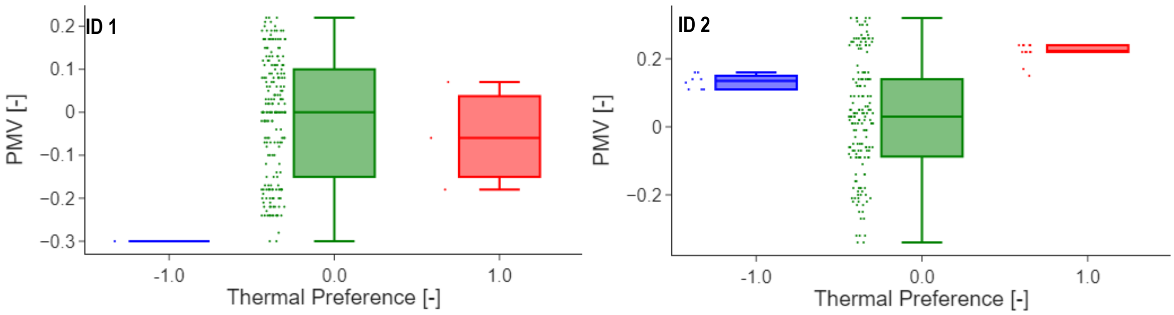



Figure 22: PMV vs TP for ID1 and ID2 (top to bottom) in phase 2

Figure 23 illustrates the variations in outdoor temperature, indoor temperature and heating valve states in the rooms 184 and 183 (occupied by ID1 and ID2 respectively) during 4 selected days in the phase 2. The period chosen for this analysis was between Mar 18-22, 2025, when the PTCM-embedded DRL controller was already deployed in both the rooms. The temperature variations reflect the data shown in Figures 19 and 20, with the indoor temperatures in both the rooms varying in the range between 21.5-24 °C. The heating valve states variations in both the rooms illustrate the adaptation of the controller towards personal preference. In the phase 1, it was seen that ID1 was quite satisfied with the indoor conditions irrespective of it lying in Categories 3 and 4. Thus in the phase 2, the controller minimized the heating energy usage by modulating the valve states accordingly. The heating valve opening percentage in the room 184 was consistently lower by as much as 10% as compared to the valve states in room 183. There were some timesteps where the valve opening in room 183 attained values lower than that in room 184. That was mainly to prevent the overshoot of the indoor air temperature beyond the comfortable temperature ranges and minimize heating energy expenditure. In the phase 2, the TP votes where the ID2 indicated "no change" as their TP increased by about 19% as compared to 74% "no change" votes during the phase 1. The valve states in both the rooms also exhibit an inverse correlation with the outdoor temperature variations to optimize personalized comfort and heating energy usage. These results demonstrate the effectiveness of the PTCM-embedded DRL controller in adapting to individual thermal preferences and external environmental conditions, leading to significant improvements in occupant satisfaction while optimizing heating energy usage. This personalized approach challenges conventional standardized comfort categories by showing that individual preference can vary substantially from prescribed norms. Furthermore, the system's ability to balance comfort and energy consumption suggests potential for broader applications in building management systems, where personalized thermal control could simultaneously enhance occupant well-being and contribute to energy conservation goals.

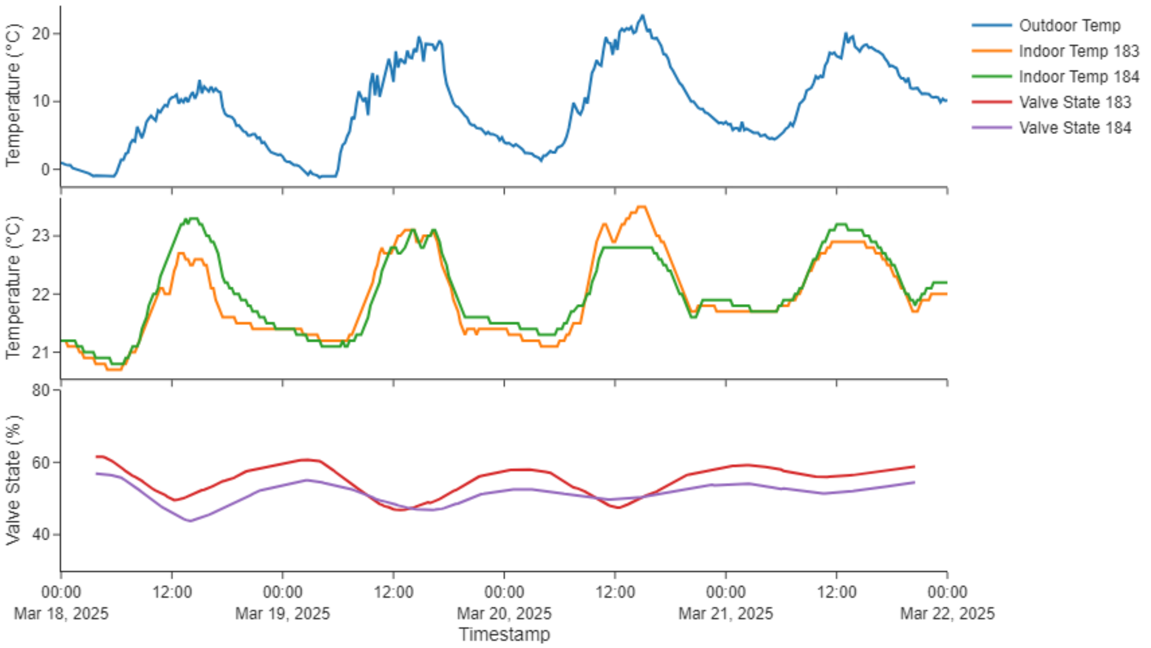



Figure 23: Variations in outdoor temperature, indoor temperature and heating valve states (top to bottom) during 4 selected days in the phase 2

The final analysis consisted of the comparison of the Heating Degree Days (HDD) with the heating energy usage in both the phases. HDD serve as a critical metric for quantifying the impact of weather conditions on building heating requirements, representing the cumulative temperature difference below a specified base temperature over a given period. The analysis of heating energy consumption against HDD provides a powerful method for normalizing energy usage across varying weather conditions, enabling fair comparisons between different time periods or before and after energy efficiency interventions. Figure 24 for ID1 (Room 184) demonstrates a clear relationship between daily heating energy consumption HDD across both phases. In Phase 1 (before controller deployment), the daily heating energy consumption ranged from approximately 2.3 to 7 kWh as HDD increases from 0 to 15 Kd. After the PTCM-embedded DRL controller deployment (Phase 2), energy consumption was slightly lower at minor HDD values (0-2 Kd) and significantly lower at higher HDD values (8-15 Kd), ranging from around 2.1 to 6 kWh. The Phase 2 trendline was slightly flatter (slope of 0.26 kWh/Kd in phase 2 compared to 0.31 kWh/Kd in phase 1), indicating a consistent energy reduction especially in colder outdoor environmental conditions.

For ID2 (Room 183), the impact of the controller was even more pronounced as shown in Figure 25. During Phase 1, daily heating energy consumption ranged from about 3.6 to 5.8 kWh. Following controller deployment in Phase 2, consumption dramatically decreased to between 1.8 to 4.8 kWh. The vertical separation between Phase 1 and Phase 2 trendlines is considerably higher than observed in Room 184. The controller maintained the positive correlation between HDD and energy use while reducing consumption by up to 1.8 kWh per day across the range of outdoor conditions. The PTCM-embedded DRL controller has successfully reduced heating energy consumption in both rooms, but with notably different magnitudes. While the room 184 achieved the highest heating energy use reduction of between 1 kWh at higher HDD values of 15 Kd, the highest heating energy use reduction in case of room 183 of about 1.8 kWh was at lower HDD values. These results demonstrate that personalized thermal comfort controllers can achieve substantial energy savings while respecting individual preferences, with the magnitude of savings varying according to room dimensions, external conditions and occupant-specific factors.

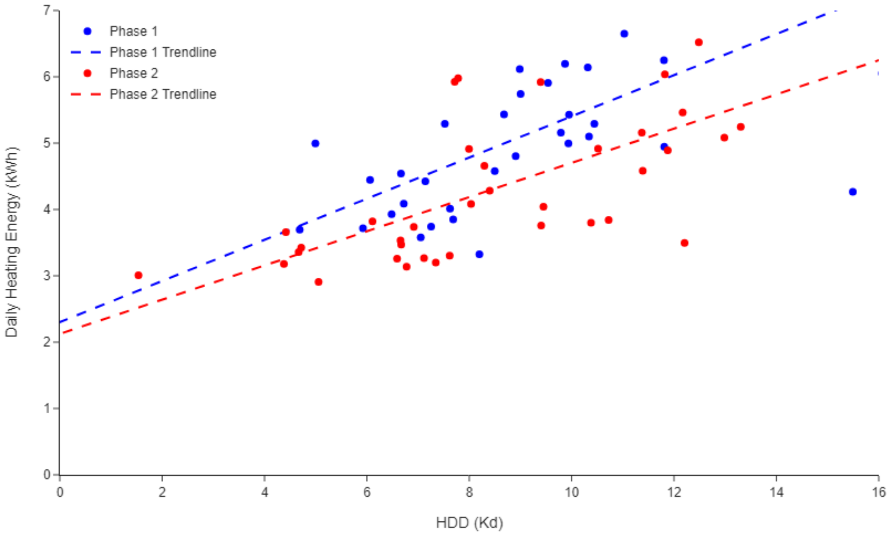



Figure 24: HDD vs heating energy usage in both phases for ID1 (Room 184)

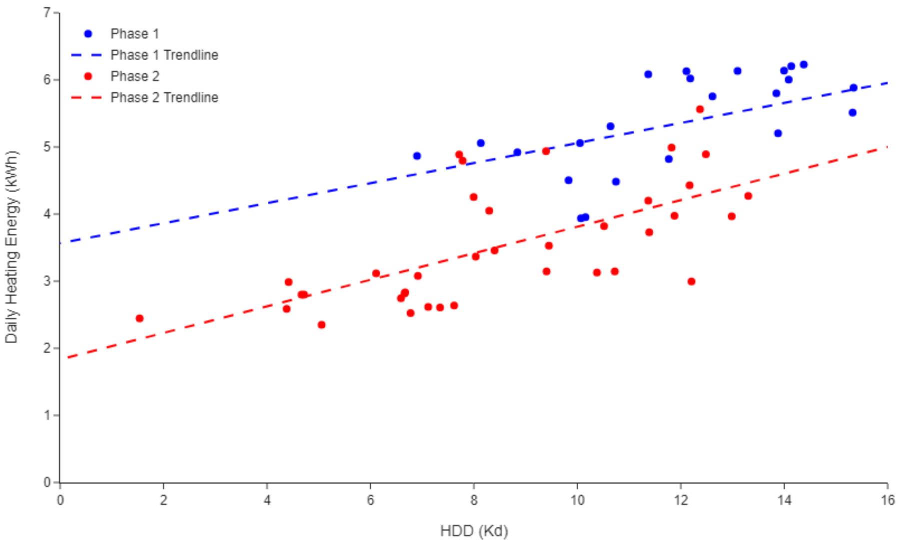



Figure 25: HDD vs heating energy usage in both phases for ID2 (Room 183)

# 5 Conclusions and outlook

This chapter summarizes the project's success in developing personalized thermal comfort control systems that significantly improved occupant satisfaction and reduced energy consumption, while acknowledging limitations including small sample size (two participants), biased model performance toward majority classes, and lack of diverse environmental conditions, recommending future research with larger populations, more varied conditions, and systematic tracking of behavioural adaptations to enhance the generalizability and robustness of personalized thermal comfort systems.

The UC-DPC project was aimed at developing new control algorithms that leverage user data on thermal comfort needs to provide user-specific comfort throughout the day automatically. The project consisted of three main parts: i) decision about what data (from which sensors and/or which survey questions) to take into consideration to determine user needs for comfort, ii) developing a comfort satisfaction model, and iii) developing the control algorithm and testing it in simulation and the real environment at the NEST building at Empa in Dübendorf. To this objective the project was divided into two primary phases. Phase 1 consisted of data collection from two single office units (Rooms 183 and 184) from Nest Sprint unit. During the data collection which took place in the winter months, the spaces were occupied by two test subjects and TP votes were gathered periodically using Survey app deployed in desk-mounted tablet. In addition, environmental data consisting of outdoor variables (air temperature, relative humidity and solar irradiation) and indoor variables (air temperature, relative humidity, $CO_2$ concentration and heating valve states) and physiological data (heart rate variability and wrist skin temperature) were recorded. Using the data gathered, SVM-based PTCMs were developed for both the subjects. These PTCMs provided 89% and 83% accurate TP predictions for ID1 and ID2 respectively. However, the performance was slightly biased given better performance in majority classes ("no change" TP votes) and worse in minority classes ("colder" and "warmer" TP votes). It was largely due to the limited variations in the indoor environmental conditions in the Nest sprint units, resulting in under-exploration of TP votes at temperatures ranging between 18-22 °C.

The developed PTCM was then embedded in a DRL-based controller framework modulating the heating valve states independently in each of the rooms with the objective to provide personalized thermal comfort and minimize heating energy usage. Thus, phase 2 of the UC-DPC project involved deploying the PTCM embedded DRL-controller framework in the Nest Sprint unit and validating its performance through periodic collection of TP votes as in phase 1. The findings of phase 2 are listed below:

1) The indoor environmental conditions resulting from PTCM-DRL controller resulted in 18% and 19% increase in "no change" TP votes from ID1 and ID2 respectively.
2) With reference to standard EN 16798-1, in phase 2 the indoor environmental conditions improved significantly in the case of rooms 183 and 184, with the conditions almost always lying in Categories 1-3 in phase 2 of the UC-DPC project.
3) The controller ensured personalization in the variation of the heating valve states for ID1 and ID2 (Room 184 and Room 183 respectively) ensuring consistently more valve opening in room 183 as compared to room 184 while maintaining similar indoor air temperature range in both rooms.
4) It was not only effective in lowering the baseline energy requirements (in both rooms) but also it ensured reduced heating energy usage at higher HDD values (in both rooms).

This study faced several methodological constraints that warrant acknowledgment. First, the limited variations in indoor environmental conditions across both experimental rooms resulted in PTCMs that exhibited performance bias, demonstrating higher accuracy for majority TP classes while underperforming in minority classes. Additionally, the research design did not account for significant confounding variables related to occupant behavioral adaptations, including clothing adjustments, window operation behaviors, varying physical activity levels prior to data collection periods, and metabolic variations due to food intake. Perhaps most significantly, the restricted sample size comprising only two participants (one male, one female) substantially limits the generalizability of findings to broader populations. Future research should address these limitations by incorporating more diverse indoor environmental conditions to train more robust PTCMs capable of accurately predicting responses across the full spectrum of thermal preferences. Expanded studies should implement systematic tracking of behavioral adaptations, potentially through wearable sensors or structured self-reporting protocols, to better isolate the causal relationships between environmental parameters and comfort perceptions. Furthermore, larger-scale deployments across diverse demographic groups, building types, and climatic regions would enhance external validity while enabling more nuanced analyses of individual differences in thermal preferences. Integration of these enhanced methodological approaches would substantially strengthen the evidence base for personalized thermal comfort systems while potentially revealing important interactions between personal, behavioral, and environmental factors that were not captured in the present investigation.

# 6 Publications and other communications

The following work-in-progress documents have been prepared using this report's scientific outputs partially documented,

- Dodig, A., Svetozarevic, B., Nagy, Z. and Heer, P. "Random Forests and Deep Reinforcement Learning based Occupant Centered control with Personal comfort models", *Building and Environment*, to be submitted (by end of May 2025).

- Chatterjee, A., Svetozarevic, B. and Heer, P. "Personalized Thermal Comfort Control via SVM-based Models and Deep Reinforcement Learning: A Field Study in Office Environments", *Building and Environment*, currently in preparation (to be submitted by end of October 2025).

The following conference paper has been prepared and submitted based on the data from UC-DPC project,

- Chatterjee, A., Heer, P. "Data-driven personalized thermal comfort model for office buildings in Switzerland", CISBAT Conference 2025.

The methodology developed and shallow insight into phase 1 and phase 2 analysis of the UC-DPC project was also presented in the following events,

- Chatterjee, A., Heer, P. "UC-DPC Project: Personalization in Thermal Comfort" was presented at SWEET Lantern Site-Visit on 16 April 2025.

- Methodology and initial findings of the UC-DPC Project was discussed at IEA Annex 95 (Human-Centric Buildings for a Changing Climate) meeting in Sevilla, Spain on 18-19 November 2024.

The materials generated from the UC-DPC project were used in the development of slides for "Lecture 3: Evolution of thermal comfort models" for the course master's in integrated building systems at ETH Zurich.

# 7 References

[1] B. W. Olesen and K. C. Parsons, “Introduction to thermal comfort standards and to the proposed new version of EN ISO 7730,” Energy and Buildings, vol. 34, no. 6, pp. 537–548, Jul. 2002, doi: 10.1016/S0378-7788(02)00004-X.
[2] R. J. de Dear and G. S. Brager, “Thermal comfort in naturally ventilated buildings: revisions to ASHRAE Standard 55,” Energy and Buildings, vol. 34, no. 6, pp. 549–561, Jul. 2002, doi: 10.1016/S0378-7788(02)00005-1.
[3] N. E. Klepeis et al., “The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants,” J Expo Sci Environ Epidemiol, vol. 11, no. 3, pp. 231–252, Jul. 2001, doi: 10.1038/sj.jea.7500165.
[4] D. Ormandy and V. Ezratty, “Health and thermal comfort: From WHO guidance to housing strategies,” Energy Policy, vol. 49, pp. 116–121, Oct. 2012, doi: 10.1016/j.enpol.2011.09.003.
[5] L. Lan, P. Wargocki, and Z. Lian, “Quantitative measurement of productivity loss due to thermal discomfort,” Energy and Buildings, vol. 43, no. 5, pp. 1057–1062, May 2011, doi: 10.1016/j.enbuild.2010.09.001.
[6] T. Akimoto, S. Tanabe, T. Yanai, and M. Sasaki, “Thermal comfort and productivity - Evaluation of workplace environment in a task conditioned office,” Building and Environment, vol. 45, no. 1, pp. 45–50, Jan. 2010, doi: 10.1016/j.buildenv.2009.06.022.
[7] L. Lan, Z. Lian, and L. Pan, “The effects of air temperature on office workers’ well-being, workload and productivity-evaluated with subjective ratings,” Applied Ergonomics, vol. 42, no. 1, pp. 29–36, Dec. 2010, doi: 10.1016/j.apergo.2010.04.003.
[8] O. Lucon et al., Buildings. Working Group III contribution to the IPCC 5th assessment report, climate change 2014: mitigation of climate change. Cambridge, UK/New York, NY, USA: Cambridge University Press, 2014.
[9] J. Y. Park et al., “A critical review of field implementations of occupant-centric building controls,” Building and Environment, vol. 165, p. 106351, Nov. 2019, doi: 10.1016/j.buildenv.2019.106351.
[10] T. I. Salsbury, “A SURVEY OF CONTROL TECHNOLOGIES IN THE BUILDING AUTOMATION INDUSTRY,” IFAC Proceedings Volumes, vol. 38, no. 1, pp. 90–100, Jan. 2005, doi: 10.3182/20050703-6-CZ-1902.01397.
[11] R. J. de Dear and G. S. Brager, “Developing an Adaptive Model of Thermal Comfort and Preference,” p. 19.
[12] G. Serale, M. Fiorentini, A. Capozzoli, D. Bernardini, and A. Bemporad, “Model Predictive Control (MPC) for Enhancing Building and HVAC System Energy Efficiency: Problem Formulation, Applications and Opportunities,” Energies, vol. 11, no. 3, p. 631, Mar. 2018, doi: 10.3390/en11030631.
[13] E. T. Maddalena, Y. Lian, and C. N. Jones, “Data-driven methods for building control — A review and promising future directions,” Control Engineering Practice, vol. 95, p. 104211, Feb. 2020, doi: 10.1016/j.conengprac.2019.104211.
[14] B. Svetozarevic, C. Baumann, S. Muntwiler, L. Di Natale, M. N. Zeilinger, and P. Heer, “Data-driven control of room temperature and bidirectional EV charging using deep reinforcement learning: Simulations and experiments,” Applied Energy, p. 118127, Nov. 2021, doi: 10.1016/j.apenergy.2021.118127.
[15] L. Di Natale, B. Svetozarevic, P. Heer, and C. N. Jones, “Deep Reinforcement Learning for room temperature control: a black-box pipeline from data to policies,” J. Phys.: Conf. Ser., vol. 2042, no. 1, p. 012004, Nov. 2021, doi: 10.1088/1742-6596/2042/1/012004.
[16] T. Wei, Y. Wang, and Q. Zhu, “Deep Reinforcement Learning for Building HVAC Control,” in Proceedings of the 54th Annual Design Automation Conference 2017, Austin TX USA, Jun. 2017, pp. 1–6. doi: 10.1145/3061639.3062224.
[17] R. Jia, M. Jin, K. Sun, T. Hong, and C. Spanos, “Advanced Building Control via Deep Reinforcement Learning,” Energy Procedia, vol. 158, pp. 6158–6163, Feb. 2019, doi: 10.1016/j.egypro.2019.01.494.
[18] “DFAB HOUSE – Digital Fabrication and Living, Empa, Duebendorf, Switzerland.” https://www.empa.ch/web/nest/digital-fabrication (accessed Oct. 17, 2019).
[19] “Urban Mining & Recycling unit, Empa, Duebendorf, Switzerland.” https://www.empa.ch/web/nest/urban-mining (accessed Oct. 17, 2019).
[20] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[21] J. Kim, S. Schiavon, and G. Brager, “Personal comfort models – A new paradigm in thermal comfort for occupant-centric environmental control,” Building and Environment, vol. 132, pp. 114–124, Mar. 2018, doi: 10.1016/j.buildenv.2018.01.023.
[22] M. André, R. De Vecchi, and R. Lamberts, “User-centered environmental control: a review of current findings on personal conditioning systems and personal comfort models,” Energy and Buildings, vol. 222, p. 110011, Sep. 2020, doi: 10.1016/j.enbuild.2020.110011.
[23] T. K. Ho, “Random decision forests,” in Proceedings of 3rd International Conference on Document Analysis and Recognition, Aug. 1995, vol. 1, pp. 278–282 vol.1. doi: 10.1109/ICDAR.1995.598994.
[24] J. A. Anderson, An Introduction to Neural Networks. MIT Press, 1995.
[25] Nkurikiyeyezu, K., Suzuki, Y., & Lopez, G. (2018). Heart rate variability as a predictive biomarker of thermal comfort. Journal of Ambient Intelligence and Humanized Computing, 9, 1465-1477. https://doi.org/10.1007/s12652-017-0567-4.
[26] Nkurikiyeyezu, K., Suzuki, Y., Tobe, Y., Lopez, G., & Itao, K. (2017). Heart rate variability as an indicator of thermal comfort state. 2017 56th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE), 1510-1512. https://doi.org/10.23919/SICE.2017.8105506.
[27] Yao, Y., Lian, Z., Liu, W., & Shen, Q. (2008). Experimental study on physiological responses and thermal comfort under various ambient temperatures. Physiology & Behavior, 93, 310-321. https://doi.org/10.1016/j.physbeh.2007.09.012.
[28] Mansi, S., Barone, G., Forzano, C., Pigliautile, I., Ferrara, M., Pisello, A., & Arnesano, M. (2021). Measuring human physiological indices for thermal comfort assessment through wearable devices: A review. Measurement, 183, 109872. https://doi.org/10.1016/J.MEASUREMENT.2021.109872.
[29] Jiang, L., & Yao, R. (2016). Modelling personal thermal sensations using C-Support Vector Classification (C-SVC) algorithm. Building and Environment, 99, 98-106. https://doi.org/10.1016/J.BUILDENV.2016.01.022.
[30] Zhou, X., Xu, L., Zhang, J., Niu, B., Luo, M., Zhou, G., & Zhang, X. (2020). Data-driven thermal comfort model via support vector machine algorithms: Insights from ASHRAE RP-884 database. Energy and Buildings, 211, 109795. https://doi.org/10.1016/j.enbuild.2020.109795.
[31] Liu, K., Nie, T., Liu, W., Liu, Y., & Lai, D. (2020). A machine learning approach to predict outdoor thermal comfort using local skin temperatures. Sustainable Cities and Society, 59, 102216. https://doi.org/10.1016/j.scs.2020.102216.
[32] Megri, A., & Naqa, I. (2016). Prediction of the thermal comfort indices using improved support vector machine classifiers and nonlinear kernel functions. Indoor and Built Environment, 25, 16 - 6. https://doi.org/10.1177/1420326X14539693.
[33] H. Guan, S. Hu, M. Lu, M. He, X. Zhang, and G. Liu, “Analysis of human electroencephalogram features in different indoor environments,” Building and Environment, vol. 186, p. 107328, Dec. 2020, doi: 10.1016/j.buildenv.2020.107328.
[34] J. Kim, Y. Zhou, S. Schiavon, P. Raftery, and G. Brager, “Personal comfort models: Predicting individuals’ thermal preference using occupant heating and cooling behavior and machine learning,” Building and Environment, vol. 129, pp. 96–106, Feb. 2018, doi: 10.1016/j.buildenv.2017.12.011.
[35] T. Chaudhuri, D. Zhai, Y. C. Soh, H. Li, and L. Xie, “Thermal comfort prediction using normalized skin temperature in a uniform-built environment,” Energy and Buildings, vol. 159, pp. 426–440, Jan. 2018, doi: 10.1016/j.enbuild.2017.10.098.
[36] J.-H. Choi and D. Yeom, “Study of data-driven thermal sensation prediction model as a function of local body skin temperatures in a built environment,” Building and Environment, vol. 121, pp. 130–147, Aug. 2017, doi: 10.1016/j.buildenv.2017.05.004.
[37] T. Chaudhuri, Y. C. Soh, H. Li, and L. Xie, “Machine learning driven personal comfort prediction by wearable sensing of pulse rate and skin temperature,” Building and Environment, vol. 170, p. 106615, Mar. 2020, doi: 10.1016/j.buildenv.2019.106615.
[38] W. Li, J. Zhang, T. Zhao, and R. Liang, “Experimental research of online monitoring and evaluation method of human thermal sensation in different active states based on wristband device,” Energy and Buildings, vol. 173, pp. 613–622, Aug. 2018, doi: 10.1016/j.enbuild.2018.05.056.
[39] W. Liu, Z. Lian, and Y. Liu, “Heart rate variability at different thermal comfort levels,” Eur J Appl Physiol, vol. 103, no. 3, pp. 361–366, Jun. 2008, doi: 10.1007/s00421-008-0718-6.

[40] D. Li, C. C. Menassa, and V. R. Kamat, “Non-intrusive interpretation of human thermal comfort through analysis of facial infrared thermography,” Energy and Buildings, vol. 176, pp. 246–261, Oct. 2018, doi: 10.1016/j.enbuild.2018.07.025.
[41] W. Li, J. Zhang, T. Zhao, and J. Ren, “Experimental study of an indoor temperature fuzzy control method for thermal comfort and energy saving using wristband device,” Building and Environment, vol. 187, p. 107432, Jan. 2021, doi: 10.1016/j.buildenv.2020.107432.
[42] X. Cheng, B. Yang, T. Olofsson, G. Liu, and H. Li, “A pilot study of online non-invasive measuring technology based on video magnification to determine skin temperature,” Building and Environment, vol. 121, pp. 1–10, Aug. 2017, doi: 10.1016/j.buildenv.2017.05.021.
[43] X. Cheng et al., “A Contactless Measuring Method of Skin Temperature based on the Skin Sensitivity Index and Deep Learning,” Applied Sciences, vol. 9, no. 7, Art. no. 7, Jan. 2019, doi: 10.3390/app9071375.
[44] A. Chatterjee and D. Khovalyg. "Dynamic indoor thermal environment using reinforcement learning-based controls: Opportunities and challenges." Building and Environment (2023). https://doi.org/10.1016/j.buildenv.2023.110766.

# 8 Appendix

## 8.1 Heating degree solar air days (HDSD)

In addition to the HDD, Heating degree solar air days (HDSD) was also calculated. HDSD offer an advantage over traditional HDD by incorporating the effects of solar gains on building heating demand. While HDD considers only outdoor air temperature, HDSD provides a more accurate reflection of a building's true heating needs by accounting for passive solar heat contributions. The equation below illustrates the calculation of HDSD,

$$HDSD = T_{out} + \frac{G_h * \propto}{h_o} - T_{base} \quad \text{..............................} \quad \text{Equation 5}$$

Where,

$T_{out}$ is the outdoor air temperature

$G_h$ is the horizontal global irradiation

$\propto$ is the solar absorptivity of the clear triple glazing unit in Sprint (≈0.15)

$h_o$ is the external surface heat transfer coefficient (= 23 as prescribed by EN ISO 6946, ISO 13790)

$T_{base}$ is the typical baseline temperature for heating in Switzerland

The figures below illustrate the variation of daily heating energy usage with respect to variations in HDSD for both rooms. Figure 26 for ID1 (Room 184) demonstrates a clear relationship between daily heating energy consumption and HDSD across both phases. In Phase 1, the daily heating energy consumption ranged from approximately 2.3 to 7.4 kWh as HDSD increases from 0 to 16 Kd. After the PTCM-embedded DRL controller deployment (Phase 2), energy consumption was slightly lower at minor HDSD values (0-2 Kd) and significantly lower at higher HDD values (8-16 Kd), ranging from around 2.4 to 6.5 kWh. The Phase 2 trendline was slightly flatter, indicating a consistent energy reduction especially in colder outdoor environmental conditions.

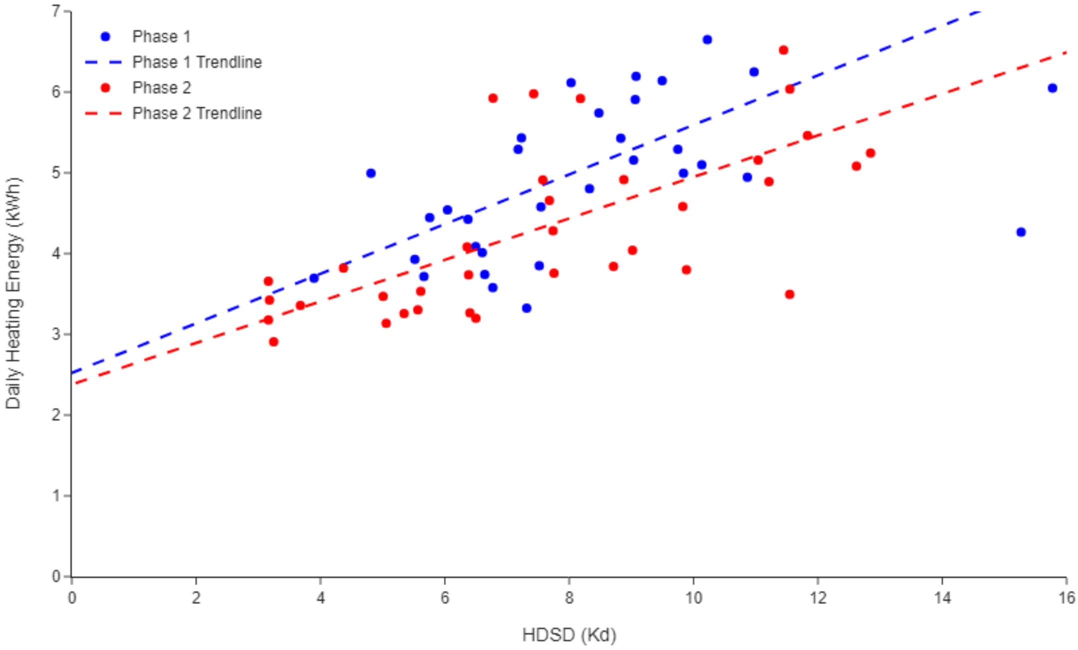



Figure 26: HDSD vs daily heating energy usage in both phases for ID1 (Room 184)

For ID2 (Room 183), the impact of the controller was even more pronounced as shown in Figure 27. During Phase 1, daily heating energy consumption ranged from about 3.7 to 6 kWh. Following controller deployment in Phase 2, consumption dramatically decreased to between 2 to 5.2 kWh. The vertical separation between Phase 1 and Phase 2 trendlines is considerably higher than observed in Room 184. The controller maintained the positive correlation between HDSD and energy use while reducing consumption by up to 1.7 kWh per day across the range of outdoor conditions. While the room 184 achieved the highest heating energy use reduction of between 0.9 kWh at higher HDD values of 16 Kd, the highest heating energy use reduction in case of room 183 of about 1.7 kWh was at lower HDD values. The results again demonstrate that personalized thermal comfort controllers can achieve substantial energy savings while respecting individual preferences, with the magnitude of savings varying according to room dimensions, external conditions and occupant-specific factors.

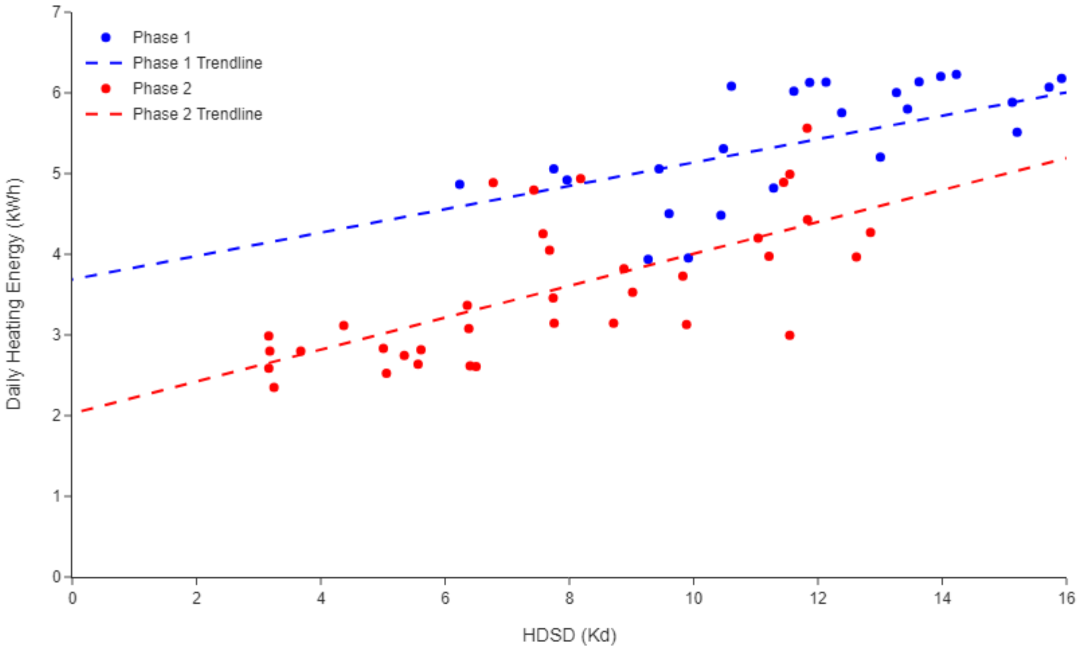



Figure 27: HDSD vs daily heating energy usage in both phases for ID2 (Room 183)

## 8.2 Error in heating energy calculation

The equation 4 in the chapter 3 (Procedures and methodology) provided the mathematical formulation for calculating the instantaneous heating energy used by room 183. However, the equation assumed that the difference in temperature of the hot water between the inlet and outlet to the room 183 and the whole Sprint unit was the same. The assumption was employed in this study given that hot water temperature measurements are only available for the whole Sprint unit and not for each room. The equation below illustrates the calculation of heating energy consumption at every timestep in room 183 without the above assumption,

$$Q_{heat,183} = \frac{vs_{183} * \dot{m}_{183} * \Delta T_{183}}{\sum(vs_{i,sprint} * \dot{m}_{i,sprint}) * \Delta T_{sprint}} * Q_{heat,tot} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad \text{Equation 6}$$

Where,

$Q_{heat,183}$ is the heating energy supplied by the heating panels to room 183
$vs_{183}$ is the instantaneous valve position in the heating panels in room 183
$\dot{m}_{183}$ is the maximum design flow rate of hot water to room 183
$\Delta T_{183}$ is the temperature difference between the inlet and outlet hot water in room 183
$vs_{i,sprint}$ is the instantaneous valve position in the heating panels in room i of Sprint unit
$\dot{m}_{i,sprint}$ is the maximum design flow rate of hot water to room i of Sprint unit
$\Delta T_{sprint}$ is the temperature difference between the inlet and outlet hot water in Sprint unit
$Q_{heat,tot}$ is the total heating energy supplied to all the rooms in the Sprint unit

Figure 28 shows the variation in the difference between inlet and outlet hot water to room 183 in Phase 2. Mean temperature difference between the inlet and outlet water temperature was about 4 °C. Figure 29 shows the uncertainty in the heating energy usage with deviation in outlet hot water temperature from room 183. The blue line shows the calculated energy usage based on equation 4 ranging. The chart includes three error bands representing different levels of measurement uncertainty: ±0.25 K (green), ±0.5 K (orange), and ±1 K (red), with the red band showing the largest potential variation in energy estimates. The error bands demonstrate that the energy consumption estimates can vary by 20-35% if the outlet water temperature from room 183 and the whole sprint unit differ by as much as ±1 K.

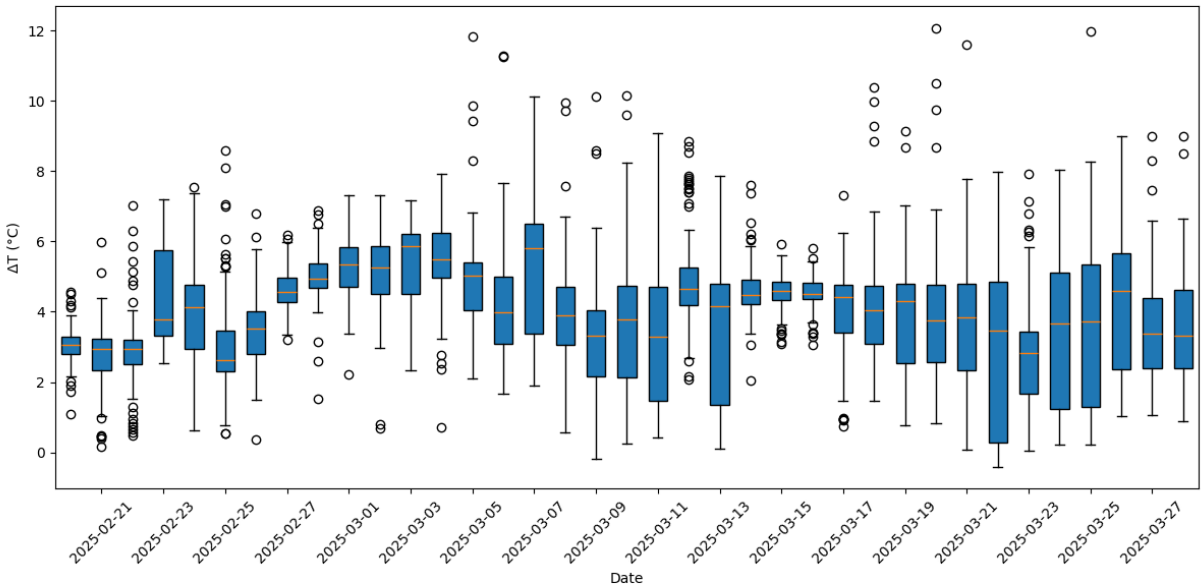



Figure 28: Variation in the difference between inlet and outlet hot water in Sprint unit in Phase 2

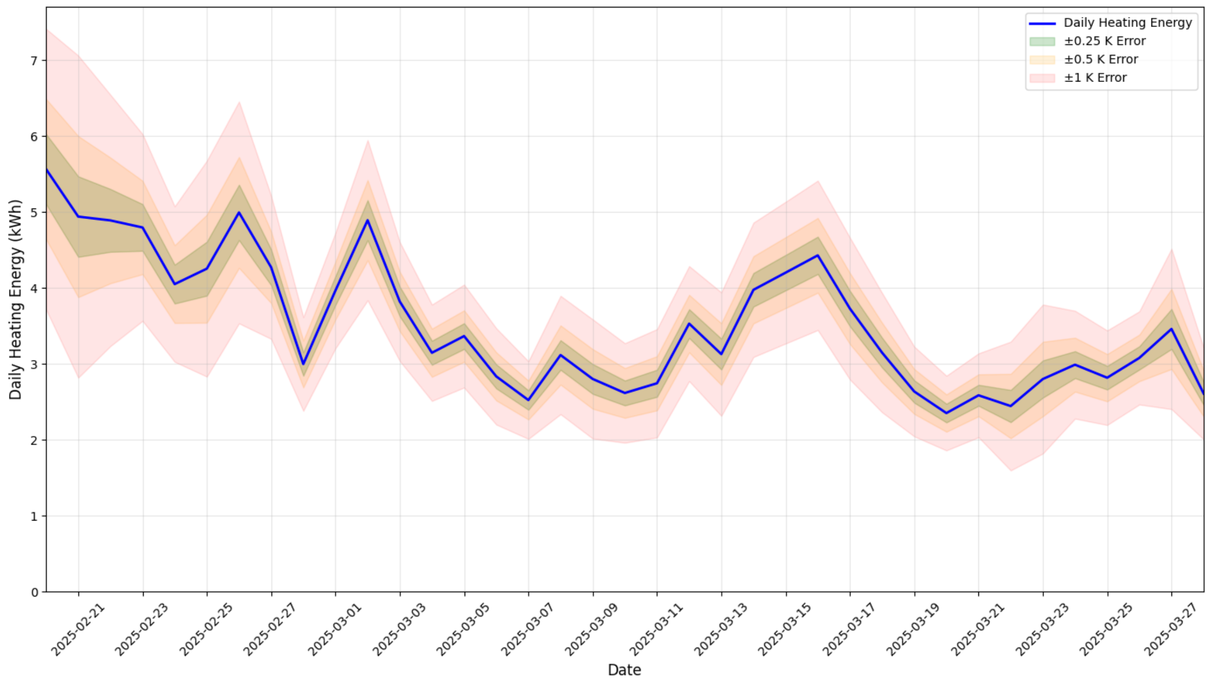



Figure 29: Uncertainty in the heating energy usage with deviation in outlet hot water temperature from room 183