



Evaluation of thermal control based on spatial thermal comfort with reconstructed environmental data

Youngkyu Kim ^{a,1}, Byoungyun Yoo ^{a,c}, Ji Young Yun ^b, Hyeokmin Lee ^a,
Sehyeon Park ^a, Jin Woo Moon ^b, Eun Ji Choi ^d,*

^a Intelligence and Interaction Research Center, Korea Institute of Science and Technology, 5 Hwarangro14-gil, Seongbuk-gu, Seoul, 02792, South Korea

^b School of Architecture and Building Science, Chung-Ang University, Dongjak-gu, Seoul, 06974, South Korea

^c AI-Robotics, KIST School, Korea National University of Science and Technology, 5 Hwarangro14-gil, Seongbuk-gu, Seoul, 02792, South Korea

^d School of Architecture, University of Southern California, 850 Bloom Walk, Los Angeles, 90089, CA, United States of America

ARTICLE INFO

Keywords:

Indoor environment
Thermal comfort
Predicted mean vote
Data reconstruction
Gappy POD

ABSTRACT

Achieving thermal comfort while maintaining energy efficiency is a critical objective in building control. Conventional thermal comfort models, such as the Predicted Mean Vote (PMV), rely on both environmental and personal variables. However, the use of fixed-location sensors limits the ability to capture spatial variability, thereby reducing the accuracy of occupant-specific comfort estimation. To address this limitation, this study proposes a new PMV estimation method that utilizes spatial environmental data reconstructed using the Gappy Proper Orthogonal Decomposition (Gappy POD) algorithm. In addition, a novel control framework based on group PMV is developed to represent the thermal comfort of multiple occupants. The Gappy POD method enables fast and accurate reconstruction of indoor temperature fields from sparse sensor measurements. Using these reconstructed fields and occupant location data, spatially resolved PMV values are calculated. Group-level thermal conditions are then derived through statistical aggregation methods and used to control indoor temperature in a multi-occupant living lab environment. Experimental results show that the Gappy POD algorithm achieves an average relative error below 3% in temperature reconstruction. PMV distributions varied by up to 1.26 scale units depending on occupant location. Moreover, thermal satisfaction outcomes varied depending on the group PMV method employed. These findings underscore the importance for adaptive thermal control strategies that incorporate both spatial and individual variability, offering valuable insights for future occupant-centric building operations.

1. Introduction

To address the growing demands for sustainable architecture, energy efficiency, and enhanced occupant comfort, the development of advanced HVAC control strategies has become a critical focus in the field of architectural engineering. In particular, the information of occupant thermal comfort is considered a key factor in improving both occupant well-being and building energy performance [1–3]. With the increasing adoption of intelligent buildings, there is growing interest in quantitatively assessing thermal

* Corresponding author.

E-mail addresses: youngkyu_kim@berkeley.edu (Y. Kim), yoo@byoo.org (B. Yoo), yjy350@cau.ac.kr (J.Y. Yun), hyeokminlee@kist.re.kr (H. Lee), sehyeon.park@kist.re.kr (S. Park), gilerbert73@cau.ac.kr (J.W. Moon), ec_922@usc.edu (E.J. Choi).

¹ Equal contribution.

Nomenclature

Acronyms

AI	Artificial Intelligence
CFD	Computational Fluid Dynamics
CNN	Convolutional Neural Network
DNN	Deep Neural Network
FPS	Frames per Second
POD	Proper Orthogonal Decomposition
Gappy POD	Gappy Proper Orthogonal Decomposition
MAE	Mean Absolute Error
ML	Machine Learning
PMV	Predicted Mean Vote
SVD	Singular Value Decomposition
TSV	Thermal Sensation Vote

Symbols

RH	Relative humidity (%)
T_{AIR}	Air temperature (°C)
T_{MRT}	Mean radiant temperature (°C)
T_{SET}	Setpoint temperature
V_{AIR}	Air velocity (m/s)
M	Metabolic rate (met)
I_{cl}	Clothing insulation (clo)
$\text{PMV}_{\text{MEDIAN}}$	Median value of PMV data
PMV_{WA}	Weighted average value of PMV data
PMV_{MAD}	Mean absolute deviation value of PMV data
X	Snapshot matrix
x	High-dimensional state vector
$\hat{x}$	POD coefficients
x_{ref}	Reference state vector
Φ	POD basis matrix
U	Left singular vectors
Σ	Singular values matrix
V	Right singular vectors
Z^T	Sampling matrix
n_s	Number of POD basis vectors
n_r	Number of measurement points
$\dagger$	Moore–Penrose inverse of a matrix

comfort and integrating it as a core parameter in building operations [4–6]. As a result, various methods for assessing thermal comfort have been developed over time [7–11]. Thermal comfort is determined by a complex set of factors, including indoor environmental conditions, dynamic personal characteristics, and psychological responses. Early thermal comfort models primarily focused on environmental variables due to technological limitations. Over time, however, these models have been extended to incorporate personal factors, grounded in the human heat balance equation [8,12].

There are various models for assessing thermal comfort, and they continue to evolve to improve their accuracy. Among them, this study employed the Predicted Mean Vote (PMV), which has been consistently used as an international standard [7]. The PMV model incorporates four environmental variables indoor air temperature (T_{AIR}), relative humidity (RH), mean radiant temperature (T_{MRT}), and air velocity (V_{AIR}) as well as two personal factors. Accurately accounting for these input variables can enhance the predictive accuracy of PMV [13]. The PMV value is expressed on a 7-point integer scale: −3 (cold), −2 (cool), −1 (slightly cool), 0 (neutral), +1 (slightly warm), +2 (warm), and +3 (hot). A PMV of 0 represents thermal neutrality, and a PMV range of −0.5 to +0.5 is generally considered thermally comfortable for approximately 90% of people [12].

In earlier applications, the primary limitation of using PMV for building control was the difficulty of obtaining accurate personal factors. Due to technological limitations, personal factors were traditionally estimated using fixed or assumed values [12,14]. However, recent advances in image-based AI and activity recognition now enable non-intrusive, real-time estimation of occupants' metabolic rates and clothing insulation [15–20]. These developments now allow for more accurate and individualized PMV calculations based on real-time personal data.

Another key limitation in thermal comfort estimation lies in the measurement of environmental variables. While real-time environmental data are essential for accurate PMV calculations, discrepancies between sensor locations and actual occupant positions can significantly compromise the reliability of thermal comfort estimation. In typical buildings, sensors are installed at fixed points,

and their readings are often assumed to represent the conditions of the entire space. However, previous studies have demonstrated that factors such as system layout, building geometry, and surrounding objects can cause considerable spatial variation in thermal conditions [21,22]. Consequently, thermal sensation may vary depending on an occupant's location, and such discrepancies become more pronounced in cases of localized heating or cooling. To improve the accuracy of thermal comfort assessments, it is therefore necessary to incorporate environmental data that reflect the occupant's actual position [22,23].

However, deploying a high density of sensors around occupants in real building environments poses significant economic and technical challenges, making it impractical for application in typical indoor spaces outside of laboratory settings. To overcome these limitations in sensor deployment, recent studies have increasingly applied Digital Twin technologies that can predict and reconstruct environmental data by accounting for the complex physical characteristics of building spaces [24–27].

Among Digital Twin-based environmental prediction approaches, physics-based simulations using Computational Fluid Dynamics (CFD) and data-driven models based on Machine Learning (ML) have been most widely employed [28–30]. CFD enables high-resolution analysis of airflow and temperature distributions within indoor spaces, making it useful for capturing detailed environmental characteristics. However, its application in real-time scenarios and large-scale buildings is limited due to the complexity of model setup, the difficulty of defining boundary conditions, and the high computational cost. ML-based models, on the other hand, are capable of learning and predicting nonlinear and high-dimensional patterns from environmental data. Nevertheless, they typically require large training datasets and extensive hyperparameter tuning to achieve high performance. These constraints limit their applicability across diverse domains and varying environmental conditions.

Therefore, a faster and more efficient data reconstruction method is required for real-time environmental monitoring applicable to building control. To address this, data reconstruction algorithms that estimate full spatial environmental fields from a limited number of sensors have been introduced [31,32]. A representative technique in this context is the Gappy Proper Orthogonal Decomposition (Gappy POD) algorithm, which has been widely adopted as a prominent approach for reconstructing missing data from spatially sparse measurement (gappy data) [33–35]. By extracting dominant patterns through principal component decomposition, Gappy POD effectively infers unmeasured values. Its computational efficiency and stability make it particularly suitable for real-time applications in building environments. Data reconstruction algorithms provide access to spatially non-uniform indoor environmental information, enabling localized assessment of the conditions surrounding each occupant. This spatial resolution offers the potential for more accurate and individualized thermal comfort evaluation.

In addition, PMV model is designed to evaluate the thermal comfort of individual occupants. As a result, many previous studies that use PMV as a control variable have focused on single-occupant environments. In real-world multi-occupant settings, where a single HVAC system serves a shared space, it is inherently difficult to satisfy the thermal preferences of all occupants simultaneously. To improve overall comfort in such environments, a representative value of thermal comfort that reflects multiple occupants must be provided as a control variable for the HVAC system. However, studies that incorporate such metrics into building control remain limited, and a standardized methodology for representing group thermal comfort indices is still lacking.

Therefore, there is a pressing need to develop control strategies that leverage a representative thermal comfort index tailored for multi-occupants. This study aims to address this need by enhancing PMV estimation accuracy through real-time reconstruction of indoor environmental data and by evaluating the effectiveness of control strategies based on the actual PMV of multiple occupants. To this end, the Gappy POD algorithm is used to reconstruct spatial indoor temperature fields, enabling the analysis of real-time PMV accuracy and its variation based on occupant location. A group PMV-based control experiment is conducted in a living lab environment with actual occupants, and the resulting thermal comfort levels are quantitatively analyzed according to different control strategies. Through this approach, the study demonstrates the effectiveness of using Gappy POD for real-time PMV estimation and highlights the potential of group PMV-based control strategies to improve thermal comfort in multi-occupant environments.

2. Methodology

2.1. Data reconstruction algorithm

The PMV assessment depends on two key inputs: accurately reconstructed environmental conditions at occupant locations and real-time estimates of personal factors. To infer environmental conditions at an occupant's location, this study employs a data reconstruction algorithm that estimates spatial environmental fields from sparse sensor measurements.

Specifically, the Gappy POD method is adopted, a classical technique for reconstructing full-field data from limited observations [33–35]. This method relies on a linear subspace derived from a POD of snapshot data. POD reduces high-dimensional state data into a space with low-dimensional orthonormal basis using singular value decomposition (SVD), and approximates the full solution in this reduced space.

Once the POD basis is computed, Gappy POD reconstructs the full state using only partial measurements at selected spatial locations. The reconstruction is performed by estimating the POD coefficients (also known as generalized coordinates) that minimize the discrepancy between the measured and reconstructed values. Since the POD basis matrix is reused and only matrix–vector operations are needed, the method is computationally efficient and easy to implement.

In this study, Gappy POD is applied to indoor thermal data to reconstruct temperature fields from a limited number of sensors. This enables high-resolution estimation of the indoor thermal environment without the need for dense sensor deployment, forming the basis for the thermal comfort analysis and control strategies described in later sections.

It is demonstrated that Gappy POD achieves accurate spatial field reconstruction even with a small number of sensors, making it a suitable approach for indoor environmental monitoring.

The mathematical formulation of Gappy POD is as follows. The high-dimensional state $\mathbf{x} \in \mathbb{R}^{N_s}$ is approximated using the POD basis $\Phi \in \mathbb{R}^{N_s \times n_s}$ as:

$$\mathbf{x} \approx \tilde{\mathbf{x}} = \mathbf{x}_{ref} + \Phi \hat{\mathbf{x}}, \quad (1)$$

where $\mathbf{x}_{ref}$ is a reference solution and $\hat{\mathbf{x}} \in \mathbb{R}^{n_s}$ denotes the generalized coordinate (POD coefficient) in the reduced space, with $N_s \gg n_s$. The POD basis is obtained by performing SVD on a snapshot matrix $\mathbf{X}$, which contains data collected across various times and parameter conditions:

$$\mathbf{X} = \mathbf{U} \Sigma \mathbf{V}^T, \quad (2)$$

where $\mathbf{U}$, Σ , and $\mathbf{V}$ represent the left singular vectors, singular values, and right singular vectors, respectively. The first n_s columns of $\mathbf{U}$ are selected to construct the POD basis, yielding $\Phi = [\mathbf{u}_1, \dots, \mathbf{u}_{n_s}]$. When only sparse measurements are available, the POD coefficients $\hat{\mathbf{x}}$ are estimated by solving the following least-squares problem:

$$\hat{\mathbf{x}} = \underset{\hat{\mathbf{x}} \in \mathbb{R}^{n_s}}{\operatorname{argmin}} \|\mathbf{Z}^T (\mathbf{x} - \Phi \hat{\mathbf{x}})\|_2^2, \quad (3)$$

where $\mathbf{Z}^T \in \mathbb{R}^{n_r \times N_s}$ is a sampling matrix that extracts the rows corresponding to the measurement locations, n_r is the number of measurement points, and $\mathbf{Z}^T \mathbf{x}$ is the sparse measurement vector. This problem admits a closed-form solution:

$$\hat{\mathbf{x}} = (\mathbf{Z}^T \Phi)^\dagger \mathbf{Z}^T \mathbf{x}, \quad (4)$$

where † denotes the Moore–Penrose inverse of a matrix. The reconstructed full state is then computed using Eq. (1). Note that $(\mathbf{Z}^T \Phi)^\dagger$ is computed once and reused during reconstruction, which contributes to the overall computational efficiency of the method.

2.2. Thermal comfort evaluation

To estimate personal factors of PMV in real time, this study used the Personal Factor estimation model (PF model) developed based on the prior work of Choi et al. [6]. The PF model is a computer vision-based AI framework designed to estimate metabolic rate and clothing insulation from indoor image data in real time. The model integrates multiple neural network architectures including OpenPose, convolutional neural networks (CNNs), and Deep Neural Networks (DNNs) to simultaneously infer both metabolic rate and clothing insulation from a single image. A key advantage of the PF model is its ability to estimate individual personal factors for each occupant, even in multi-occupant environments.

The core process of the model consists of the following steps. First, individuals are detected from indoor images. Then, the detected person's joint coordinates are extracted, and their posture is classified to estimate activity level. At this stage, the model is capable of detecting movement from continuously captured images of occupants and identifying real-time changes in their activities. Simultaneously, the model identifies the garments worn by the individual in the same image and classifies them to estimate clothing insulation. The model estimates metabolic rates for five activity levels, including 0.7 met (sleeping), 1.0 met (sitting), 1.2 met (standing), 1.7 met (walking), and 3.5 met (exercising). In addition, the model can detect 16 garments across categories such as top, bottom, outerwear, dress, and pajama, enabling the estimation of clothing insulation for over 34 clothing combinations.

The trained model was evaluated using experimental data obtained from the building occupants, achieving an average F1 score above 0.93, demonstrating high classification performance. Additionally, the applicability of the model was validated in real-world settings, where the mean absolute error (MAE) was 0.04 met for metabolic rate and 0.03 clo for clothing insulation [6]. These results confirm the model's high feasibility for use in actual environments. Accordingly, this study employed the PF model to estimate personal factors for individual occupants. During the experiments, image data were collected at a rate of 2 frames per second (FPS), and personal factors were computed at 20-s intervals. For each interval, the mode value from the captured data was used as the representative estimate. Fig. 1 presents example images illustrating the PF model's estimation outputs for activity level and clothing insulation used in this study.

While the PF model enables simultaneous estimation of personal factors for multiple occupants, controlling a shared HVAC system requires a single representative value that reflects the thermal comfort of all occupants. In this study, the conventional PMV metric is modified to derive such a representative value, which is then used as a control variable. To achieve this, three statistical aggregation methods are applied to compute group-level PMV values from individual thermal comfort data [36].

The first method, $\text{PMV}_{\text{MEDIAN}}$, uses the median value of all individual PMV data. If the number of data points is odd, the central value is selected; if even, the average of the two middle values is used. This approach is less sensitive to outliers and effectively captures the central tendency of the distribution.

The second method, PMV_{WA} , applies a weighted average technique, in which each individual PMV value is assigned a weight based on its relative importance. The weighted average is computed using Eq. (5), where n denotes the number of individuals (i.e., the number of individual PMV values). Unlike the median-based approach, this method gives greater weight to PMV values associated with thermal discomfort, thereby incorporating the influence of outliers more deliberately. The weighting scheme is defined in Eq. (6). PMV values with absolute magnitudes less than or equal to 1 are considered to indicate relatively comfortable conditions and are assigned a weight of 1. For PMV values outside this range, the square of the PMV is used as the weight. This approach is based on the characteristic that the Predicted Percentage of Dissatisfied (PPD) increases nonlinearly with the square of the PMV, meaning that higher discomfort levels result in greater weights [37].

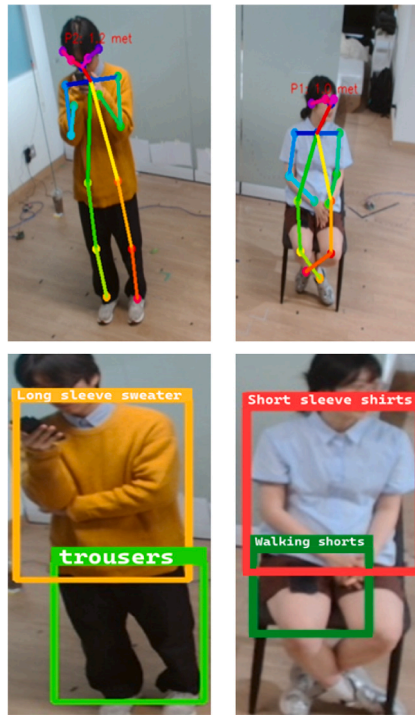


Fig. 1. Examples of PF model outputs for estimating personal factors.

$$PMV_{WA} = \frac{\sum_{i=1}^n w_i \cdot PMV_i}{\sum_{i=1}^n w_i}, \quad (5)$$

where

$$w_i = \begin{cases} 1, & \text{if } -1.0 \leq PMV_i \leq 1.0 \\ PMV_i^2, & \text{otherwise} \end{cases} \quad (6)$$

The third method, PMV_{MAD} , incorporates the mean absolute deviation (MAD) from the median to compute a representative group PMV. As shown in Eq. (7), the group PMV is calculated by adding the median of individual PMV values to the MAD, which is defined in Eq. (8), where n is the number of individuals (i.e., the number of individual PMV values). This approach is informed by observations that, particularly during the heating season, occupants may still perceive the environment as comfortable even when PMV values are slightly lower than neutral [37,38]. In this study, the MAD value was added to the PMV median as a correction factor, yielding a higher representative value than the median alone. This adjustment was intended to bridge the gap between objective PMV-based estimates and occupants' subjective thermal sensations.

$$PMV_{MAD} = PMV_{median} + MAD, \quad (7)$$

where

$$MAD = \frac{1}{n} \sum_{i=1}^n |PMV_i - PMV_{median}| \quad (8)$$

3. Experiments

3.1. Living lab

The living lab is located on the second floor of a research facility in Seoul, South Korea, with a total area of $14.80 \text{ m} \times 9.57 \text{ m}$. As shown in Fig. 2, the space consists of a control room and a living lab space. The living lab space occupies $11.1 \text{ m} \times 8.39 \text{ m}$ and includes a living room, kitchen, two rooms, and a bathroom. It is designed as an independent space, physically separated from the control room. The experiment was conducted in the experimental space measuring $11.1 \text{ m} \times 3.96 \text{ m}$, which includes the kitchen and living room areas. The control room was used by the administrator to monitor and manage the experiment. The living lab space includes a main entrance and a connecting door to the control room, but no windows. The control room has its own entrance, and a window is installed along one side of the wall. The living lab is equipped with an LRD-N837T inverter-type HVAC system, installed

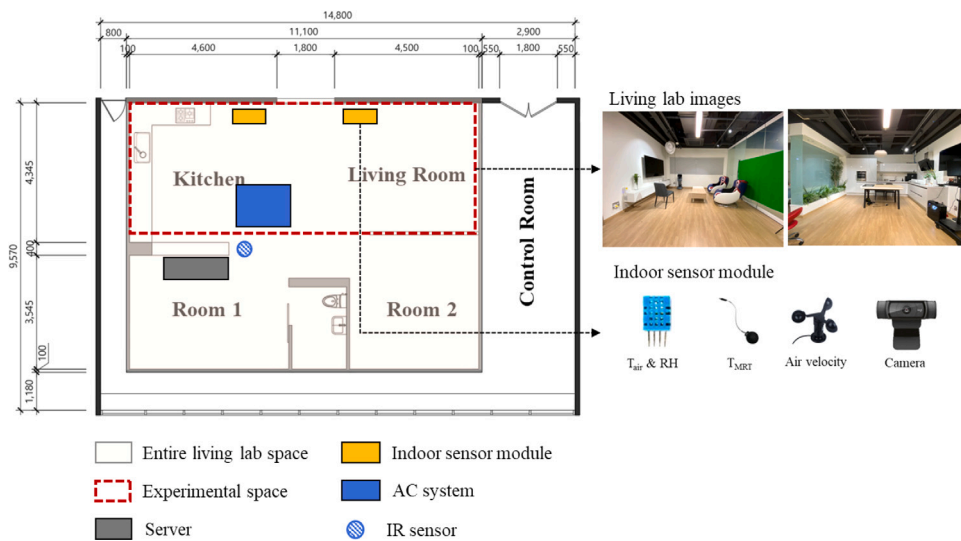


Fig. 2. Overview of the living lab environment.

as a four-way ceiling cassette indoor unit. The system is located between the kitchen and living room, with a cooling capacity of 8.3 kW and a heating capacity of 9.3 kW.

Sensors were installed in the experimental space to collect indoor environmental and image data. The environmental sensor module consisted of three types of sensors: sensors capable of measuring T_{AIR} and RH, sensors for T_{MRT} , and sensors for V_{AIR} . For real-time data acquisition, a microcontroller unit composed of a Raspberry Pi and Arduino-based sensors was used. DHT 11 sensors were employed to measure T_{AIR} , RH, and T_{MRT} , while a three-cup anemometer was used to measure V_{AIR} . To measure T_{MRT} , the DHT 11 sensor was enclosed in a black globe. All sensors were calibrated prior to field deployment. Two individual sensor modules were installed at regular intervals on the front wall at a height of 1.2 m. This height was selected to collect environmental data that closely represent the conditions typically perceived by occupants in general workspaces during both seated and standing activities. Environmental sensor data were collected at 5-s intervals. In addition to the above sensor modules, approximately 200 temperature sensors were used solely during the thermal data collection phase for constructing the snapshot data for Gappy POD, with their deployment described in detail in Section 3.2.

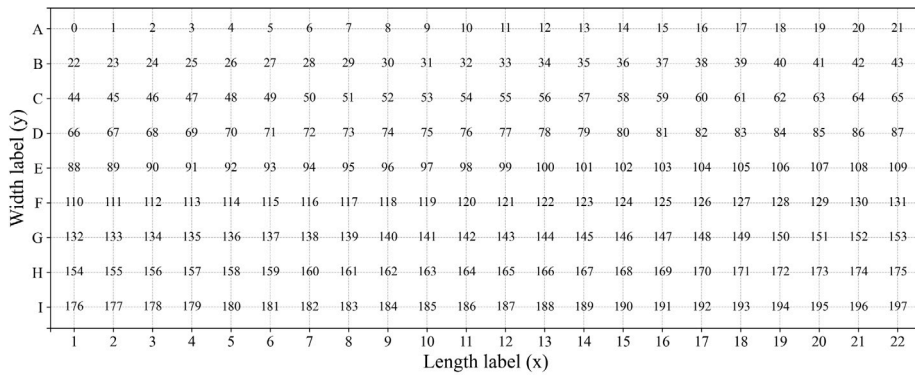
To capture indoor images for occupant detection and personal factor estimation, webcams (Logitech C920) were used, recording at 2 FPS. Two cameras were installed at a height of 2 m, positioned to cover the left and right sides of the space. Additionally, an infrared transmitter was installed near the HVAC system to enable automatic control during the experiment. All sensor data were transmitted and stored in real time on a local server within the living lab.

3.2. Data acquisition

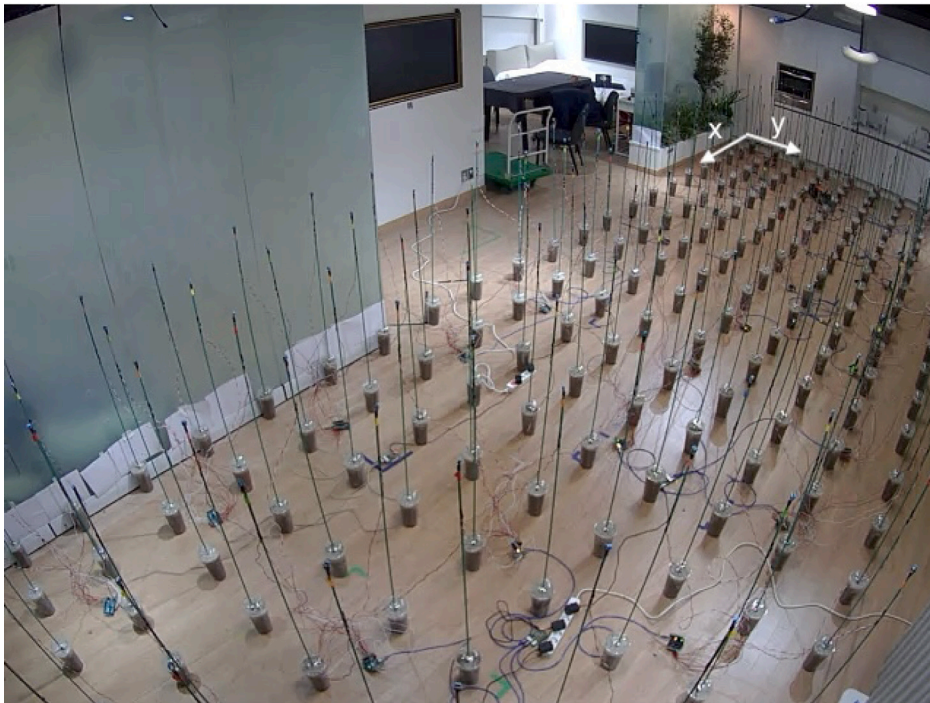
To obtain a training data set (i.e., snapshot data) for constructing the POD basis matrix, the living lab space (9.45 m $\times$ 3.2 m), excluding interior wall thickness and kitchen fixtures, was divided into a 9 $\times$ 22 grid. The row and column intervals were set to 40 cm and 45 cm, respectively, resulting in a total of 198 grid points within the rectangular space, as shown in Fig. 3(a). Temperature sensors were installed at a height of 1.2 m at each grid point, as illustrated in Fig. 3(b). According to the coordinate axes in Fig. 3(a), the row direction corresponds to the vertical (y-axis), and the column direction corresponds to the horizontal (x-axis). Each sensor measured temperature at 2-s intervals, and the data were continuously stored in a database.

Snapshot data were obtained across the temperature range that encompassed the heating system settings used during the thermal control experiments. Each snapshot scenario consisted of three stages of temperature control: an initial temperature, a first setpoint, and a second setpoint. The specific control sequence was as follows. First, after setting the initial temperature, the heating system operated for 10 min, followed by a 10-min off period. Next, the system was turned on for 15 min at the first setpoint temperature, and then switched to the second setpoint for another 15 min. Finally, the heating system was turned off and maintained in the off state for an additional 10 min. Thus, each scenario lasted for 60 min in total. The initial temperature was set to either 22 °C or 23 °C, and the first and second setpoint temperatures were selected from the set {22 °C, 23 °C, 24 °C, 25 °C, 26 °C}. This resulted in a total of $2 \times 5 \times 5 = 50$ parameter combinations.

For each scenario, data were collected during the last 2 min of three key time points: the first setpoint change, the second setpoint change, and immediately before the final heating system off time. A moving window average with a one-minute interval was applied to each period, resulting in 90 time-series samples per scenario. Consequently, each scenario yielded a two-dimensional snapshot array of size 90 $\times$ 198.



(a) Grid layout (9 × 22 grid, 198 points). The numbers 0–197 indicate the flat indices corresponding to the tuple of coordinate indices.



(b) Temperature sensors at grid points

Fig. 3. Layout of the living lab.

Among the 50 total scenarios, 8 specific combinations – with initial temperatures of 22 °C or 23 °C, first setpoints of 23 °C or 25 °C, and second setpoints of 23 °C or 25 °C – were used as the validation dataset. The remaining 42 scenarios were used as the snapshot dataset for constructing the POD basis matrix.

3.3. Experiment set-up

The participants consisted of four adults in their twenties—two males and two females. To minimize the effect of body mass index (BMI) on thermal sensation, only individuals within the normal BMI range ($18.5 \text{ kg/m}^2 < \text{BMI} < 24.9 \text{ kg/m}^2$) were recruited. The average age of the participants was 25.5 years, and the mean BMI was 20.5 kg/m^2 . The experiment was conducted over three days, from January 20 to January 22, 2025, during the heating season. Each day, the experiments were held from 9:00 a.m. to 6:00 p.m., including scheduled breaks.

Thermal control experiments were conducted by varying both the occupant's spatial location and personal factors. The experimental design included four occupant scenarios (A–D), with personal factor conditions arranged into four cases to ensure that subjects experienced distinct thermal sensation levels in each case. As shown in Table 1, personal factors were categorized into

Table 1
Personal factor conditions applied for the experiment.

Case	1	2	3	4
Metabolic rate (met)	1.0	1.2	1.0	1.2
Clothing insulation (clo)	0.5	0.5	1.0	1.0

Table 2
Subjects scenarios for the experiment.

Scenario	Subject 1	Subject 2	Subject 3	Subject 4
A	Case 2	Case 3	Case 4	Case 1
B	Case 1	Case 4	Case 1	Case 4
C	Case 2	Case 1	Case 1	Case 4
D	Case 3	Case 4	Case 4	Case 1

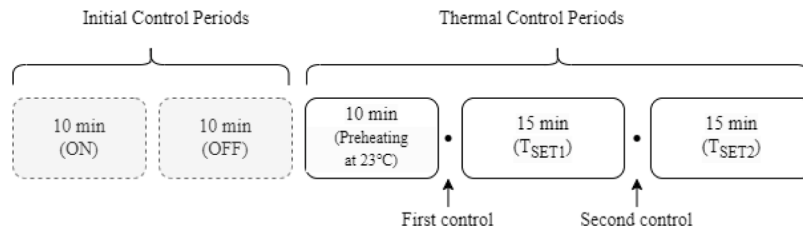


Fig. 4. Thermal control sequence.

two levels for each variable, resulting in a total of four cases. Metabolic rate was classified as either sitting (1.0 met) or standing (1.2 met), while clothing insulation was divided into light clothing (0.5 clo) and heavy clothing (1.0 clo). Table 2 presents the configuration of the four experimental scenarios (A–D) assigned to each participant. In Scenario A, all participants were assigned different cases, while Scenarios B, C, and D included overlapping cases among some participants. Under identical environmental conditions, the PMV values corresponding to the cases in Table 1 increased in the following order: Case 1 < Case 2 < Case 3 < Case 4.

Each of the four scenarios was repeated three times according to the group PMV methods described in Section 2.2, resulting in a total of 12 experimental runs. Subjects 1 through 4 participated from different locations within the space, and temperature data reconstructed for each participant's specific location was used during the experiment. The occupant locations were assigned based on the living lab grid shown in Fig. 3(a), corresponding to grid positions 114 (row F, column 5), 140 (row G, column 9), 147 (row G, column 16), and 106 (row E, column 19), respectively. For brevity, grid locations are referred to by their row–column indices (e.g., F5 represents the point at row F and column 5) throughout the remainder of this paper, unless coordinate values are explicitly provided.

During the experiment, participants reported their subjective thermal sensation and satisfaction through a structured questionnaire. The survey employed a 7-point Thermal Sensation Vote (TSV) scale, which corresponds to the PMV scale. It followed the point-in-time (short-term) method recommended by ASHRAE 55 [37], with data collected every 2 min. To ensure accurate responses, participants used a mobile survey platform and avoided actions that could affect thermal sensation, such as eating, drinking, or wearing masks.

3.4. Thermal control algorithm

The thermal control experiment for each scenario followed the sequence shown in Fig. 4. To ensure a consistent initial indoor temperature, an initial control period was first applied: the system was turned on at 22 °C for 10 min and then turned off for the following 10 min. Following this step, the thermal control period begins, comprising a preheating phase and two subsequent control points. The preheating phase set the initial temperature to 23 °C and lasts for 10 min. The control points refer to the moments when the system adjusts to the designated setpoint temperatures (T_{SET1} and T_{SET2}). Each control action is performed at 10 and 25 min, and the experiment concludes at 40 min.

A procedural difference exists between the Gappy POD training data acquisition and the thermal control experiment. While both processes share the same initial control period, the control process includes an additional 10-min preheating phase at 23 °C prior to the start of the thermal control period. This difference accounts for real occupancy conditions during the experiments. Unlike the data acquisition phase, which was conducted without occupant movement, participants in the experimental scenarios entered and exited the space during breaks between trials. These door operations introduced cold outdoor air into the room, potentially causing fluctuations in the initial indoor temperature. To minimize this impact and ensure consistent starting conditions, the additional preheating step was incorporated into the experimental procedure.

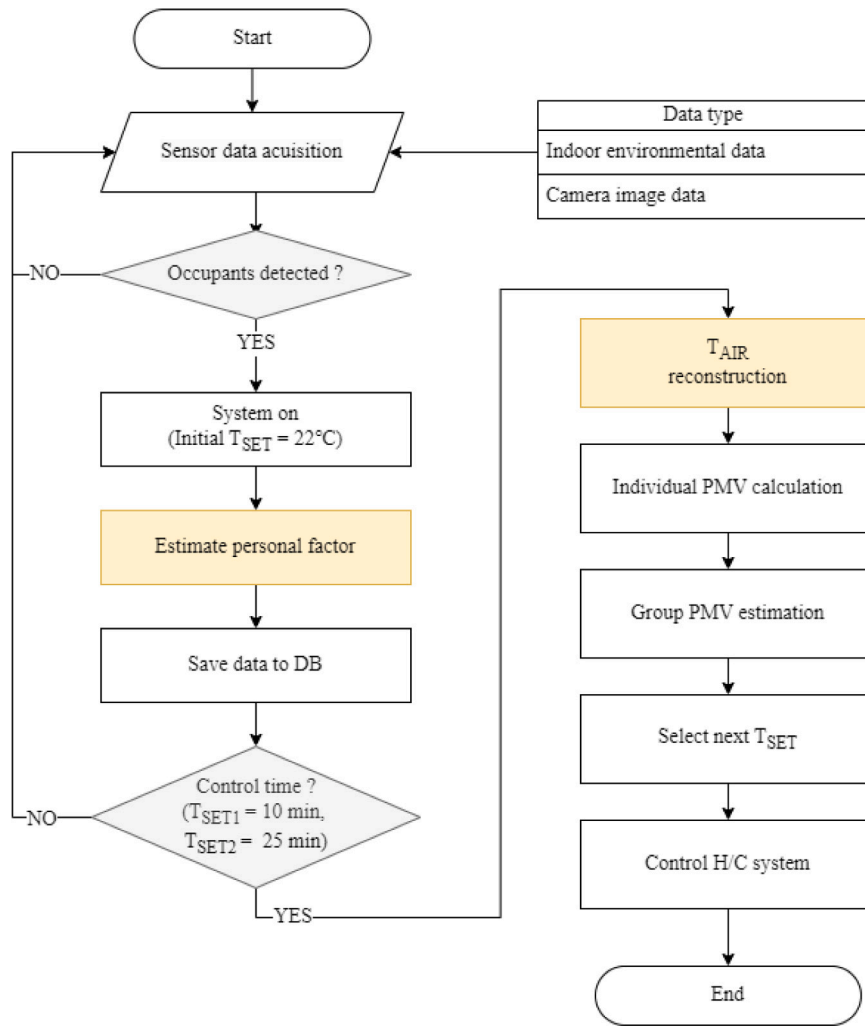


Fig. 5. Thermal control algorithm.

The thermal control algorithm is illustrated in Fig. 5. To determine the system's setpoint temperature, the control process iteratively follows the sequence outlined in the figure. First, real-time data are collected from sensors installed in the living lab. These include indoor environmental variables as well as indoor image data. Among them, the air temperature data used for the reconstruction algorithm are obtained every 2 s from the sensors located at the positions indicated in Fig. 6. Occupancy detection is performed using the collected image data. Upon detecting occupancy, the HVAC system is activated with a preheating of 23 °C. Subsequently, occupant personal factors—metabolic rate and clothing insulation—are estimated based on the captured images using the personal factor estimation model described in Section 2.2. All collected data are stored in a centralized database.

At each control interval, the system evaluates whether an adjustment is necessary. If so, the indoor temperature distribution is reconstructed using 12 temperature measurements taken along the room boundaries. The data reconstruction algorithm estimates zone-specific indoor temperatures based on measurements collected during the two minutes preceding each control point. Using these reconstructed temperature values, along with previously obtained occupant-specific activity levels and clothing insulation data, individual PMV values are calculated. These individual PMVs are then aggregated to estimate a group PMV. The system's optimal setpoint temperature is determined such that the group PMV approaches zero, and this setpoint is subsequently used to control the HVAC system.

During the heating experiments in the living lab, several physical constraints limited the range of feasible system setpoint temperatures. The experimental space was located within the interior of a building, which reduced exposure to outdoor temperature fluctuations and made it difficult to achieve indoor temperatures below 22 °C. Additionally, pre-tests revealed that due to aging HVAC equipment and decreased building airtightness, the system was unable to raise the indoor temperature above 26 °C. Considering these limitations, the operational setpoint temperature range was restricted to between 22 °C and 26 °C throughout the experiment.

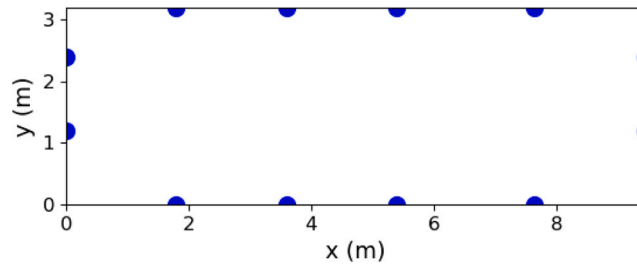


Fig. 6. 12 measurement points (boundary-based locations).

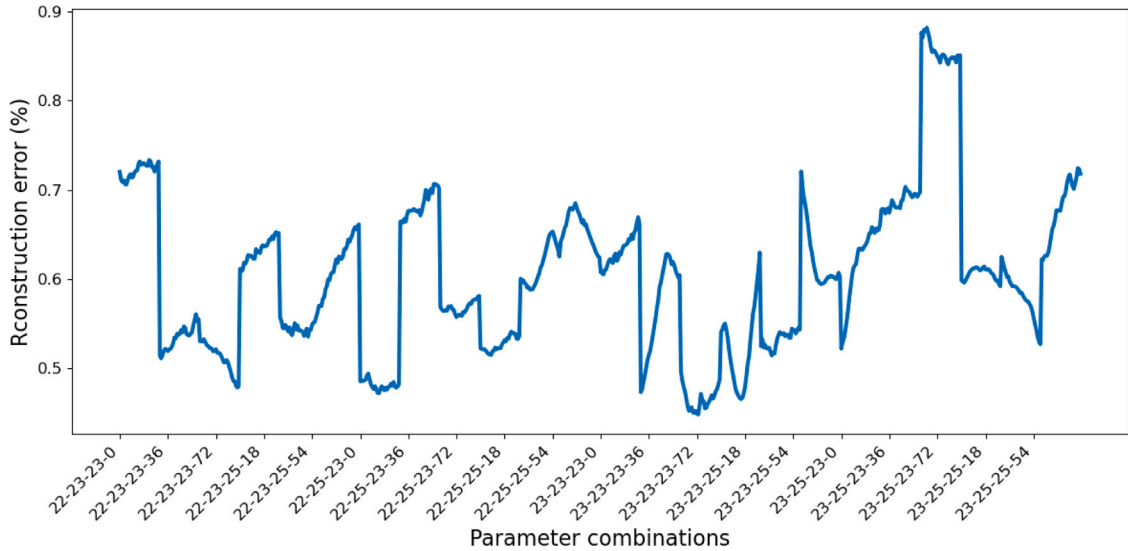


Fig. 7. Reconstruction error on the validation dataset using Gappy POD.

4. Results and discussion

4.1. Data reconstruction performance

4.1.1. Offline phase

The number of POD modes used to construct the POD basis matrix should be greater than the intrinsic dimension of the snapshot data, and at the same time, should be less than half the number of measurements to ensure numerical stability in Gappy POD. The former condition ensures that the key features of the data are adequately captured, while the latter helps avoid numerical instability caused by a large condition number during the computation of POD coefficients.

In this study, the intrinsic dimension of the snapshot data is four, consisting of three temperature setting parameters and one temporal domain and the number of available measurement points is 12, therefore, the number of POD modes is set to be 5. To evaluate reconstruction performance, we selected 12 sparse measurement locations evenly distributed along the boundaries of the living lab, as shown in Fig. 6.

The reconstruction error was computed using the following definition:

$$\text{reconstruction error (\%)} = \frac{\|\mathbf{x} - \tilde{\mathbf{x}}\|_2^2}{\|\mathbf{x}\|_2^2} \times 100 \quad (9)$$

where $\mathbf{x}$ represents the ground truth temperature field, and $\tilde{\mathbf{x}}$ denotes the reconstructed result.

The reconstruction performance on the validation dataset is presented in Fig. 7. The x-axis labels are formatted as i-j-k-t, where i is the initial temperature (°C), j and k represent the first and second setpoint temperatures (°C), respectively, and t indicates the index of the moving average window applied to the final 2 min of each control point. The average reconstruction error was approximately 0.603%, with a minimum of 0.448% and a maximum of 0.882%.

The reconstruction results for the best and worst cases, along with the differences from the ground truth, are shown in Fig. 8. Each column presents the reconstructed temperature field, the ground truth, and their difference. All fields are smoothed using cubic interpolation on a 4 times finer grid for visual clarity.

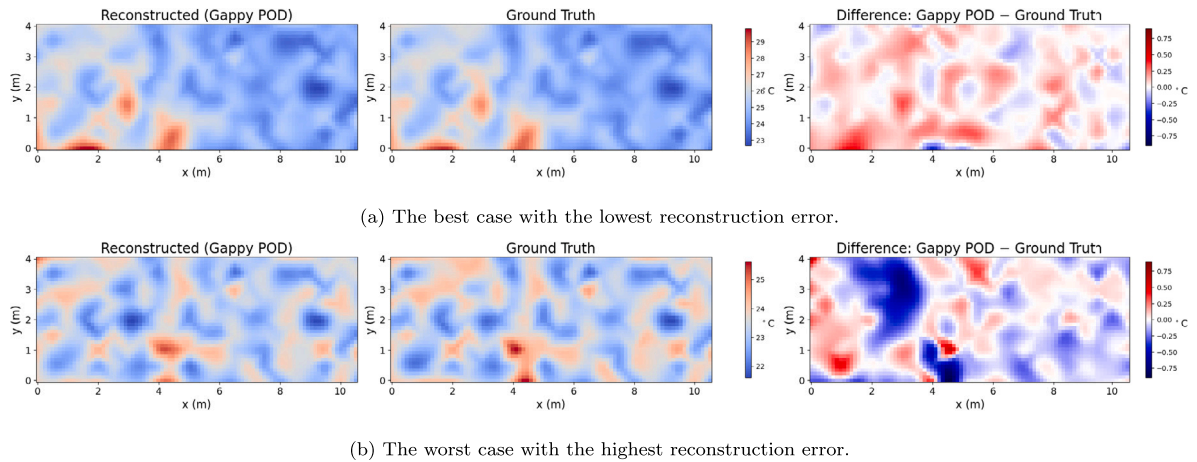


Fig. 8. Comparison of the best and worst reconstruction cases in the validation dataset.

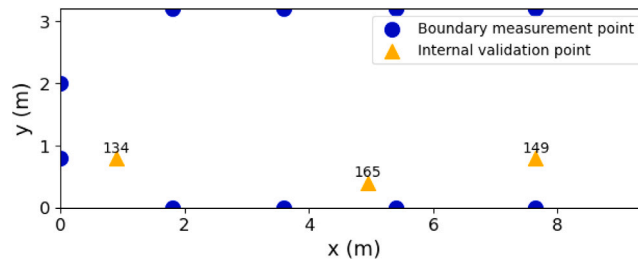


Fig. 9. Locations of measurement (boundary) and validation (internal) points.

Even in the worst case, the temperature difference remained below 1 °C. This is because the validation dataset falls within the interpolation range of the parameter sweep used to generate the snapshot (training) data.

4.1.2. Online phase

During the actual PMV control experiments, the environmental conditions differed from those of the data acquisition phase. Factors such as the presence of occupants, a reduced number of sensors, and varying door-opening frequency in the living lab may have negatively affected reconstruction accuracy. Since it was not feasible to measure the entire temperature field during the experiment, reconstruction performance was evaluated using three additional internal measurement points, in addition to the original 12 boundary sensors. The validation points were carefully selected to represent areas where relatively large temperature fluctuations were observed during data collection, while ensuring that their placement did not interfere with the experiment. The measurement locations are shown in Fig. 9, where the boundary points are marked by blue circles and the internal validation points by orange triangles. The numbers above the orange markers denote the flat indices of the corresponding grid points (refer to Fig. 3(a)).

At each PMV control point during the experiment, the relative error between the reconstructed temperature and the ground truth measured at the internal validation points was computed. The error was computed as $(T_{\text{measured}} - T_{\text{reconstructed}})/T_{\text{measured}} \times 100$. The results are plotted in Fig. 10. The average relative error is 2.13%, corresponding to a temperature difference of 0.56 °C.

4.2. Thermal comfort analysis

This section is organized into three parts to present the building thermal control results. Section 4.2.1 provides an auxiliary analysis of spatial temperature non-uniformity using reconstructed fields. Section 4.2.2 reports the PMV estimation results obtained from subject experiments. Finally, Section 4.2.3 analyzes the building control outcomes based on the method used to derive group PMV values.

4.2.1. Effect of reconstructed temperature on spatial PMV

The experiments were conducted based on the methodologies described in Sections 3.3 and 3.4. Indoor environmental sensors measured data at the locations shown in Fig. 2. These locations were selected to account for differences in the influence of the heating system, so that environmental variables other than air temperature could provide representative average values for the

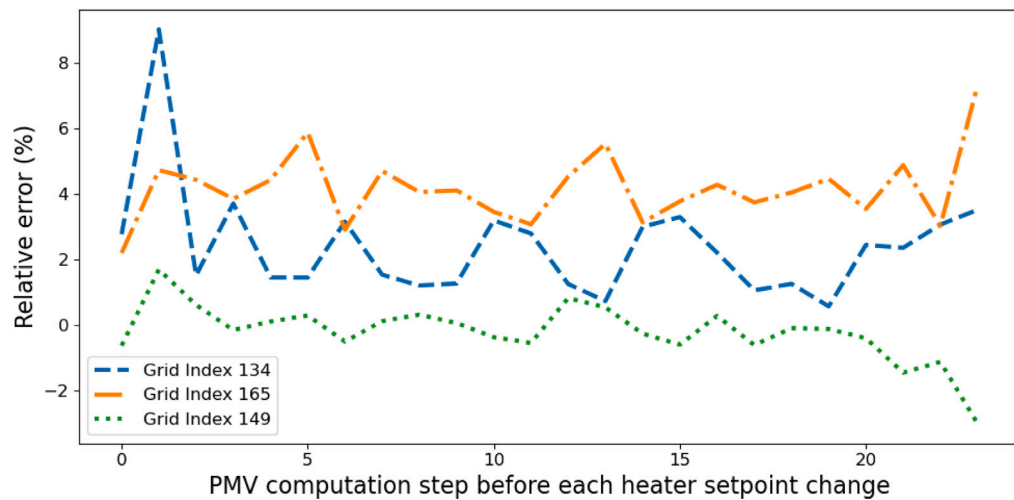


Fig. 10. Relative error between reconstructed and measured temperatures.

living lab space. During the experiments, the heating conditions ranged between approximately 22–26 °C. Under these conditions, the RH, obtained as the spatial average across the indoor environmental sensors and then averaged over the experimental period, was 20% (SD = 1.03%), and the corresponding V_{air} was 0.27 m/s (SD = 0.08 m/s). Using the environmental data obtained from the experiments and the reconstructed temperature data, the PMV distribution across the entire living lab was analyzed to examine differences in thermal comfort by location.

Fig. 11 presents an example of the spatial distribution of PMV values calculated based on the reconstructed indoor temperature data within the living lab space. Specifically, PMV was calculated using the reconstructed temperature field. Temperature data for computing PMV as shown in Figs. 11(a) and 11(b) are reconstructed after 10 min of heating with a setpoint temperature of 23 °C and 15 min of heating with a setpoint temperature of 26 °C, respectively. The environmental variables other than temperature for the PMV calculation were obtained from indoor sensor modules, with the module-averaged values at each measurement used as representative inputs, while personal factors were assumed to be 1.2 met and 0.5 clo. In these subfigures, thermally comfortable regions ($-0.5 < \text{PMV} < +0.5$) are indicated in white, while thermally uncomfortable zones ($\text{PMV} < -0.5$) are shown in increasingly darker shades of blue as the magnitude of discomfort (i.e., $|\text{PMV}|$) increases.

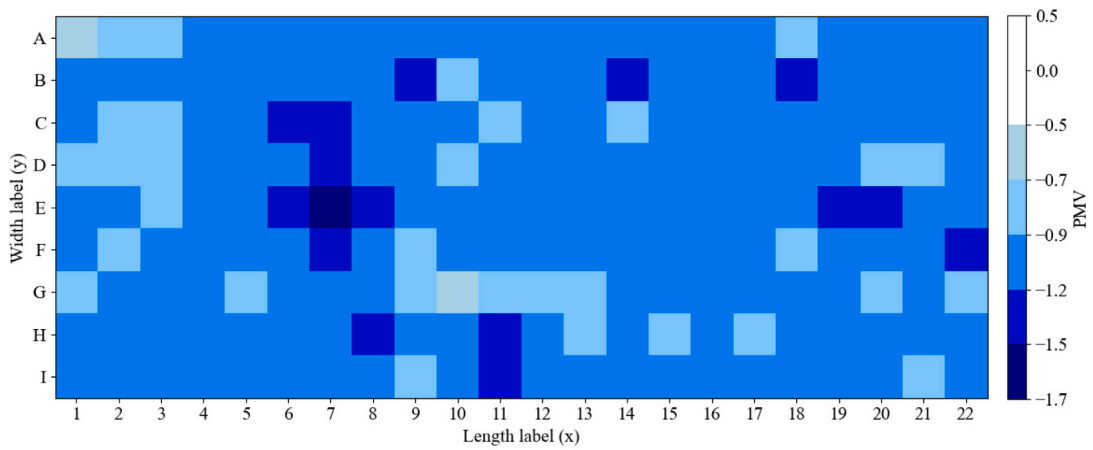
In Fig. 11(a), PMV values range from -1.59 to -0.62 , indicating a predominantly thermally uncomfortable environment under the 23 °C setpoint condition. In contrast, Fig. 11(b) shows a wider range from -1.06 to 0.20 , reflecting a general shift toward thermal comfort range. However, the PMV variation of up to 1.26 across the space highlights the strong influence of occupant location. These findings underscore the importance of accounting for spatial temperature distribution when estimating PMV in multi-occupant environments.

As shown in Fig. 11, the non-uniform distribution of room temperature leads to variations in PMV. This non-uniformity can be attributed to factors such as the location of the heating system supply outlet, and surrounding partitions and furniture (i.e., the overall layout of the space). The living lab itself is situated within a building and includes an automatic sliding door on the north wall. Although the door remained closed during the experiments, thermal variations arising from differences in insulation across walls and the door were evident, contributing to the observed non-uniform temperature patterns. In particular, the heating system in the living lab is located near the bottom-left corner of the space (around grid points H7, H8, I7, and I8), and the supply air outlet is a 4-way diffuser. The supply direction is fixed slightly upward so that airflow does not blow directly toward the occupants. As a result, uneven temperature distribution occurs, which is especially evident toward the upper and right region of the space. As a consequence of these factors, the overall temperature distribution remained non-uniform, leading to noticeable differences in PMV values, including areas close to the HVAC unit. These findings confirm the presence of spatial thermal non-uniformity, which can be effectively quantified using the data reconstruction algorithm. Temperature field reconstruction contributes to a more precise assessment of PMV distribution throughout the space.

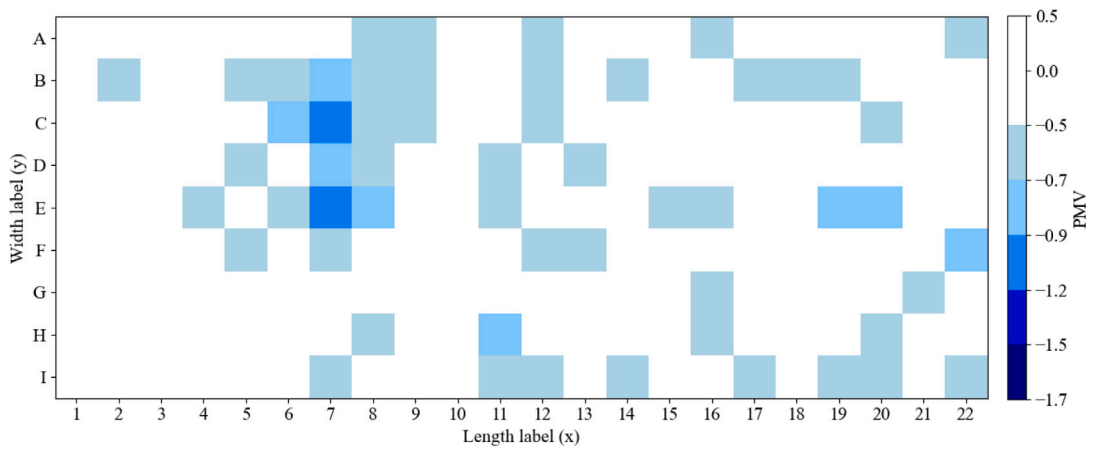
4.2.2. Experiment results of PMV estimation

During the thermal control experiments, the occupants were located at positions F5, G9, G16, and E19 in the living lab, based on the scenario-specific personal factor conditions described in Table 2. During the experiments, participants adjusted their activity levels and clothing types according to each scenario. Representative images illustrating these experimental conditions are provided in Fig. 12.

The estimation accuracy of personal factor for each occupant was evaluated using the mean absolute error (MAE) during the experimental procedure. The PF model estimates of metabolic rate and clothing insulation from image data at 2 FPS, storing the most frequent value every 20 s. As a result, three data points per minute were obtained. Over the total experiment duration (40 min



(a) After 10 minutes of heating with a setpoint of 23°C



(b) After 15 minutes of heating with a setpoint of 26°C

Fig. 11. Spatial distribution of PMV based on reconstructed indoor temperature fields.

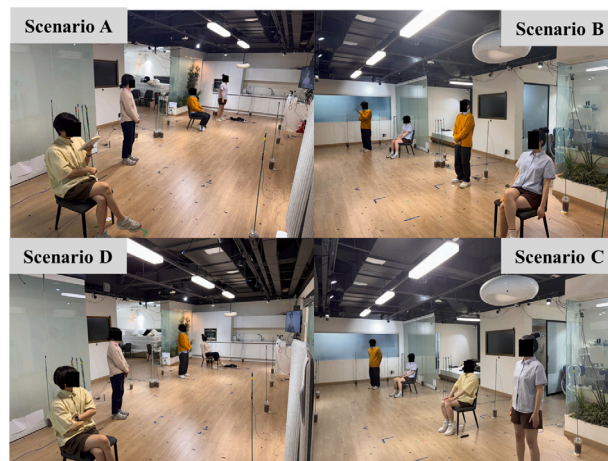


Fig. 12. Experimental images for each scenario.

Table 3

PMV values calculated after 10 min of a 23 °C preheating phase, averaged across three repeated experiments.

Scenario	Subject 1	Subject 2	Subject 3	Subject 4
A	-1.13	0.05	-0.07	-1.88
B	-1.44	0.31	-1.37	-0.11
C	-0.99	-0.78	-1.46	-0.03
D	-0.48	0.16	-0.16	-2.02

per scenario $\times$ 4 scenarios $\times$ 3 control points = 480 min), a total of 1440 personal factor data points were collected per participant. In this study, data from four participants yielded 5652 personal factor entries in total. Some missing values were observed due to sensor noise and communication errors. The overall MAE between the estimated and ground truth values was 0.011 met for metabolic rate and 0.029 clo for clothing insulation, consistent with the performance level reported in previous studies on the PF model [6].

Based on the estimated personal factors and the reconstructed indoor temperature data obtained during the experiment, average PMV values were compared for each participant, as summarized in Table 3. These values represent the mean PMV computed across three repeated experimental runs and were evaluated at the end of the 10-min preheating phase. Notable differences in PMV were observed depending on both personal factors and occupant location, even under the same controlled temperature condition of 23 °C during that period.

In Scenario A, all participants had different activity and clothing conditions, resulting in distinct PMV values for each individual. In Scenarios C and D, subjects 2 and 3 shared identical personal factor conditions. Specifically, in Scenario C, both subjects performed Case 1, which involved low activity and light clothing; however, their average PMV values differed by approximately 0.68 (−0.78 for subject 2 and −1.46 for subject 3). In Scenario D, where both subjects performed Case 4 with high activity and heavy clothing, the PMV difference was approximately 0.32. This outcome is consistent with the spatial distribution of PMV shown in Fig. 11, where Subject 2 (G9) exhibited a higher PMV than Subject 3 (G16), despite identical personal conditions. This discrepancy is attributed to differences in local thermal environments.

These results demonstrate that incorporating reconstructed temperature data and individual personal factors allows for precise PMV estimation at each occupant's location. The thermal control results based on these estimations are presented in the following section.

4.2.3. Thermal control results based on group PMV

This study implemented indoor thermal control using group PMV values representing the overall thermal comfort of multiple occupants. The corresponding setpoint decisions and indoor temperature changes are shown in Fig. 13. Dashed lines indicate the setpoint temperatures (T_{SET}) sent to the system at 10 min (T_{SET1}) and 25 min (T_{SET2}), while solid lines represent the average of real-time indoor temperature measurements under the control condition. This average was calculated using three sources: (1) two temperature measurements from the indoor sensor modules described in Section 3.1, (2) the 12 boundary measurement points used for temperature reconstruction as shown in Section 4.1.1, and (3) the 3 internal validation points introduced in Section 4.1.2 for evaluating reconstruction performance.

As shown in Fig. 4, the initial control period and the 10-min preheating at 23 °C were consistently applied. Nevertheless, in some cases, a noticeable increase in indoor temperature was observed during the early phase. This phenomenon was particularly observed in experiments conducted after 2 p.m. (i.e., Scenario A and B under PMV_{MEDIAN} and PMV_{WA} , and Scenario C). The elevated indoor temperatures compared to the setpoint are not attributed to errors in the control algorithm but rather to the thermal inertia of the building caused by heat accumulation from outdoor conditions.

In all scenarios, the setpoint temperature for PMV_{MEDIAN} was determined within the range of 25–26 °C, while PMV_{WA} consistently resulted in a fixed setpoint of 26 °C. In contrast, PMV_{MAD} produced a wider range of setpoint temperatures, spanning from 22 °C to 26 °C. In Scenarios A, B, and D, PMV_{MAD} produced a different setpoint temperature (T_{SET}) compared to the other two methods. In contrast, all three methods yielded the same setpoint of 26 °C in Scenario C.

Due to the thermal characteristics of the living lab where heat tends to accumulate within the experimental space a time lag was observed between the control input and the actual indoor temperature response, even when the HVAC system functioned as intended. The elevated indoor temperature persisted due to delayed cooling, as observed in Scenario A (PMV_{MAD}), where temperatures remained above the setpoint in the final phase. These findings highlight the importance of accounting for thermal inertia in real building control and suggest the need for adaptive control strategies in future applications.

Although there is a difference between the HVAC setpoint and the actual indoor temperature, PMV the control variable for the HVAC system was calculated using temperature fields reconstructed by the Gappy POD method. By incorporating the ambient temperature conditions experienced by occupants on-site, this approach ensures that thermal comfort is more accurately reflected. Furthermore, although absolute differences remain between setpoints and actual temperatures, the direction of change is consistent; when occupants feel too warm or too cold, the PMV-based control reliably indicates whether the temperature should be lowered or raised. Through continuous adjustment, the system can therefore converge to and maintain thermal comfort conditions.

Fig. 14 summarizes the control outcomes across all scenarios, presenting both individual PMV values and the corresponding group PMV values computed using three different aggregation methods. These results are shown at three key time points: First

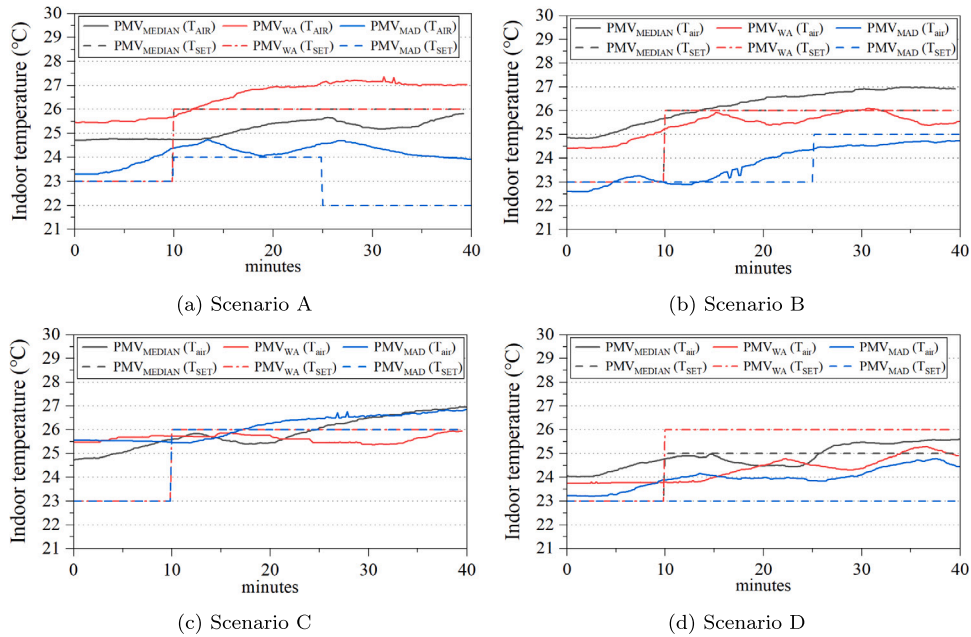


Fig. 13. Results of indoor temperature under group PMV-based control.

control (T_{SET1}), Second control (T_{SET2}), and End of control. The first control point refers to the moment after 10 min of preheating (23 °C) at the beginning of the thermal control period. The second control and end of control points correspond to PMV values recorded after 15 min of control at T_{SET1} and T_{SET2} , respectively.

As discussed in Section 4.2.2, the variation in PMV among participants differed across scenarios. At the first control (T_{SET1}), PMV_{MAD} generally produced the highest group PMV values, while PMV_{WA} consistently resulted in the lowest. The differences between the group PMV methods were pronounced depending on the scenario. This is primarily because PMV_{WA} is designed to assign greater weight to participants experiencing thermal discomfort. For example, in Scenarios B and D, where participants experienced greater thermal discomfort, PMV_{WA} strongly reflected these extremes. In contrast, in Scenario C, where individual PMV values showed minimal variation, the differences among all methods were less pronounced. Additionally, greater deviation in individual PMVs led to larger differences between PMV_{MAD} and PMV_{MEDIAN} , highlighting the influence of participant variability on group-level estimates.

Across all control points, the group PMV values for each method gradually approached the thermal comfort range (± 0.5) as the experiment progressed. For example, in Scenario A, PMV_{WA} improved from -0.97 at the first control (T_{SET1}) to -0.11 by the end of the experiment. Similar improvements were observed for PMV_{MEDIAN} and PMV_{MAD} as shown in Fig. 14. These results indicate that adjusting the setpoint temperature using group PMV as a control variable effectively enhanced overall thermal comfort.

Individual PMV responses varied depending on the group PMV method. PMV_{WA} effectively improved comfort for participants with extremely low PMV values—for example, Subject 4 in Scenario A showed improvement in PMV from -1.7 to -0.9 . Although this occasionally led to increased discomfort for others (e.g., Subject 2, whose PMV increased from 0.18 to 0.6), it resulted in a tighter distribution around PMV 0 overall. In contrast, PMV_{MEDIAN} was less sensitive to outliers but could overlook extreme discomfort in participants with high variability, such as in Scenario D. In Scenario D, Subject 4 showed only limited improvement in comfort under the PMV_{MEDIAN} method, with a PMV increase of just 0.19 (from -1.74 to -1.55), whereas PMV_{WA} resulted in a considerably larger improvement of 0.76 (from -2.28 to -1.49). PMV_{MAD} , which adjusts for deviation, generally produced higher group PMVs when participant differences were large. However, it sometimes led to overheating, especially when indoor temperatures rose excessively.

These results highlight that different group PMV methods can lead to significantly different control decisions, even under identical occupancy conditions. To further evaluate the effectiveness of these control strategies, participants' subjective thermal sensations were analyzed across different scenarios. Across 12 experiments, minor issues such as mobile communication errors and missed responses led to an average of 227 TSV entries per participant, totaling 909 responses. Of these, 697 TSV entries collected after control actions were used for analysis. Fig. 15 shows the proportion by TSV level.

Perceived TSV results varied by group PMV method, even under the same environmental conditions. Higher indoor temperatures generally led to higher TSV scores, with Scenario A under PMV_{WA} showing the highest temperature (26.8 °C) and the greatest discomfort. In this case, most participants except Subject 4, who had the lowest activity and clothing insulation reported TSV values above 1.5, indicating a warm sensation. As shown in Fig. 15(a), the proportion of TSV = 2 responses was also highest for PMV_{WA} .

In Scenario B, participants with similar personal factors (Subjects 1 and 3) reported high thermal comfort (TSV = 0) under warmer conditions (PMV_{MEDIAN} , PMV_{WA}), while Subjects 2 and 4 experienced greater discomfort (TSV = 1 or 2) due to contrasting

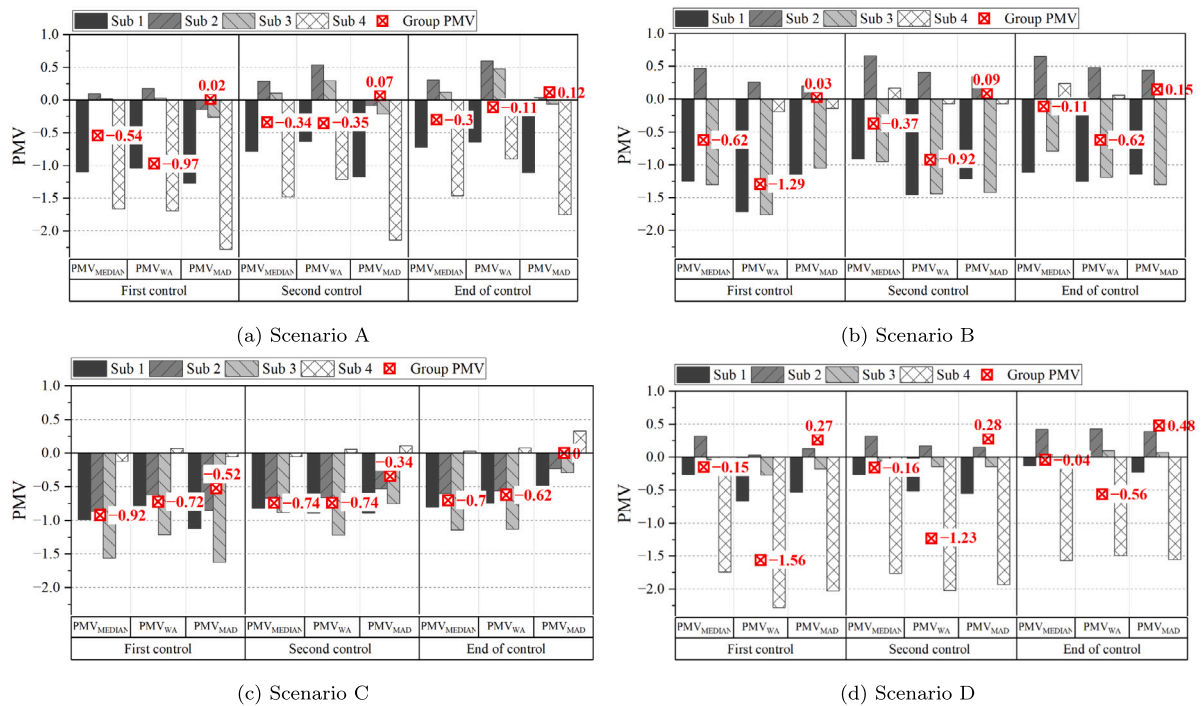


Fig. 14. Control variable estimation results under different experimental scenarios.

Table 4

Occupants thermal satisfaction ratio by thermal comfort range.

Comfort range	TSV = 0			$-1 \leq \text{TSV} \leq +1$		
	PMV _{MEDIAN}	PMV _{WA}	PMV _{MAD}	PMV _{MEDIAN}	PMV _{WA}	PMV _{MAD}
Scenario A	67.5%	30.0%	63.6%	85.0%	57.5%	100%
Scenario B	65.0%	70.5%	31.8%	97.5%	95.5%	93.2%
Scenario C	54.5%	58.9%	61.4%	86.4%	93.2%	100%
Scenario D	52.3%	65.0%	54.5%	95.5%	100%	100%
Total mean	59.8%	56.1%	52.8%	91.1%	86.5%	98.3%

personal factor conditions. The lower indoor temperature under PMV_{MAD} resulted in a higher proportion of TSV = −1 responses, indicating that all subjects experienced discomfort. In Scenarios C and D, despite comparable average indoor temperatures across the three control methods, variations in TSVs were observed. These differences are attributed to localized environmental factors, such as late-stage temperature increases and occupant positioning within the space. For example, in Scenario D, Subject 3 who is positioned near the heater, reported higher thermal sensation than Subject 2, despite identical personal conditions. These findings emphasize that spatial factors like occupant position significantly affect perceived comfort.

To assess the control performance of each group PMV method, participant TSV responses were analyzed using two comfort ranges, as shown in Table 4: complete neutrality (TSV = 0) and broader comfort with low thermal discomfort ($-1 \leq \text{TSV} \leq +1$).

Based on the analysis, the proportion of participants reporting complete neutrality (TSV = 0) followed the order PMV_{MEDIAN} (62.2%) > PMV_{WA} (58.0%) > PMV_{MAD} (52.8%). None of the methods exceeded 70% on average, highlighting limitations in achieving universal thermal comfort across all occupants. Among the three approaches, the PMV_{MEDIAN} method exhibited the most consistent performance across varying personal factor conditions. In contrast, PMV_{WA} and PMV_{MAD} showed lower TSV = 0 proportions in scenarios with high inter-individual variation, indicating reduced effectiveness in settings with diverse occupant profiles.

Analysis of the broader comfort range ($-1 \leq \text{TSV} \leq +1$) revealed that all three group PMV estimation methods achieved average thermal comfort rates exceeding 86.5%, with PMV_{MAD} showing the highest performance at 98.4%. In comparison to the PMV values presented in Fig. 14, the actual TSV responses exhibited a tendency toward higher perceived thermal sensation. For instance, PMV values between −2 and −1 often corresponded to TSV = −1, while PMV > +0.5 frequently aligned with TSV = +2. This indicates a positive perceptual bias, where subjective thermal sensation is higher than objective PMV estimates. The PMV_{MAD} method helped reduce this discrepancy by accounting for distributional deviations. While PMV_{MEDIAN} was less sensitive to individual extremes, it provided consistent control across scenarios. In contrast, PMV_{WA} effectively addressed extreme discomfort but showed greater variability, particularly under various occupant conditions.

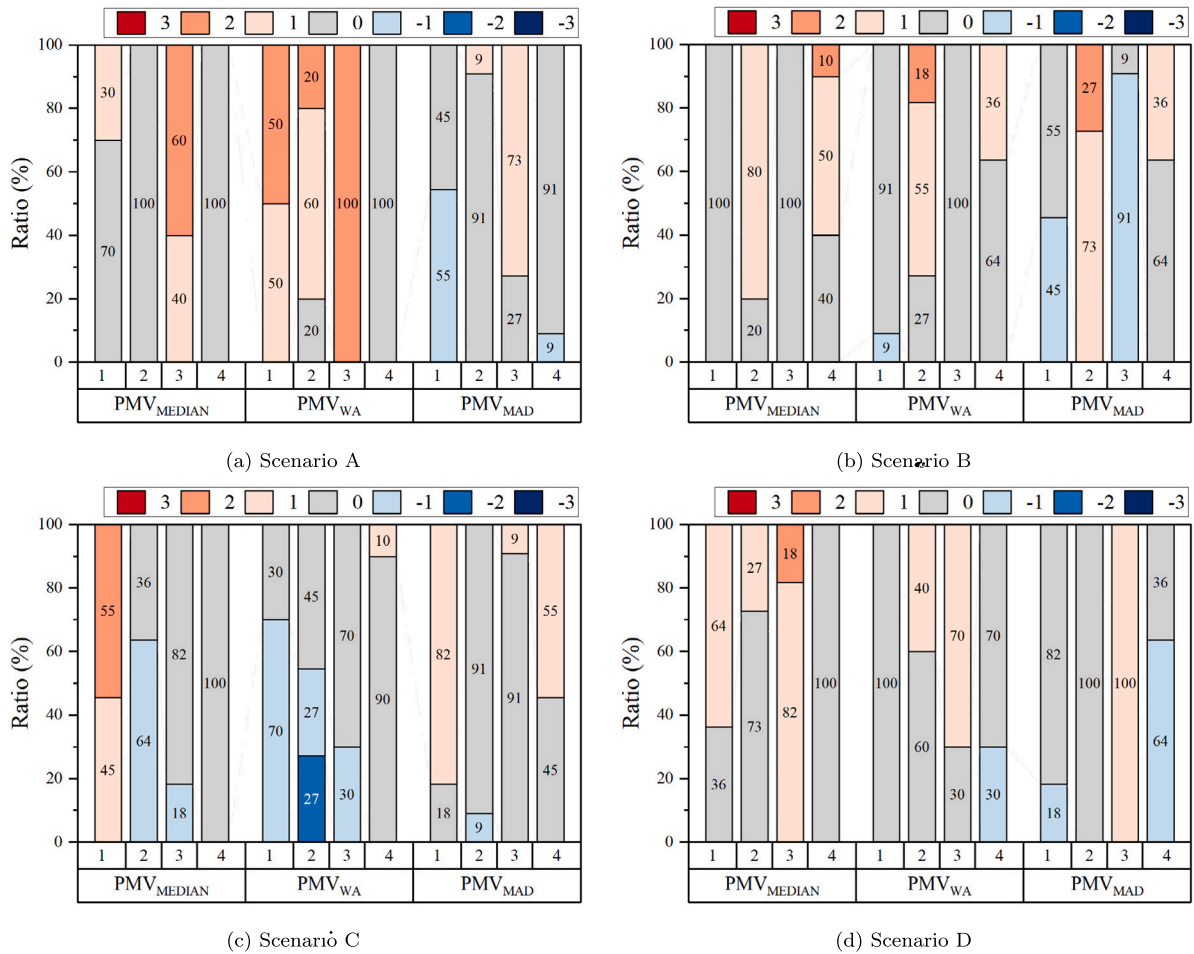


Fig. 15. Distribution of TSV and proportional comfort levels under group PMV-based control methods.

On average, PMV_{MAD} was the most effective in satisfying the thermal comfort of multiple occupants under various scenarios. Nevertheless, each group PMV estimation method presents distinct trade-offs that must be considered when designing control strategies. PMV_{MEDIAN} offers stable and balanced performance across varied conditions, making it well-suited for general comfort control. PMV_{MAD} is effective in mitigating extreme discomfort by incorporating variability, though it may lower the proportion of neutral responses. PMV_{WA} enhances comfort for highly dissatisfied occupants but exhibits inconsistent results at the group level. Therefore, the selection of an appropriate group PMV method should be based on the specific context of application, including the space's function, occupant characteristics, and control objectives.

4.3. Discussion

The main innovation of this study lies in proposing a novel methodology that combines a gappy data reconstruction algorithm, Gappy POD, with an occupant-information recognition model to estimate actual thermal comfort. Recent studies have made extensive efforts to improve the accuracy of thermal comfort prediction, with most focusing on developing equation-based or AI-driven models for real-time estimation of occupant personal factors [6,16,39]. In contrast, this study introduces an integrated thermal comfort estimation framework that incorporates reconstructed temperature fields with individual parameters such as metabolic rate and clothing insulation. The proposed framework enables personalized thermal comfort assessment by comprehensively reflecting occupant location and state using only a limited number of sensor measurements. The results revealed that spatial variations in the reconstructed PMV values reached up to 1.26 scale units, highlighting the need for more localized control strategies beyond zone-averaged approaches.

Furthermore, while many previous studies have focused on thermal control within single-occupant spaces [15,16,40], actual buildings typically involve multiple occupants with diverse comfort needs. Considering this, the present study proposed three statistical approaches for deriving a single control variable applicable to multi-occupancy environments. Unlike previous studies that conducted limited evaluations of thermal comfort in multi-occupant settings, often lacking experimental validation [6,41], this

study experimentally evaluated multiple statistical methodologies under various occupancy scenarios. By employing group-based PMV metrics and analyzing the results experimentally, the study demonstrated that a representative comfort index can be effectively derived and applied across different occupancy conditions.

Since the experiment was conducted under predefined and controlled scenarios, it may not fully capture the dynamic and unpredictable nature of real-world occupancy. To enhance applicability in actual buildings, the proposed framework should be integrated with real-time occupant detection and location-tracking technologies, in addition to the environmental reconstruction algorithm and personal-factor model. Systematic collection of thermal data across space under varying HVAC operating modes and occupancy conditions will support the development of more robust and generalizable models for real-world applications. Moreover, privacy concerns associated with image-based AI models should be carefully addressed to ensure feasible and ethical deployment in large-scale smart-building environments.

Despite the contributions of this study, several technical limitations were identified, highlighting important directions for future research. First, regarding the data reconstruction algorithm, the effect of sensor placement on reconstruction performance was not investigated. This is because sensor placement was assumed, in which sensors were evenly distributed along each boundary to reflect practical constraints related to building design and occupant convenience. In addition, the number of sensors used in the living lab (12 sensors in a 3.2 m × 9.45 m space) is relatively large. Further research is needed to achieve high reconstruction accuracy with fewer sensors. While Gappy POD is effective for reconstructing quantities with minimal advection effects, its performance may degrade in scenarios where strong internal airflow introduces dominant advection dynamics. Moreover, the reconstruction accuracy was validated primarily under HVAC setpoint variations, while other influencing factors such as boundary conditions, heat sources, and detailed HVAC operating settings were not explicitly tested. Although the use of real experimental data implicitly incorporated some of these variations, the reconstruction capability of Gappy POD under systematically altered conditions remains uncertain.

To overcome these limitations, several future directions are planned. Specifically, the influence of sensor placement on reconstruction accuracy will be quantitatively evaluated. Optimization algorithms such as the Discrete Empirical Interpolation Method (DEIM) [42–44] and S-OPT [45,46] will be employed for sensor placement, and their reconstruction performance will be systematically compared. DEIM is a greedy sampling algorithm that selects interpolation points to minimize the upper bound of reconstruction error. In addition to the original DEIM formulation, an oversampled variant [46,47] is considered, where the number of sampling points is allowed to exceed the number of POD modes. S-OPT is a quasi-optimal sampling method that seeks to maximize the orthogonality of the sampled POD basis vectors and the determinant of the corresponding Gram matrix constructed from them, thereby enhancing numerical stability and reducing reconstruction error.

Furthermore, the application of Gappy Autoencoder (Gappy AE) [48], a nonlinear manifold-based reconstruction algorithm, is planned to be explored. Gappy AE offers the potential to achieve high reconstruction accuracy with fewer sensors and is particularly suitable for reconstructing environmental quantities governed by advection-dominant dynamics. In particular, Gappy AE is expected to perform well in reconstructing airflow velocity fields, which typically exhibit strong advection and complex flow structures. In this way, techniques need to be developed for reconstructing not only indoor temperature but also the distributions of other indoor environmental variables such as RH and V_{AIR} . In addition, to enhance the model's ability to generalize for application to diverse building environments, it will be necessary to acquire simulation-based datasets that explicitly vary key environmental factors such as boundary conditions, heat sources, and HVAC configurations. Such datasets will enable a more systematic evaluation of reconstruction accuracy under diverse scenarios and provide the basis for enhancing the model's adaptability to unseen conditions while improving its robustness across different environments.

In terms of thermal comfort assessment, various models have been proposed, each with inherent limitations in predictive accuracy under real building conditions. In this study, the PMV model was adopted as one indicator of thermal comfort. Although it is known to have limitations in accuracy, as noted by Tartarini and Schiavon [49], efforts were made to enhance the reliability of its predictions by improving the accuracy of the input parameters [13]. Nevertheless, from a long-term perspective, the development of alternative approaches that can provide more precise evaluations of thermal comfort beyond the PMV model remains an important direction for future research.

The control experiment also had several limitations. Due to the structural characteristics of the experimental space and the aging HVAC system, precise temperature control was constrained, and the setpoint range was limited to 22–26 °C. Additionally, some data loss occurred due to sensor noise, communication delays, and storage errors. However, to minimize the impact on control performance, data collected during the two minutes preceding each control point were preprocessed using mean and most frequent values. Furthermore, the experiment was conducted under predefined and controlled scenarios, which did not fully reflect the dynamic and unpredictable nature of real-world occupancy. To enhance applicability in actual environments, the integration of real-time occupant detection technologies is necessary, in addition to the environmental reconstruction algorithm and the PF model proposed in this study.

This study was conducted during the heating season with a limited number of participants. Therefore, future research should involve long-term data collection encompassing a wider range of seasonal conditions and more diverse occupancy patterns. Lastly, a method for determining a control variable that represents group thermal comfort and dynamically integrates real-time information such as occupancy status, number of occupants, metabolic rate, and clothing insulation—is essential. Such capabilities would advance the system from scenario-based control to an adaptive, user-centered thermal management framework capable of responding effectively to real-world environmental changes.

5. Conclusion

This study proposes a PMV estimation approach that utilizes an indoor temperature reconstruction algorithm based on Gappy POD to account for occupant location. Spatial variations in PMV across multi-occupant conditions were quantitatively analyzed, and a group PMV-based control framework was applied to evaluate thermal control performance. The Gappy POD method reconstructs real-time indoor temperature distributions using a limited number of sensor measurements. This approach enables fast and efficient reconstruction by leveraging a pre-trained POD basis. Also, it supports accurate estimation of spatially varying PMV values. Experimental results from the living lab indicate that the average relative error in temperature reconstruction remained below 3%, demonstrating the method's feasibility and reliability.

Thermal comfort analysis based on the reconstructed indoor temperature data in the living lab revealed that spatial variations in PMV across the space can reach up to 1.26 scale units. By integrating reconstructed environmental fields with occupant-specific personal factors, accurate location-based PMV values were estimated and applied to group PMV-based control strategies. The results demonstrated that each group PMV method produced distinct thermal comfort outcomes depending on the occupancy scenario. When thermal satisfaction was assessed by the proportion of occupants within the comfort range ($TSV \leq \pm 1$), the MAD-based method which accounts for distributional variability proved more effective in enhancing thermal comfort during the heating season.

In summary, the findings of this study highlight the feasibility of real-time thermal comfort estimation and control based on occupant location, even within spatially heterogeneous indoor environments. In addition, the results underscore the importance of selecting appropriate group PMV metrics in accordance with occupant diversity, spatial configurations, and control priorities of the HVAC system. Furthermore, the study emphasizes the necessity of thermal control strategies that integrate both spatial and individual variability to ensure optimal comfort in shared environments.

CRediT authorship contribution statement

Youngkyu Kim: Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Conceptualization. **Byounghyun Yoo:** Validation, Supervision, Funding acquisition, Formal analysis. **Ji Young Yun:** Resources, Investigation, Data curation. **Hyeokmin Lee:** Resources, Investigation, Data curation. **Sehyeon Park:** Resources, Investigation, Data curation. **Jin Woo Moon:** Supervision, Funding acquisition. **Eun Ji Choi:** Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Funding acquisition, Conceptualization.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used OpenAI's tool ChatGPT in order to improve readability and language. After using this tool/service, the authors carefully reviewed and edited the content as needed and take full responsibility for the content of the published article.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2023-00217322, RS-2023-00276529, and NRF-2021R1A2C2093065).

This study was reviewed and approved by the Institutional Review Board of Chung-Ang University (IRB No. 1041078-20241114-HR-326).

Data availability

Data will be made available on request.

References

- [1] Z. Afroz, G.M. Shafiqullah, T. Urmee, M.A. Shueb, G. Higgins, Predictive modelling and optimization of HVAC systems using neural network and particle swarm optimization algorithm, *Build. Environ.* 209 (2022) 108681, <http://dx.doi.org/10.1016/j.buildenv.2021.108681>.
- [2] S. Altomonte, S. Schiavon, M.G. Kent, G. Brager, Indoor environmental quality and occupant satisfaction in green-certified buildings, *Build. Res. Inf.* 47 (3) (2019) 255–274, <http://dx.doi.org/10.1080/09613218.2018.1383715>.
- [3] L. Magnier, F. Haghighat, Multiobjective optimization of building design using trnsys simulations, genetic algorithm, and artificial neural network, *Build. Environ.* 45 (3) (2010) 739–746, <http://dx.doi.org/10.1016/j.buildenv.2009.08.016>.
- [4] F. Jazizadeh, A. Ghahramani, B. Becerik-Gerber, T. Kichkaylo, M. Orosz, Human-building interaction framework for personalized thermal comfort-driven systems in office buildings, *J. Comput. Civ. Eng.* 28 (1) (2014) 2–16, [http://dx.doi.org/10.1061/\(ASCE\)CP.1943-5487.000030](http://dx.doi.org/10.1061/(ASCE)CP.1943-5487.000030).
- [5] C. Zhong, J.-H. Choi, Development of a data-driven approach for human-based environmental control, *Procedia Eng.* 205 (2017) 1665–1671, <http://dx.doi.org/10.1016/j.proeng.2017.10.341>.
- [6] E.J. Choi, J.Y. Yun, Y.J. Choi, M.C. Seo, J.W. Moon, Impact of thermal control by real-time PMV using estimated occupants personal factors of metabolic rate and clothing insulation, *Energy Build.* 307 (2024) 113976, <http://dx.doi.org/10.1016/j.enbuild.2024.113976>.
- [7] P.O. Fanger, *Thermal Comfort. Analysis and Applications in Environmental Engineering*, Danish Technical Press., Copenhagen, 1970, p. 244.
- [8] A.P. Gagge, A. Fobelets, L. Berglund, A standard predictive index of human response to the thermal environment, *Ashrae Trans.* 92 (1986) 709–731.
- [9] P.O. Fanger, J. Toftum, Extension of the PMV model to non-air-conditioned buildings in warm climates, *Energy Build.* 34 (6) (2002) 533–536, [http://dx.doi.org/10.1016/S0378-7788\(02\)00003-8](http://dx.doi.org/10.1016/S0378-7788(02)00003-8).
- [10] R. De Dear, G.S. Brager, Developing an adaptive model of thermal comfort and preference, *ASHRAE Trans.* 104 (1) (1998) 67–75.
- [11] J. Ngarambe, G.Y. Yun, M. Santamouris, The use of artificial intelligence (AI) methods in the prediction of thermal comfort in buildings: Energy implications of AI-based thermal comfort controls, *Energy Build.* 211 (2020) 109807, <http://dx.doi.org/10.1016/j.enbuild.2020.109807>.
- [12] *ASHRAE Standard 55, Thermal Environmental Conditions for Human Occupancy*, American Society of Heating, Refrigerating and Air-Conditioning Engineers, Atlanta, GA, USA, 2020.
- [13] F.R.d. Alfano, B.I. Palella, G. Riccio, The role of measurement accuracy on the thermal environment assessment by means of PMV index, *Build. Environ.* 46 (7) (2011) 1361–1369, <http://dx.doi.org/10.1016/j.buildenv.2011.01.001>.
- [14] W.T. Sung, S.J. Hsiao, The application of thermal comfort control based on smart house system of IoT, *Meas.* 149 (2020) 106997, <http://dx.doi.org/10.1016/j.measurement.2019.106997>.
- [15] E.J. Choi, Y.J. Choi, N.H. Kim, J.W. Moon, Seasonal effects of thermal comfort control considering real-time clothing insulation with vision-based model, *Build. Environ.* 235 (1) (2023) 110255, <http://dx.doi.org/10.1016/j.buildenv.2023.110255>.
- [16] H. Choi, B. Jeong, J. Lee, H. Na, K. Kang, T. Kim, Deep-vision-based metabolic rate and clothing insulation estimation for occupant-centric control, *Build. Environ.* 221 (2022) 109345, <http://dx.doi.org/10.1016/j.buildenv.2022.109345>.
- [17] J. Liu, I.W. Foged, T.B. Moeslund, Automatic estimation of clothing insulation rate and metabolic rate for dynamic thermal comfort assessment, *Pattern Anal. Appl.* 25 (3) (2022) 619–634, <http://dx.doi.org/10.1007/s10044-021-00961-5>.
- [18] J.Y. Yun, E.J. Choi, M.H. Chung, T.W. Kim, J.W. Moon, Development and performance evaluation of an indoor thermal environment control algorithm incorporating MET estimation model with object detection, *Build. Environ.* 267 (2025) 112217, <http://dx.doi.org/10.1016/j.buildenv.2024.112217>.
- [19] K. Lee, H. Choi, H. Kim, D.D. Kim, T. Kim, Assessment of a real-time prediction method for high clothing thermal insulation using a thermoregulation model and an infrared camera, *Atmosphere* 11 (1) (2020) 106, <http://dx.doi.org/10.3390/atmos11010106>.
- [20] J. Dziedzic, D. Yan, V. Novakovic, Measurement of Dynamic Clothing Factor (D-CLO), NTNU, Trondheim, 2018.
- [21] X. Shan, W.-Z. Lu, An integrated approach to evaluate thermal comfort in air-conditioned large-space office, *Sci. Technol. Built Environ.* 27 (4) (2020) 436–450, <http://dx.doi.org/10.1080/23744731.2020.1796420>.
- [22] E.J. Choi, Y. Kim, J.W. Moon, B. Yoo, Improving the accuracy of spatial thermal comfort assessment through gappy data reconstruction, *Korean J. Comput. Des. Eng.* 29 (4) (2024) 506–520.
- [23] J.A.O. García, A review of general and local thermal comfort models for controlling indoor ambiances, *Air Qual.* (2010) 309–326.
- [24] A. Brunello, A. Urgolo, F. Pittino, A. Montvay, A. Montanari, Virtual sensing and sensors selection for efficient temperature monitoring in indoor environments, *Sens.* 21 (8) (2021) 2728, <http://dx.doi.org/10.3390/s21082728>.
- [25] B. Yang, Z. Lv, F. Wang, Digital twins for intelligent green buildings, *Build.* 12 (6) (2022) 856, <http://dx.doi.org/10.3390/buildings12060856>.
- [26] V.A. Arowoiya, R.C. Moehler, Y. Fang, Digital twin technology for thermal comfort and energy efficiency in buildings: A state-of-the-art and future directions, *Energy Build. Environ.* 5 (5) (2024) 641–656, <http://dx.doi.org/10.1016/j.enbenv.2023.05.004>.
- [27] A. Clausen, K. Arendt, A. Johansen, F.C. Sangogboye, M.B. Kjaergaard, C.T. Veje, B.N. Jørgensen, A digital twin framework for improving energy efficiency and occupant comfort in public and commercial buildings, *Energy Inform.* 4 (2021) 1–19, <http://dx.doi.org/10.1186/s42162-021-00153-9>.
- [28] R. Molinaro, J.-S. Singh, S. Catsoulis, C. Narayanan, D. Lakehal, Embedding data analytics and CFD into the digital twin concept, *Comput. & Fluids* 214 (2021) 104759, <http://dx.doi.org/10.1016/j.compfluid.2020.104759>.
- [29] A. Arsiwala, F. Elghaish, M. Zoher, Digital twin with machine learning for predictive monitoring of CO2 equivalent from existing buildings, *Energy Build.* 284 (2023) 112851, <http://dx.doi.org/10.1016/j.enbuild.2023.112851>.
- [30] J.C. Cheng, J. Zhang, H.H. Kwok, J.C. Tong, Thermal performance improvement for residential heritage building preservation based on digital twins, *J. Build. Eng.* 82 (2024) 108283, <http://dx.doi.org/10.1016/j.jobbe.2023.108283>.
- [31] A. Liguori, R. Markovic, T.T.H. Dam, J. Frisch, C. van Treeck, F. Causone, Indoor environment data time-series reconstruction using autoencoder neural networks, *Build. Environ.* 191 (2021) 107623, <http://dx.doi.org/10.1016/j.buildenv.2021.107623>.
- [32] F. Zhao, S.H. Lee, G. Augenbroe, Reconstructing building stock to replicate energy consumption data, *Energy Build.* 117 (2016) 301–312, <http://dx.doi.org/10.1016/j.enbuild.2015.10.001>.
- [33] R. Everson, L. Sirovich, Karhunen–Loève procedure for gappy data, *J. Opt. Soc. Amer. A* 12 (8) (1995) 1657–1664, <http://dx.doi.org/10.1364/JOSAA.12.001657>.
- [34] T. Bui-Thanh, M. Damodaran, K. Willcox, Aerodynamic data reconstruction and inverse design using proper orthogonal decomposition, *AIAA J.* 42 (8) (2004) 1505–1516, <http://dx.doi.org/10.2514/1.2159>.
- [35] K. Willcox, Unsteady flow sensing and estimation via the gappy proper orthogonal decomposition, *Comput. & Fluids* 35 (2) (2006) 208–226, <http://dx.doi.org/10.1016/j.compfluid.2004.11.006>.
- [36] E.J. Choi, J.Y. Yun, J.Y. Hyun, J.W. Moon, PMV estimation methods for thermal control in multi-occupant buildings, *KIEAE J.* 25 (2) (2025) 73–80.
- [37] R.J.D. Dear, G.S. Brager, Thermal comfort in naturally ventilated buildings: revisions to ASHRAE Standard 55, *Energy Build.* 34 (6) (2002) 549–561, [http://dx.doi.org/10.1016/S0378-7788\(02\)00005-1](http://dx.doi.org/10.1016/S0378-7788(02)00005-1).
- [38] R. De Dear, G. Schiller Brager, The adaptive model of thermal comfort and energy conservation in the built environment, *Int. J. Biometeorol.* 45 (2001) 100–108, <http://dx.doi.org/10.1007/s004840100093>.
- [39] K. Watanabe, H.B. Rijal, T. Nakaya, Investigation of clothing insulation and thermal comfort in Japanese houses, in: *PLEA*, 2013.
- [40] S. Liu, S. Schiavon, H.P. Das, M. Jin, C.J. Spanos, Personal thermal comfort models with wearable sensors, *Build. Environ.* 162 (2019) 106281.

- [41] F. Topak, G.S. Pavlak, M.K. Pekerikli, J. Wang, F. Jazizadeh, Collective comfort optimization in multi-occupancy environments by leveraging personal comfort models and thermal distribution patterns, *Build. Environ.* 239 (2023) 110401, <http://dx.doi.org/10.1016/j.buildenv.2023.110401>, URL <https://www.sciencedirect.com/science/article/pii/S0360132323004286>.
- [42] S. Chaturantabut, D.C. Sorensen, Nonlinear model reduction via discrete empirical interpolation, *SIAM J. Sci. Comput.* 32 (5) (2010) 2737–2764, <http://dx.doi.org/10.1137/090766498>.
- [43] Z. Drmac, S. Gugercin, A new selection operator for the discrete empirical interpolation method—Improved a priori error bound and extensions, *SIAM J. Sci. Comput.* 38 (2) (2016) A631–A648, <http://dx.doi.org/10.1137/15M1019271>.
- [44] Z. Drmac, A.K. Saibaba, The discrete empirical interpolation method: Canonical structure and formulation in weighted inner product spaces, *SIAM J. Matrix Anal. Appl.* 39 (3) (2018) 1152–1180, <http://dx.doi.org/10.1137/17M1129635>.
- [45] Y. Shin, D. Xiu, Nonadaptive quasi-optimal points selection for least squares linear regression, *SIAM J. Sci. Comput.* 38 (1) (2016) A385–A411, <http://dx.doi.org/10.1137/15M1015868>.
- [46] J.T. Lauzon, S.W. Cheung, Y. Shin, Y. Choi, D.M. Copeland, K. Huynh, S-OPT: A points selection algorithm for hyper-reduction in reduced order models, 2022, <http://dx.doi.org/10.48550/arXiv.2203.16494>, arXiv:2203.16494.
- [47] K. Carlberg, C. Bou-Mosleh, C. Farhat, Efficient non-linear model reduction via a least-squares Petrov–Galerkin projection and compressive tensor approximations, *Internat. J. Numer. Methods Engrg.* 86 (2) (2011) 155–181, <http://dx.doi.org/10.1002/nme.3050>.
- [48] Y. Kim, Y. Choi, B. Yoo, Gappy AE: A nonlinear approach for Gappy data reconstruction using auto-encoder, *Comput. Methods Appl. Mech. Engrg.* 426 (2024) 116978, <http://dx.doi.org/10.1016/j.cma.2024.116978>.
- [49] F. Tartarini, S. Schiavon, Comparative analysis of PMV models accuracy implemented in the ISO 7730: 2005 and ASHRAE 55: 2023, *Build. Environ.* 275 (2025) 112766.