



**HAL**  
open science

# Co-zyBench: using co-simulation and digital twins to benchmark thermal comfort provision in smart buildings

Jun Ma, Dimitrije Panic, Roberto Yus, Georgios Bouloukakis

## ► To cite this version:

Jun Ma, Dimitrije Panic, Roberto Yus, Georgios Bouloukakis. Co-zyBench: using co-simulation and digital twins to benchmark thermal comfort provision in smart buildings. 22nd International Conference on Pervasive Computing and Communications (PerCom), IEEE, Mar 2024, Biarritz, France. hal-04514006

**HAL Id: hal-04514006**

**<https://inria.hal.science/hal-04514006>**

Submitted on 20 Mar 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Co-zyBench: Using Co-Simulation and Digital Twins to Benchmark Thermal Comfort Provision in Smart Buildings

Jun Ma<sup>†</sup>, Dimitrije Panic<sup>†</sup>, Roberto Yus<sup>\*</sup>, Georgios Bouloukakis<sup>†</sup>

{jun\_ma, dimitri.panic, georgios.bouloukakis}@telecom-sudparis.eu, ryus@umbc.edu

<sup>†</sup>Télécom SudParis, Institut Polytechnique de Paris, France

<sup>\*</sup>Dept. of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, USA

**Abstract**—Heating, Ventilation, and Air Conditioning (HVAC) systems account for 40% to 50% of energy usage in commercial buildings. Thus, innovative ways to control and manage HVAC systems while preserving occupants’ comfort are required. State-of-the-art solutions employ pervasive systems with sensors or smart devices to gauge individual thermal sensations, yet assessing these methods is challenging. Real-world experiments are expensive, limited in access, and often overlook occupant and regional diversity. To address this, we introduce Co-zyBench, a benchmark tool using a Digital Twin (DT) approach for evaluating personalized thermal comfort systems. It employs a co-simulation middleware interfacing between a DT of the smart building and its HVAC system and another DT representing occupants’ dynamic thermal preferences in various spaces. The DTs that support Co-zyBench are generated based on information, including data captured by sensors, of the space in which the thermal comfort system has to be evaluated. Co-zyBench incorporates metrics for energy consumption, thermal comfort, and occupant equality. It also features reference DTs based on standard buildings, HVAC systems, and occupants with diverse thermal preferences.

**Index Terms**—personalized thermal comfort, benchmarking, energy consumption, digital twin

## I. INTRODUCTION

Efficient energy management in buildings is a pressing concern due to Heating, Ventilation, and Air Conditioning (HVAC) systems contributing nearly half of the total energy consumption in commercial buildings [1]. By 2050, cooling demand is projected to triple driven by global warming and urbanization [2]. While energy conservation is vital, maintaining thermal comfort for occupants is equally important. Research shows a direct link between the indoor thermal environment and the cognitive functions and productivity of occupants [3]. However, achieving thermal comfort is challenging because it varies among individuals due to several factors [4] including physiological (such as age, sex and fitness); psychological (including effects of personal control, thermo-specific self-efficacy and personality); and the built environment’s contextual factors [5].

The increasing focus on both occupant comfort and energy efficiency has highlighted the importance of advanced thermal comfort systems capable of adapting to individual sensations. Leveraging pervasive computing and the Internet of Things

(IoT), recent research has demonstrated the feasibility of modeling and responding to individual thermal preferences [6], [7]. Particularly, wearable devices such as wrist-worn temperature sensors or infrared cameras can be used to monitor human body temperature change, and Machine Learning (ML) algorithms can predict the thermal sensations of individuals. Alternatively, a more straightforward approach involves occupants reporting their sensations via smartphones. These methods facilitate a balance between personal comfort and energy conservation [8]. For instance, understanding occupant behavior can reduce energy usage by anticipating occupancy patterns [9], and optimizing control systems through predictive modeling can further improve energy efficiency [10].

A challenge that hinders the applicability of such solutions is that evaluating them in diverse scenarios is difficult. First, using devices to gather occupant data can raise privacy issues, and regularly soliciting occupant feedback is not always practical. Additionally, installing sensor infrastructure and integrating it with HVAC systems can be beyond the reach of many researchers. Past studies have primarily focused on buildings in the US, Europe, or China, often overlooking the unique climate needs of other regions [6]. To effectively assess these systems, a standardized benchmark is crucial [11]. This benchmark would allow for consistent evaluation, enabling researchers to refine their systems through comparative analysis and assisting building administrators in choosing the most appropriate system for their specific contexts.

Co-zyBench, our proposed benchmark, aims to assess and enhance thermal comfort systems using Digital Twins (DTs) of buildings, HVAC systems, and occupants. DTs are virtual representations of physical entities, capable of simulating the behaviors of HVAC systems, occupancy patterns, and thermal preferences. This provides a valuable tool for analysis and decision-making. Integrating DTs into the benchmarking process offers multiple advantages. Firstly, it is a cost-effective alternative to real-world building evaluations, enabling researchers to consider a diverse range of environments that may not be feasible in physical settings. Second, building administrators can assess the potential performance of these systems in their buildings without actual implementation.

Co-zyBench employs state-of-the-art simulation tools to

model occupant movement and thermal sensations within these DTs, using historical occupant data, and to simulate changes in temperature caused by HVAC manipulation. A co-simulation middleware in Co-zyBench facilitates dynamic interaction between the DTs. This middleware moves occupants within the virtual building, generating their expected thermal sensations in response to the dynamic environmental conditions controlled by the evaluated system. The benchmark proposes three metrics to evaluate the performance of the thermal comfort provision systems with respect to their ability to: 1) Ensuring general occupant comfort; 2) Aiming to prevent consistent discomfort for a minority group; and 3) Effectiveness in reducing energy consumption. In summary, the main contributions of this paper are:

- We present a novel benchmark to evaluate the performance of thermal comfort provision systems based on pervasive computing technology.
- We create a Digital Twin modeler for the user’s building that reflects and predicts occupant behavior and building energy performance.
- We introduce a middleware that integrates a co-simulation strategy and mediates between DTs of occupants and buildings to provide a realistic and dynamically evolving context.
- We propose an individual thermal sensation modeler designed to account for occupant movement in the building and diversity in thermal preference.

The rest of the paper is structured as follows. Section II overviews the state of the art. Section III presents the overall methodology of the benchmark by the framework design and utilization process. Section IV details the design of each component of the benchmark and the evaluation metrics. Section V shows a scenario implementation in the benchmark and some analysis problems based on the evaluation. Finally, Section VI concludes the paper.

## II. RELATED WORK

The proliferation of pervasive computing technologies like smartphones and sensors has spurred significant research into personalized thermal comfort in buildings [6], [7], [12]. These studies primarily address the challenge of estimating or capturing occupants’ thermal sensations and preferences, employing two main approaches. On the one hand, given that most occupants likely possess a personal device like a smartphone, several methods involve soliciting direct input from users. This includes collecting demographic data (age, height, weight, gender) to infer thermal sensation or directly asking them to report their current thermal comfort [8], [13]. The main applicability challenge of this approach is motivating consistent participation from occupants and managing incomplete data. On the other hand, other works utilize interconnected networks of sensors, actuators, and devices to continuously gather and transmit data on energy usage, environmental conditions, and occupant behavior [6], [7], [12]. Various devices, including wearables [6] and infrared cameras [7], have been used to

monitor temperature changes and predict thermal preferences. Additional data like heartbeat rates [12], obtained via smartphones [14] and in-ear microphones [15], or physical activity levels, detected through wearables [16], [17], also contribute to understanding occupants’ thermal sensations. The main applicability challenges for sensor-based systems are the costs associated with widespread deployment and encouraging users to share data from personal devices.

The prevalent method in the literature for evaluating thermal comfort systems primarily involves using datasets gathered through monitoring or questionnaires, which track energy consumption in lighting, HVAC, and appliances [18]. However, this approach is often tedious, costly, and sometimes not feasible. While some studies conduct experiments in actual buildings [19], simulations have emerged as an alternative to address these challenges [20]. Realistic datasets containing comfort data, such as the ASHRAE Global Thermal Comfort Database I and II [21], [22], argued that synthetic data may outperform real datasets for comfort as they remove the invalid data and potentially contain more diversity.

Utilizing Digital Twin (DT) models for thermal comfort systems is a promising avenue for the design, optimization, and maintenance of HVAC systems [23]–[25]. Co-zyBench leverages this approach by facilitating the benchmarking of such systems within a DT environment. For instance, Khayatian et al. [20] used DTs to analyze the energy consumption of various HVAC controllers, providing a platform for comparative assessments in identical environmental conditions. They also measured thermal comfort by tracking deviations from the recommended indoor temperature range. However, current methodologies often overlook the subjective nature of comfort and the diversity of human profiles [26]. Very few studies have considered performance across different regions and climate zones [18], [26], which significantly influence overall energy consumption [27]. To the best of our knowledge, no prior work has fully integrated simulations that encompass both synthetic yet realistic buildings and occupants, including simulating occupant interactions with their environment and the consequent thermal sensations. Crucially, these simulations should reflect the diversity of occupants and climate regions, an aspect that remains underrepresented in current evaluation methods. Co-zyBench aims to fill this gap by providing a comprehensive and realistic benchmarking platform.

## III. CO-ZYBENCH OVERVIEW

The goal of Co-zyBench is to provide a simulation-based testing platform specifically designed for evaluating thermal comfort systems that utilize sensors and smart devices. Co-zyBench is an open-source platform and the code is available on GitHub [28]. The benchmark includes three primary components (see Figure 1):

- **Occupant Digital Twin** which creates a virtual representation of building occupants, incorporating both static information (age, gender, height, weight) and dynamic data. Utilizing sensor inputs, it generates occupant movement trajectories within the building and dynamically simulates

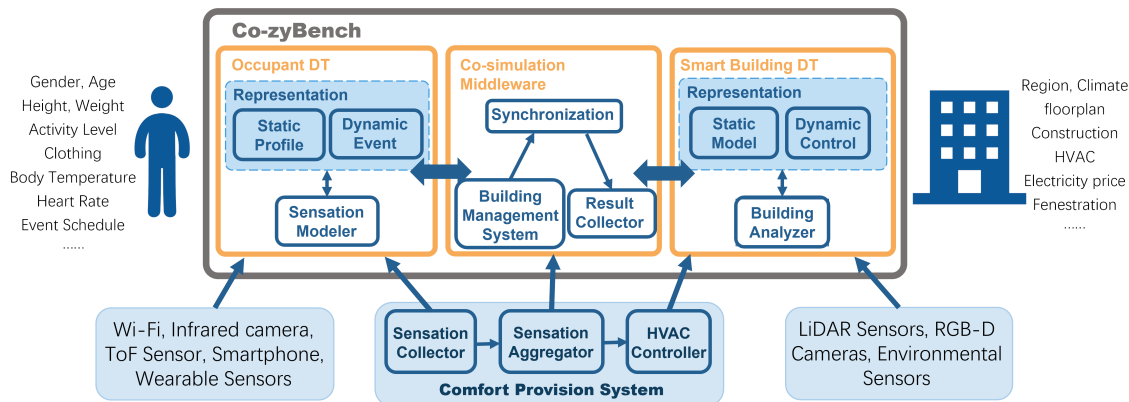


Fig. 1: Overview of the Co-zyBench benchmark.

their thermal sensations. These simulations are based on individual profiles and the prevailing environmental conditions, ensuring a personalized and realistic experience.

- **Smart Building Digital Twin** which maintains a DT of the building itself, including its HVAC system. It is responsible for implementing changes to the HVAC system and simulating temperature variations in different spaces. These simulations consider both environmental factors and dynamic influences, such as occupant presence and activities, providing a comprehensive view of the building’s thermal dynamics. In this paper, our primary focus is on the thermal comfort provided by HVAC systems. While other energy systems (e.g., lighting control) could potentially be integrated using our co-simulation middleware, these are beyond the scope of our current work.
- **Co-simulation Middleware** which serves as the communication bridge between the two DTs. The main task is to exchange information about the indoor thermal conditions from the smart building DT with the Occupant DT to generate realistic changes in the thermal sensations of people. It also interacts with the thermal comfort provision system that is being evaluated.

For users to evaluate their thermal comfort provision systems using Co-zyBench, they need to integrate these systems into the platform. A typical thermal comfort system comprises three key components: 1) Sensation Collector responsible for gathering data related to occupant thermal comfort. It collects individual sensation data, which could include subjective feedback from occupants or objective data from sensors. 2) Sensation Aggregator which aggregates the individual thermal sensations into a collective measure of comfort for a group of occupants. 3) HVAC Controller which translates the aggregated thermal comfort data into actionable HVAC control commands. To facilitate the evaluation of these systems in Co-zyBench, a wrapper is required around the user’s thermal comfort system. This wrapper communicates information from the co-simulator (e.g., thermal sensation of users, state of the HVAC system, etc.) to the system and from the system (e.g., the specific setpoint for the general HVAC system or each room) to the co-simulator. Co-zyBench provides users with

reference wrapper code and sample code for standard thermal comfort provision systems. This is particularly helpful as it allows users to focus on modifying only the aspects of the system they are interested in, rather than building everything from scratch.

Co-zyBench is structured to be a highly adaptable and user-friendly platform, enabling the evaluation of various thermal comfort provision systems within DTs that reflect the user’s actual buildings or a range of pre-included templates. Figure 2 depicts the overall steps required to set up the benchmark and use it. The user begins by importing the thermal comfort provision system they wish to evaluate into Co-zyBench (**Step 1**). Then, the user selects a benchmark scenario which can be either: a DT based on a specific building and occupant characteristics (**Step 2.a**); or one of the Co-zyBench reference DTs which are standard scenarios Co-zyBench provides (**Step 2.b**). Finally, the Co-zyBench user runs the benchmark (**Step 3**) and obtains the evaluation results (**Step 4**). A more detailed guideline is provided at [28].

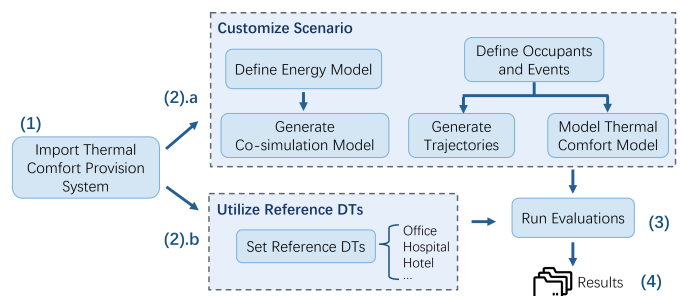


Fig. 2: Co-zyBench workflow.

#### IV. CO-ZYBENCH COMPONENTS

In this section, we detail each component of Co-zyBench including its evaluation metrics.

##### A. Evaluation Metrics

Co-zyBench introduces three evaluation metrics that are designed to explore the impact of a diverse workforce and a

sustainability focus on reducing energy consumption, as well as the actual thermal comfort provided to individuals:

**Individual Thermal Comfort (ITC):** The Thermal Sensation Vote (TSV) [29] is de facto standard to represent the thermal comfort preferences of users in the literature. The TSV scale models thermal sensation in a 7-point scale from -3 (very cold) to 3 (very hot). Hence, thermal comfort provision systems attempt to predict the specific value in the scale that would match the current thermal sensation of an individual. To assess the performance of the system, we compute the difference, in absolute value, between the system’s prediction and the real thermal sensation of each user. This way, this metric computes how uncomfortable an individual would feel if the system implements a temperature change based on its prediction. Hence, the higher this value is, the higher the discomfort the user would feel. Since there are multiple people and multiple spaces in the building, we aggregate this information as follows. Let the thermal sensation of  $n$  people in zone  $z$  at round  $r$  be  $TS_r[z] = \{TS_r[z][1], \dots, TS_r[z][n]\}$ . We define the metric,  $ITC$ , for number of people  $n$  in zones  $Z$  for all the round  $R$  as

$$ITC = \frac{\sum_{r=0}^R \sum_{z=0}^Z \sum_{i=1}^n |TS_r[z][i] - GTC_r[z]|}{\sum_{r=0}^R \sum_{z=0}^Z n}$$

where  $GTC_r[z]$  models the computed group thermal comfort value for zone  $z$  at round  $r$ . Hence, a strategy that minimizes the average discomfort that people experience will minimize the  $ITC$  value.

**Thermal Comfort Equality (TCE):** Even today, inequity in thermal comfort, particularly based on gender and/or age, has been highlighted as a challenge that requires rethinking equality in air conditioning in buildings [30]. We propose TCE by aggregating the idea from [13]. In this work, the authors proposed a mathematical model of *fairness* in participatory thermal comfort provision based on the fairness definition for the carpooling problem [31]. The idea is to compute, each round, a notion of how much of the overall discomfort of a group, based on the group thermal comfort value computed by the strategy, each individual would have to suffer to make the system fair. This value, referred to as an *Ideal Loss* (or discomfort), is computed by splitting  $ITC_r[z]$  equally among all the people in the zone. Then, based on this the authors propose a notion of *Extra Loss* which compares the discomfort of one participant  $P_i$  against this ideal situation as  $EL[P_i] = ITC_r[z][P_i] - \frac{ITC_r[z]}{n}$  where  $n$  is the number of participants in the zone  $z$  at round  $r$ . A positive, or negative, value for  $EL[P_i]$  means that this participant had a lesser, or greater, discomfort than what would be expected if the overall discomfort of the group was shared equally. Since it would be impossible for a system to be fair in the decision made for each round, the authors also introduce the notion of *Accumulated Extra Loss* of an individual at a round  $r$  as the summation of the *Extra Loss* the individual has experienced until that round. A strategy would be “fair” if, for the  $n$  users in the zone and at the round  $r$ , their *Accumulated Extra Loss* in absolute value is bounded by some constant  $M$  which is independent of  $r$  [31].

**Energy Consumption (EC):** This metric represents the amount of energy consumed (in kWh) by the HVAC system to implement the decided group thermal comfort of the user’s system. Given that there exists a mechanism to translate a group thermal comfort value to a temperature setpoint value for HVAC, this metric computes the required energy to achieve that setpoint for each of the zones and for all of the rounds. Note that the temperature setpoint value might change in every round which means that in some specific rounds, the energy consumed might be higher than others. Specifically, if the group thermal comfort obtained is feeling cold, then this might translate into setting up a higher setpoint than the current one which, if the HVAC is only cooling the zone, might not consume energy during that round. This metric can be measured using various real IoT devices or simulation approaches. This paper uses the standardized and open-source tool EnergyPlus [32] in our experimental evaluation.

## B. Smart Building Digital Twin

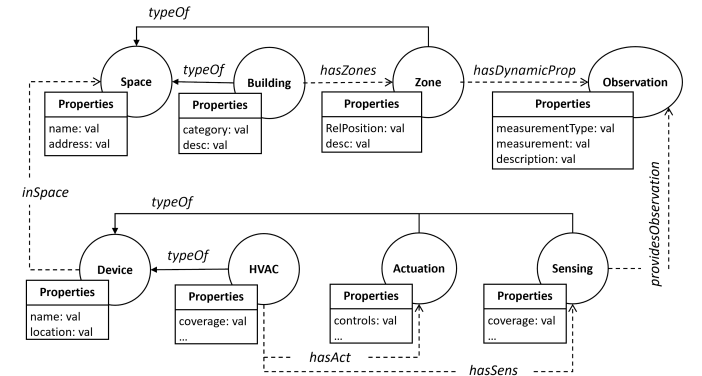


Fig. 3: The Smart Building-based NGSI-LD property graph.

Providing benchmark users with diverse buildings and HVAC models contributes to performing robustness testing of thermal comfort provision systems. However, exhaustively defining the diversity of building types worldwide is practically impossible due to its complexity [33]. To tackle this, the NGSI-LD specification offers a solution for creating DTs by providing a set of data models for smart environments [34] that we integrate and further extend in our work. As depicted in Figure 3, a building can be represented as a *property graph* which defines *entities* (e.g., room), *relationships* (e.g., hasRooms), and *properties* (e.g., RelPosition) as the key components of the building information model serialized in JSON-LD. Using the NSGI-LD building specification, CozyBench supports both the definition of a real building for creating its DT (spaces, devices, etc.), as well as a set of *reference buildings* already defined as NGSI-LD entities.

To define the DT of a real smart building, the first step involves defining the smart building’s NGSI-LD property graph (Figure 3) based on its spaces (floor plan, etc.) and exploited devices (HVAC system, IoT devices, etc.). Spaces definition includes metadata about the different floors and

zones, their type and maximum capacity. Information about the number of windows and doors (along with their thermal conductivity), insulation material of external walls, insulation material of the roof, and other construction materials are defined to determine heat transfer and heat storage surfaces. In terms of defining an HVAC system, this consists of a heating unit, a cooling unit, fans, and ducts to move the heated or cooled air around, settings such as its structure, capacity, coefficient of performance, and the number of fans with their maximum flow rate. Regarding IoT devices, environmental, occupancy, motion, devices for occupant feedback, and more, can be defined based on the NGS-LD property graph.

Another requirement for the DT of the building is the climate zone in which it is located based on the climate zone definition from the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) standard [35]. Finally, the DT includes the latest available Typical Meteorological Year (TMY) weather reports derived from the recent 15 years (2007-2021). Note that acquiring such smart building information can be time-consuming and it is out of the scope of this paper. On the other hand, building administrators can rely on building reports or on sensor-driven pervasive computing systems to facilitate this process. For example, LiDAR sensors and RGB-D cameras can achieve the automatic generation of building models [36], [37].

Regarding standard buildings, Co-zyBench includes a widely used set of 16 reference building types developed by the U.S. Department of Energy (DoE) which includes offices, warehouses, hospitals, and hotels [38] (approximately covering 70% of commercial buildings in the USA). The DoE buildings used do not include information about their internal structure in terms of rooms, corridors, etc. However, this information is essential to generate realistic trajectories to simulate realistic occupant behavior and reaction to HVAC changes. To ameliorate this issue we have generated realistic indoor floor plans based on state-of-the-art techniques [39]. Based on the type and size of the reference buildings, Co-zyBench includes different HVAC systems such as Variable Air Volume (VAV).

As already pointed out, these building representations will be utilized for simulation as part of the Co-zyBench Co-simulation middleware (Section IV-D).

### C. Occupant Digital Twin

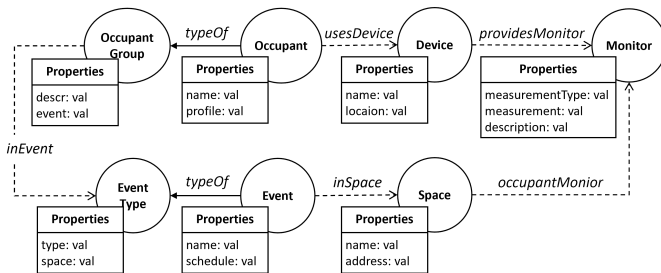


Fig. 4: The Occupant-based NGS-LD property graph.

To create the occupant DT, Co-zyBench considers three main aspects: their profiles that influence their thermal comfort preferences, their location at different points in time, and their thermal sensation in a specific location with a specific temperature. However, NGS-LD standard does not offer a model for occupants within buildings. Co-zyBench proposes an extension of the NGS-LD specification to support a standard way of defining building occupants. As depicted in Figure 4, an *Occupant* (John) has a set of properties (including his age, height, weight, gender, etc.) and belongs to an *Occupant Group* (e.g., students) that attend an *Event* (e.g., Programming in C) that belongs to an *Event Type* (e.g., Programming classes). Events are located in spaces where occupants can exploit devices while roaming from one event to another. Co-zyBench proposes a standard way of defining the occupant-based NGS-LD property graph that is then used to define the occupant DT.

In particular, occupants of a building typically do not stay in a specific space (e.g., their cubicle or a specific office) throughout the entire duration. To comprehend where the individuals are, we can define them as components of dynamic occupant groups engaged in different types of events or activities such as work, meetings, or lunch breaks. Secondly, occupants are diverse with regard to factors like age, gender, race/ethnicity and there is a correlation between these and their thermal preferences (e.g., based on their gender [40], age [5], or events [41]). Hence, to reflect this diversity, it is required to model the diverse characteristics present among individuals. This information will be abstracted and used to estimate their thermal preferences (e.g., whether they would prefer warmer or colder temperatures or whether they would be neutral). Thirdly, based on the thermal profiles there is a requirement to model how a person would feel based on the current temperature, for which we use the TSV scale, which consists of seven categories with values from -3 to 3: “cold”, “cool”, “slightly cool”, “neutral”, “slightly warm”, “warm”, and “hot” (e.g. based on [40], men feel comfortable when the temperature is 22°C, this translates to 0 on the scale, while women would feel comfortable at 25°C).

The profiles and activity schedules of the occupants enable the modeling of their moving trajectories which is essential for representing how the environmental conditions of a space will impact their thermal sensation. To this end, Co-zyBench integrates a trajectory generation model along with the NGS-LD property graph which defines the semantic representation of space, occupants and events. Their integration extracts meaningful patterns from sensorized spaces which are subsequently applied within an event-driven simulation to produce synthetic but realistic data, encompassing events, trajectories and thermal sensations. For example, to attend a meeting or take a break with friends, Co-zyBench generates realistic trajectories of occupants in a smart building and their thermal sensation based on their current location.

Algorithm 1 shows the high-level process followed by Co-zyBench to generate thermal sensations, the most important process of the occupant DT. Initially, Co-zyBench generates

---

**Algorithm 1** Thermal Sensation Generation

---

**Input:** Seed data  $data$ , Occupants  $O$ , Events  $E$ , Spaces  $S$ , Simulation time  $t$ , Temperature  $Temp$

**Output:** Occupants Thermal Comfort  $TC$

```
1: configure_learning( $data$ )
2:  $event, occupants = traj\_learning(data, MP)$ 
3:  $trajectory = traj\_generation(event, occupants)$ 
4: for  $p$  in  $occupants$  do
5:    $location\_p = locate(trajectory, p)$ 
6:   if ( $location\_p$  in  $S$ ) then
7:      $current\_temp = get\_temp(t, Temp, location\_p)$ 
8:      $tc\_p = thermal\_comfort\_predict(p, current\_temp)$ 
9:      $TC \leftarrow tc\_p$ 
10:  end if
11: end for
12: return  $TC$ 
```

---

trajectories of occupants using *seeddata* (sensor datasets for locating people such as data from WiFi connections, cameras, ToF sensors, etc.) following the process in [42]. Based on Alg. 1 the *seeddata* are given as input to the *traj\_learning* module to extract features related to *occupant* and *event* (line 2 in Alg. 1). Then, based on each occupant’s location and the indoor temperature (obtained from the smart building’s DT), their thermal sensation is generated. Eventually, the algorithm returns the predicted thermal sensations of all occupants.

The thermal comfort predictor (line 8 in Alg. 1), provides the flexibility to incorporate historical thermal sensation data that can be extracted from a real building environment. In this paper, we leverage the ASHRAE Global Thermal Comfort Database II [22], which collects 107,463 thermal comfort samples from the last two decades of thermal comfort field studies around the world. We implemented a K-Nearest Neighbors (KNN) model to predict thermal comfort. Assuming that the nearest neighbors are correctly classified, the KNN algorithm tallies the occurrences of each group among the K neighbors and identifies the group with the highest count. We trained the model with parameters of indoor temperature, occupants’ profiles as introduced in Figure 7a and their activity level which is obtained by event simulations. With the KNN model, we can simply but accurately predict how occupants would feel in their current zone.

#### D. Co-simulation Middleware

The co-simulation middleware connects the two DTs (building and occupant) through a synchronization mechanism that includes a range of functions such as time-step coordination, data exchange, and simulation control. Synchronization enables the simulators to advance the simulation at the same speed through time-step coordination, guaranteeing that data exchanges and interactions occur at the correct moments. The sequence diagram in Figure 5 illustrates the overall orchestration of an experiment by the co-simulator for a given scenario and a given HVAC control system. At the beginning of the evaluation, the Evaluator (a Co-zyBench user), initiates the process by providing the thermal comfort provision system

( $tc\_system$ ) to be evaluated and the selected scenario setup ( $building, people$ ) as inputs.

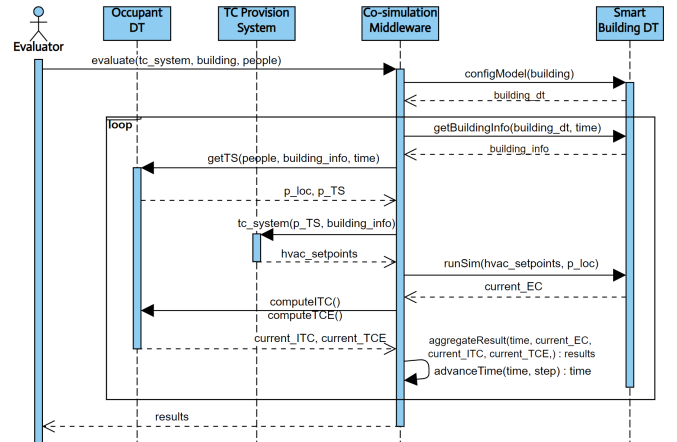


Fig. 5: Sequence diagram of Co-zyBench.

The middleware starts the building simulation to obtain building conditions (*building\_info*), such as indoor and outdoor temperature and energy consumption of HVAC system components. Building information, along with simulation time (*time*) and scenarios of the occupant (*people*), are used to simulate people’s location and thermal sensations ( $p\_loc, p\_TS$ ). By executing the user-provided  $tc\_system$ , the middleware obtains the HVAC setpoints ( $hvac\_setpoints$ ) by aggregating  $p\_TS$  based on  $building\_info$ . These parameters are used as inputs for running the building simulation for the next time step. Additionally, the parameter  $p\_loc$  is set in the building simulator, as our benchmark assumes the occupants to be accompanied by equipment that generates heat. This makes our simulation more realistic and takes into account the impact of occupants’ movements on the building’s energy performance. The building simulation outputs the energy consumption ( $current\_EC$ ) and update  $temp\_building$ . To get the metrics ITC and TCE results, the middleware requests the Occupant Simulator to compute  $current\_ITC$  and  $current\_TCE$ . Finally, the middleware advances the simulation time by step and the co-simulation is repeated until reaching the end of the simulation time.

## V. EXPERIMENTS

We showcase Co-zyBench by evaluating the performance of standard thermal comfort provision systems. Our ultimate goal is to assess: (1) How does the effectiveness of thermal sensation collections affect the overall thermal comfort and energy consumption; (2) How poorly do alternative systems perform when fairness is not taken into account; and (3) How differently do the systems perform in various climate zones.

### A. Experimental Setup

**Experimental Smart Building DT.** Figure 6a shows the floor plan of the office space, with a gross cooling area of 371 m<sup>3</sup>, in our experimental smart building DT. The common areas



(i.e., conference rooms, break room, kitchen, and restroom) are shared by different people at different points in time while the offices are shared by workers who are assigned to specific workplaces. Based on the type and size of the building, we select the VAV system because it is the most used HVAC system in commercial buildings and can achieve good thermal comfort levels while consuming less energy [43]. Figure 6b presents a diagram of the single duct VAV system that is distributed to multiple thermal zones and which is incorporated to Co-zyBench. This is mainly consisting of one outdoor air (OA) system, one Air Handling Unit (AHU), sensors, and thermal zones with VAV boxes. The mixture of the return and outdoor air from the OA system is delivered to thermal zones and conditioned by the AHU on the way to a zone. The controller supervises the supply air temperature in the duct and operates the coils and VAV fan in AHU to control the pressure and temperature of the supply air. Then, based on climate zone definition and construction suggestion from ASHRAE [35], the building for diverse climate zones has been designed.

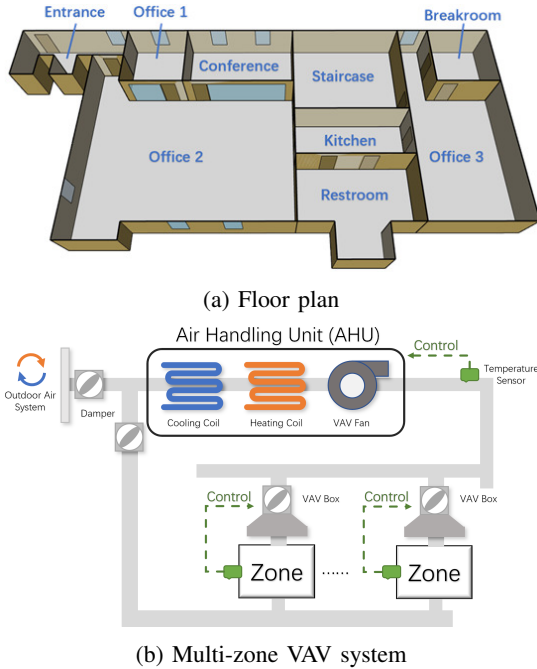


Fig. 6: Example building and HVAC system in our experiments.

**Experimental Occupant DT.** We created 18 different occupants for the experiment with diverse profiles (see Figure 7a) including static information such as their job and office, that can be used to predict their activities and location, and their gender, age, height and weight can be used to model their thermal preferences. Note the weight attribute is abstracted to Body Mass Index categories. Figure 7b shows the meeting schedules used to simulate their trajectories. People start working in the building at around 9 am and leave after 6 pm. People spend most of their working time attending a “working event” that occurs in their assigned offices. Their schedule

Office	Name	Gender	Job	Age	Height	Weight
Office Small	Michael	male	Boss	51	175	2
	Erin	female	Receptionist	33	165	2
Office Large	Jim	male	Seller	33	191	1
	Dwight	male	Seller	47	189	2
	Pam	female	Seller	39	168	1
	Andy	male	Seller	39	182	2
	Stanley	male	Seller	55	180	3
	Phyllis	female	Seller	64	166	2
	Oscar	male	Accountant	54	175	1
	Kevin	male	Accountant	40	185	3
Office Medium	Angela	female	Accountant	42	155	2
	Meredith	female	Supplier Manager	49	163	3
	Creed	female	None	70	183	3
	Toby	male	HR	46	177	1
	Gabe	male	Marketing	29	193	1
	Clark	male	Customer Service	28	175	1
	Pete	male	Customer Service	28	177	1
	Kelly	female	Sales Representative	34	163	1

(a) Occupant Profile

TIME	Monday	Tuesday	Wednesday	Thursday	Friday
9:00					
9:30	Boss, Sellers				
10:00			Boss, Accountants, Supplier Manager	Sellers	
10:30	Sellers, Marketing, Sales Representative	Sellers			Sellers, Marketing, Sales Representative
11:30					
12:00 - 14:00	Lunch Time				
14:00					
14:30	Accounts		Sellers, Marketing, Sales Representative		All
15:00					
15:30					

(b) Meeting Schedule

Fig. 7: Example occupant information in our experiments.

of weekly and daily events involves two types of meetings happening in the conference room: small meetings, held by some members of one or two teams mainly in the morning lasting 30 to 60 minutes; and large meetings, held by some members of two or more teams lasting around 90 minutes. For example, sellers and marketers have a small meeting every two days for at most one hour in the morning, all the staff have a large meeting at 14:00 for at most two hours every Friday. Besides, we model a lunchtime event from 12:00 to 14:00 and limit the amount of time spent for lunch to around one hour in the kitchen, the break room, or the outside. Moreover, people also attend non-time-fixed events like taking a break in the break room and going to the restroom.

**Experimental Systems.** We evaluated the performance of two systems based on the standard design of thermal comfort provision systems (see Section II). The main difference between the two systems is their sensation collection/estimation modules. In **System 1**, individuals employ their smartphones or a website to convey their thermal comfort sensation, while in **System 2**, we assume that sensors can monitor the occupants and obtain information about their profile and current activity and estimate their sensation based on it. For both systems, we configure the benchmark to collect individuals’ sensations every 30 minutes. To comprehensively evaluate these systems, we considered multiple versions of each, denoted as  $SX_{100}$ ,  $SX_{75}$ ,  $SX_{50}$ ,  $SX_{25}$ , and  $SX_0$  (with X representing either 1 or 2). In the case of System 1, these versions correspond to scenarios where 0% - 100% of users actively provide feedback through their smartphones, while the remainder do not. For System 2, the versions reflect instances where the sensor-based and ML-driven system encounters challenges in accurately



estimating 0% - 100% of the required features for thermal sensation prediction. In our scenarios, the 25% case only considers activity levels that can be modeled through various smart devices. And then combine data on age, gender and weight levels until 100%. We include an additional baseline that represents the standard approach used in most buildings which disregards the presence or the explicit thermal sensation of the occupants of the building [44] and just sets up a constant temperature of  $23^{\circ}\text{C}$  (observed to be a favorable temperature for a group of people [45]).

Both systems use the same sensation aggregators:

- **Majority Rule (MR).** We chose to evaluate the majority rule to reach consensus, usually considered a robust rule [46] since it always satisfies the majority’s requirement and maximizes the thermal comfort level of the overall workforce.
- **Drift Approach (DA).** Purdon et al. [8] present the following aggregation strategy to provide thermal comfort while conserving energy: If the summary of the thermal sensations of a group is higher than 0 (or lower than 0) the output group thermal sensation is 2 (or -2). If the sum is zero, which means that everybody is comfortable, instead of maintaining the current temperature, they apply a *drift* towards the temperature outside of the building to save energy (e.g., representing that the group is slightly warm or cold depending on whether is winter or summer, respectively).
- **Fairness Approach (FA).** Shin et al. [13] propose an approach to maintain *fairness* among the occupants of a building: people accumulate loss (§IV-A) when their discomfort is greater than the ideal fair discomfort. Otherwise, they reduce their accumulated loss. For example, if the aggregated thermal sensation is 0, people who feel warm (2 on the scale) get more loss than those who feel slightly warm (1 on the scale). To maintain fairness, the approach takes into account the accumulated loss of each individual in each round and chooses a group thermal sensation value that minimizes the accumulated loss of the individual with the highest loss so far.

To control the HVAC based on the estimated group thermal sensation, we follow the findings in [8]. Their experiments suggest that changing the temperature by 1 degree every 30 minutes is naturally accepted by people. Our translation function hence increases/decreases the current HVAC setpoint by  $0.5^{\circ}\text{C}$  every 10, 15, or 30 minutes if the aggregated thermal sensation is -3/3, -2/2, or -1/1. Since we collect sensations every 30 minutes the total change from one collection to the next is 1.5, 1.0, or  $0.5^{\circ}\text{C}$ , respectively.

**Experimental Configuration.** Our experiment focuses the evaluation on workdays in months of the year where most energy consumption takes place and adapts the HVAC system to perform: cooling in summer and heating in winter (132 days in total). When simulating the HVAC system, Co-zyBench controls it on a temperature setpoint where the HVAC system is switched on if receives a demand from the thermal comfort

provision system. We run the benchmark for each combination of sensation collector and sensation aggregator approach  $SX_p^{agg}$  (e.g.,  $S1_{100}^{MR}$  for smartphone-based collection with all individuals specifying their sensation and with the majority rule for aggregation). Each evaluation is run 5 times and the results are averaged to reduce the impact of randomness.

### B. Analysis of the Results

**Analysis 1: Influence of Inaccurate Thermal Sensation Estimation.** Ensuring active participation from users in feedback systems, such as System 1, poses a considerable challenge as individuals may not always have time or willingness to engage with the system. Similarly, achieving accurate thermal sensation estimation based on sensor data, as seen in System 2, can be problematic due to potential sensor failures or the acquisition of inaccurate information. In this experiment, we aim to compare the thermal discomfort experienced by individuals across different scenarios. Specifically, we assess the two systems with a percentage from 0 to 100 in comparison. The results are shown in Figure 8 presenting the Individual Thermal Comfort (*ITC*), defined in §IV-A, where higher values result to occupants discomfort.

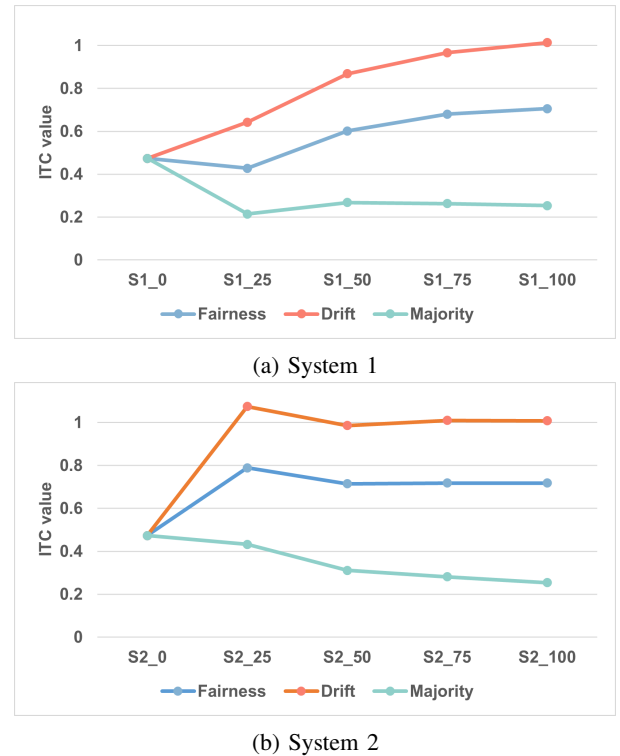


Fig. 8: Individual Thermal comfort (ITC).

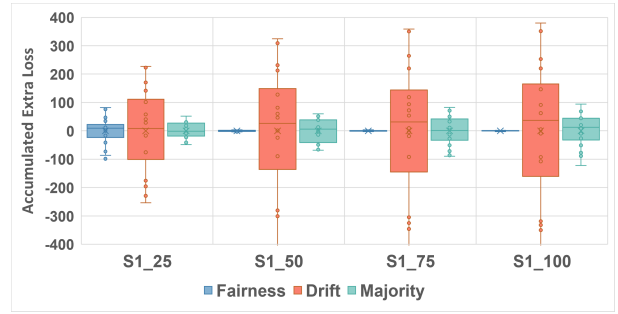
Since  $S1_0$  and  $S2_0$  represent the extreme situation in which both systems fail to estimate the thermal sensation of the occupants, they always set the temperature in the spaces to 23 degrees which is comfortable for most people [40]. Hence, they perform the same regardless of the aggregation strategy used. Also, we can think of it as an indicator that a value below it means that the algorithm is able to fulfill the occupants’

sensations better. We observe that, as expected, using the majority rule to aggregate the thermal sensations of the group outperforms all the other strategies by maintaining the lowest *ITC* value (since the *ITC* value is lower when the group comfort value computed is the median of occupants’ thermal sensations). In our scenario, like in many real buildings, most people have similar thermal preferences even if there is a minority that differs. The Fairness Approach, which sometimes introduces discomfort to a large number of people to satisfy the requirements of the minority, and the Drift Approach, which sometimes disregards all thermal preferences to lower energy consumption, as expected perform worse. As we will see next, this is a tradeoff with how fair and energy efficient such aggregation strategies are.

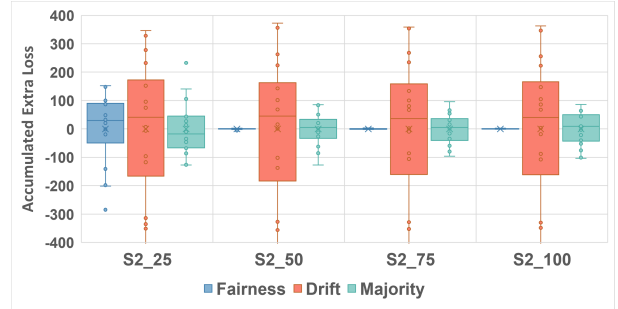
We observe that System 2, which predicts thermal sensation based on estimated parameters such as occupant’s age/weight/height and activity level collected through sensors, obtains slightly worse results than System 1 for most of the scenarios. This is due to the challenge of predicting thermal sensation vs. directly requesting this information from users. As expected, we also observe that having more information on occupants to model their thermal sensations can have a positive effect on thermal comfort. However, in System 1, using the Fairness and Drift Approaches makes occupants feel more comfortable when fewer people participate. The rationale for this outcome is evident: using these two algorithms yields worse results compared to not considering occupant comfort at all. Hence, having fewer occupants’ feedback can make more people feel comfortable but this simultaneously undermines their advantages of achieving equality and saving energy.

**Analysis 2: Evaluating Aggregation Strategies on Equality.** The choice of thermal comfort provision strategy plays an essential role, particularly in scenarios where multiple individuals share a common space. Depending on the strategy employed, issues related to fairness can arise, impacting the overall user experience. In this analysis, we focus on assessing the impact of various aggregation strategies on the pervasive personalized thermal comfort system. Figure 9 shows the results w.r.t. the Thermal Comfort Equality (*TCE*) metric (defined in §IV-A). The figure shows the values for all people as box plots displaying the minimum, maximum, median, and first and third quartiles of each scenario. We interpret these plots as the distribution of the *AccumulatedExtraLoss*. The more decentralized the data is plotted, the more unfair the strategy is (an *AccumulatedExtraLoss* of zero for all people would be the ideal Thermal Comfort Equality where all discomfort was equally distributed among them).

We can clearly find that the Fairness Approach, as expected, obtains the best results overall since it was designed with this goal in mind. The minimum standard deviation of the *AccumulatedExtraLoss* of all the scenarios using the Fairness system is around 1.4 while for other strategies it increases 146 and 28 for Drift and Majority, respectively. This highlights how unfair the latter approaches are with respect to some specific individuals who happen to be in the minority consistently throughout their working day. In addition to the



(a) System 1



(b) System 2

Fig. 9: Thermal Comfort Equality (TCE).

strategy comparisons, our findings reveal that the Fairness strategy performs worse in the 25% scenarios. This issue can be attributed to the fact that a 25% is insufficient to accurately model thermal sensation. Thus, for building managers aiming to optimize system performance w.r.t. equality, it becomes imperative to ensure an adequate percentage of either sensor data or user participation in voting processes.

**Analysis 3: Energy Consumption in Different Climate Zones.** In this analysis, we turn our attention to the impact of climate zones on system performance. To gain a comprehensive understanding, we specifically examine energy consumption as a key metric. By doing so, we aim to shed light on the system’s adaptability and effectiveness in diverse climatic conditions. Here we chose five cities including Scranton (USA) where the television show *the office* that inspired of experimental setup takes place. Table I contains information on the building and ASHRAE Climate Zone (CZ) of the five cities.

TABLE I: Buildings across different cities in Co-zyBench.

City	ASHRAE CZ	Insulation External Wall	Conductivity of Windows	Insulation of Roof
Mumbai	Extreme Hot Humid	3.4cm	2.1W/mK	17cm
Cairo	Hot Dry	4.5cm	0.042W/mK	21cm
LA	Warm Dry	5.6cm	0.019W/mK	21cm
Paris	Mixed Humid	6.8cm	0.013W/mK	21cm
Scranton	Cool Humid	7.9cm	0.013W/mK	21cm

Figure 10 presents the results for one particular case,  $S_{1,100}$  (which is equivalent to  $S_{2,100}$ ). Additionally, it compares using a personalized thermal comfort system with using a traditional HVAC control policy in buildings: Maintaining the temperature constant to a certain temperature consistently. In

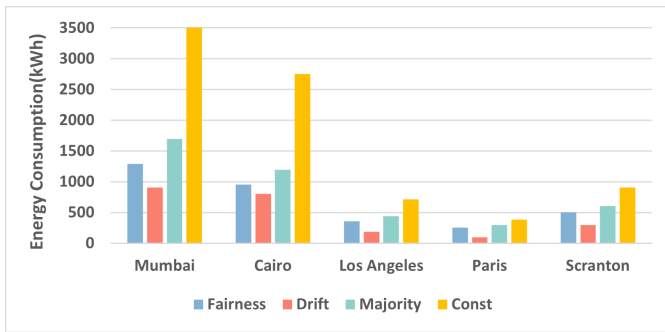


Fig. 10: Energy Consumption (EC) per city.

our case, this temperature set point is 23 degrees as explained before. Also, since some cities in our experiments do not need heating systems for the whole year, to compare fairly, Figure 10 shows only the required energy consumption to cool down the building for the three summer months. We can observe an expected and clear difference in actual consumption between warmer and colder climate zones. When focusing on specific strategies, we observe the high price that the Constant system pays. This is because, without occupants’ feedback, it regulates the HVAC system even if the space is empty. Of course, increasing the fixed temperature further would reduce energy consumption but would have a clear impact on both the fairness and thermal comfort of individuals. And this is also the reason why in warmer cities, it consumes much more energy than the others. The Drift strategy outperforms the others but at the expense of people’s comfort and equality.

**Discussion:** After analyzing the results, it becomes evident that each thermal comfort system has its advantages. In the case of the Sensation Collector system, System 1 outperforms System 2 when compared at the same percentage level. However, since System 2’s standout feature is its independence from participation of people, in the real world one would expect to compare the systems at different percentage levels (e.g.,  $S1_{50}$  vs  $S2_{75}$ ) which makes the results closer. Among the three Sensation Aggregators we assessed, each strategy excels in the feature it has been designed to optimize: overall thermal comfort for Majority Rule (MR), equality for the Fairness Approach (FA), and energy consumption for the Drift Approach (DA). However, note that when looking at the overall performance across metrics, even when there is no clear winner, FA achieves a good tradeoff in performance. We have also shown that these results are highly dependent on the distribution of occupant profiles and could be different in situations where there is a different majority (e.g., a majority of women vs. men). Co-ZyBench supports executing experiments in variations of a scenario which makes it possible, with a small configuration change, to evaluate the performance of the presented (or any other system) in multiple configurations which would be infeasible to do otherwise.

### C. Threats to Validity

This section introduces internal and external threats to validity that may impact the reliability of our Co-zyBench.

**Internal Validity.** As described in Section IV, Co-zyBench utilizes simulators for the occupants and buildings. Even though the simulators have proven their effectiveness, they are based on a limited set of parameters that may not fully capture the complexity of real-world conditions. Such inconsistencies between DTs and their counterparts potentially affect the accuracy of occupant movement modeling and energy simulations. To mitigate this effect, Co-zyBench provides available standards and real-world datasets for designing realistic scenarios. However, due to the inherent limitation of available data and current techniques, some level of deviation between simulated outcomes and real-world conditions is unavoidable.

**External Validity.** Our benchmark primarily focuses on evaluating thermal comfort provision systems on specific aspects of comfort and energy consumption. The applicability of our benchmark to systems with different concerns, for example, operating lighting and window shading, may be limited. However, such systems can be incorporated into Co-zyBench if simulation tools like EnergyPlus exist and offer an API for integration with the co-simulation middleware.

## VI. CONCLUSIONS

We presented Co-zyBench, a benchmarking platform for evaluating and optimizing pervasive computing based thermal comfort provision systems in smart buildings in terms of individual thermal comfort, thermal comfort equality, and energy consumption. In addition, our platform offers flexibility by allowing users to import their Digital Twin model for an accurate assessment of the performance of their systems in specific buildings.

To demonstrate the applicability of Co-zyBench, we conducted experiments by creating a Digital Twin of the building and occupants, analyzing three thermal comfort provision strategies to compare their performance. For the scenarios we assume, the observations can facilitate the building administrator to make a decision to employ thermal comfort systems based on their requirements. In addition, our analysis reveals that it is possible to develop techniques that can effectively balance equality and thermal comfort while simultaneously reducing energy consumption. In future work, we plan to maintain Co-zyBench extended by incorporating new standard buildings. Additionally, we plan to evaluate more thermal comfort provision systems and create a website that can help the community see the results to upload theirs.

## ACKNOWLEDGEMENTS

This work is partially supported by the Horizon Europe project DI-Hydro under grant agreement number 101122311, the Energy4Climate Interdisciplinary Center (E4C), which is in part supported by 3rd Programme d’Investissements d’Avenir [ANR-18-EUR-0006-02], and the China Scholarship Council (CSC).

## REFERENCES

- [1] F. Belic, Z. Hocenski, and D. Sliskovic, “Hvac control methods-a review,” in *2015 19th International Conference on System Theory, Control and Computing (ICSTCC)*. IEEE, 2015, pp. 679–686.

- [2] International Energy Agency, "The future of cooling," 2019, license: CC BY 4.0. [Online]. Available: <https://www.iea.org/reports/the-future-of-cooling>
- [3] J. Mazon, "The influence of thermal discomfort on the attention index of teenagers: an experimental evaluation," *International journal of biometeorology*, vol. 58, no. 5, pp. 717–724, 2014.
- [4] M. e. a. Schweiker, "Drivers of diversity in human thermal perception—a review for holistic comfort models," *Temperature*, vol. 5, no. 4, pp. 308–342, 2018.
- [5] P. Lenzuni, D. Freda, and M. Del Gaudio, "Classification of thermal environments for comfort assessment," *Annals of Occupational Hygiene*, vol. 53, no. 4, pp. 325–332, 2009.
- [6] J. Xie, H. Li, C. Li, J. Zhang, and M. Luo, "Review on occupant-centric thermal comfort sensing, predicting, and controlling," *Energy and Buildings*, vol. 226, p. 110392, 2020.
- [7] 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, 2018.
- [8] S. e. a. Purdon, "Model-free hvac control using occupant feedback," in *LCN Workshops*, 2013.
- [9] A. e. a. Albatayneh, "The impact of the thermal comfort models on the prediction of building energy consumption," *Sustainability*, vol. 10, no. 10, 2018.
- [10] S. Taheri, P. Hosseini, and A. Razban, "Model predictive control of heating, ventilation, and air conditioning (hvac) systems: A state-of-the-art review," *Journal of Building Engineering*, p. 105067, 2022.
- [11] N. Mims, S. R. Schiller, E. Stuart, L. Schwartz, C. Kramer, and R. Faesy, "Evaluation of us building energy benchmarking and transparency programs: Attributes, impacts, and best practices," Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), Tech. Rep., 2017.
- [12] K. N. Nkurikiyeyezu, Y. Suzuki, and G. F. Lopez, "Heart rate variability as a predictive biomarker of thermal comfort," *Journal of Ambient Intelligence and Humanized Computing*, vol. 9, pp. 1465–1477, 2018.
- [13] E.-J. e. a. Shin, "Exploring fairness in participatory thermal comfort control in smart buildings," in *BuildSys*, 2017.
- [14] K. Jiokeng, G. Jakllari, and A.-L. Beylot, "Handrate: Heart rate monitoring while simply holding a smartphone," in *2021 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 2021, pp. 1–11.
- [15] K.-J. Butkow, T. Dang, A. Ferlini, D. Ma, and C. Mascolo, "heart: Motion-resilient heart rate monitoring with in-ear microphones," in *2023 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 2023, pp. 200–209.
- [16] B. Khaertdinov, E. Ghaleb, and S. Asteriadis, "Deep triplet networks with attention for sensor-based human activity recognition," in *2021 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 2021, pp. 1–10.
- [17] J. Morales, N. Yoshimura, Q. Xia, A. Wada, Y. Namioka, and T. Maekawa, "Acceleration-based human activity recognition of packaging tasks using motif-guided attention networks," in *2022 IEEE international conference on pervasive computing and communications (PerCom)*. IEEE, 2022, pp. 1–12.
- [18] M. McWhinney, A. Fanara, R. Clark, C. Hershberg, R. Schmeltz, and J. Roberson, "Energy star product specification development framework: using data and analysis to make program decisions," *Energy Policy*, vol. 33, no. 12, pp. 1613–1625, 2005.
- [19] "Comparing energy and comfort metrics for building benchmarking," *Energy and Buildings*, vol. 205, p. 109539, 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0378778819318055>
- [20] F. Khayatian, H. Cai, A. Bojarski, P. Heer, and A. Bollinger, "Benchmarking hvac controller performance with a digital twin," *EnerarXiv*, May, 2022.
- [21] R. J. De Dear, "A global database of thermal comfort field experiments," *ASHRAE transactions*, vol. 104, p. 1141, 1998.
- [22] V. F. Ličina, T. Cheung, H. Zhang, R. De Dear, T. Parkinson, E. Arens, C. Chun, S. Schiavon, M. Luo, G. Brager *et al.*, "Development of the ashrae global thermal comfort database ii," *Building and Environment*, vol. 142, pp. 502–512, 2018.
- [23] H. Hosamo, M. H. Hosamo, H. K. Nielsen, P. R. Svennevig, and K. Svidt, "Digital twin of hvac system (hvacdt) for multiobjective optimization of energy consumption and thermal comfort based on bim framework with ann-moga," *Advances in Building Energy Research*, vol. 17, no. 2, pp. 125–171, 2023.
- [24] A. Clausen, K. Arendt, A. Johansen, F. C. Sangogboye, M. B. Kjærgaard, C. T. Veje, and B. N. Jørgensen, "A digital twin framework for improving energy efficiency and occupant comfort in public and commercial buildings," *Energy Informatics*, vol. 4, pp. 1–19, 2021.
- [25] Y. Fathy, M. Jaber, and Z. Nadeem, "Digital twin-driven decision making and planning for energy consumption," *Journal of Sensor and Actuator Networks*, vol. 10, no. 2, p. 37, 2021.
- [26] W. Zheng, Y. Chang, L. Zhang, W. Li, Y. Liu, and S. Cho, "A passive building energy-efficiency evaluation method considering human thermal adaptation with seasonal climate change," in *IOP Conference Series: Earth and Environmental Science*, vol. 581, no. 1. IOP Publishing, 2020, p. 012038.
- [27] D. Westphalen and S. Koszalinski, "Energy consumption characteristics of commercial building hvac systems volume i: Chillers, refrigerant compressors, and heating systems," *Final Report to the Department of Energy (Contract No. DE-AC01-96CE23798)*, 2001.
- [28] <https://github.com/satrai-lab/cozybench>, Co-zyBench".
- [29] X. Zhou, Q. Ouyang, Y. Zhu, C. Feng, and X. Zhang, "Experimental study of the influence of anticipated control on human thermal sensation and thermal comfort," *Indoor Air*, vol. 24, no. 2, pp. 171–177, 2014.
- [30] T. e. a. Parkinson, "Overcooling of offices reveals gender inequity in thermal comfort," *Scientific reports*, vol. 11, no. 1, 2021.
- [31] R. Fagin and J. H. Williams, "A fair carpool scheduling algorithm," *IBM Journal of Research and Development*, vol. 27, no. 2, pp. 133–139, 1983.
- [32] "Energyplus." [Online]. Available: <https://energyplus.net/>
- [33] D. Prowler and S. Vierra, "Whole building design," *Whole Building Design Guide. National Institute of Building Sciences. Accessed*, vol. 2, 2008.
- [34] <https://smarddatamodels.org/>, Smart Data Models project.
- [35] A. . T. C. N. A. G. 30329-2305, "Climatic Data for Building Design Standards," 2019. [Online]. Available: <https://www.ashrae.org/technical-resources/standards-and-guidelines>
- [36] F. Rottensteiner, "Automatic generation of high-quality building models from lidar data," *IEEE Computer Graphics and Applications*, vol. 23, no. 6, pp. 42–50, 2003.
- [37] S. Vidas, P. Moghadam, and M. Bosse, "3d thermal mapping of building interiors using an rgb-d and thermal camera," in *2013 IEEE international conference on robotics and automation*. IEEE, 2013, pp. 2311–2318.
- [38] M. Deru, K. Field, D. Studer, K. Benne, B. Griffith, P. Torcellini, B. Liu, M. Halverson, D. Winiarski, M. Rosenber *et al.*, "Us department of energy commercial reference building models of the national building stock," 2011.
- [39] G. Pintore, V. Garro, F. Ganovelli, E. Gobetti, and M. Agus, "Omnidirectional image capture on mobile devices for fast automatic generation of 2.5d indoor maps," in *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2016, pp. 1–9.
- [40] B. Kingma and W. van Marken Lichtenbelt, "Energy consumption in buildings and female thermal demand," *Nature climate change*, vol. 5, no. 12, 2015.
- [41] P. O. Fanger and J. Toftum, "Extension of the pmv model to non-air-conditioned buildings in warm climates," *Energy and buildings*, vol. 34, no. 6, pp. 533–536, 2002.
- [42] A. e. a. Chio, "Smartspec: Customizable smart space datasets via event-driven simulations," in *PerCom*, 2022.
- [43] G. S. Okochi and Y. Yao, "A review of recent developments and technological advancements of variable-air-volume (vav) air-conditioning systems," *Renewable and Sustainable Energy Reviews*, vol. 59, pp. 784–817, 2016.
- [44] M. F. Haniff, H. Selamat, R. Yusof, S. Buyamin, and F. S. Ismail, "Review of hvac scheduling techniques for buildings towards energy-efficient and cost-effective operations," *Renewable and Sustainable Energy Reviews*, vol. 27, pp. 94–103, 2013.
- [45] M. e. a. Campano, "Analysis of the perceived thermal comfort in a portuguese secondary school: Methodology and results," *Int. J. Eng. Technol*, vol. 9, no. 6, pp. 448–455, 2017.
- [46] P. Dasgupta and E. Maskin, "On the robustness of majority rule," *Journal of the European Economic Association*, vol. 6, no. 5, pp. 949–973, 2008.