

Routledge
Taylor & Francis Group

Journal of Asian Architecture and Building Engineering

ISSN: 1346-7581 (Print) 1347-2852 (Online) Journal homepage: www.tandfonline.com/journals/tabe20

Compensating errors in building energy simulation calibration: quantifying data-driven behavioral uncertainty

Basak Gucyeter

To cite this article: Basak Gucyeter (27 Feb 2026): Compensating errors in building energy simulation calibration: quantifying data-driven behavioral uncertainty, Journal of Asian Architecture and Building Engineering, DOI: 10.1080/13467581.2026.2631876

To link to this article: https://doi.org/10.1080/13467581.2026.2631876

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group on behalf of the Architectural Institute of Japan, Architectural Institute of Korea and Architectural Society of China.

Published online: 27 Feb 2026.

Submit your article to this journal

Article views: 39

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tabe20

JOURNAL OF ASIAN ARCHITECTURE AND BUILDING ENGINEERING
https://doi.org/10.1080/13467581.2026.2631876

Taylor & Francis
Taylor & Francis Group

ENVIRONMENTAL ENGINEERING

OPEN ACCESS Check for updates

Compensating errors in building energy simulation calibration: quantifying data-driven behavioral uncertainty

Basak Gucyeter

Department of Architecture, Eskisehir Osmangazi University, Eskisehir, Turkey

ABSTRACT

Building energy performance gap is partly attributed to simplified occupant behavior representations in simulations. Conventional calibration approaches satisfy statistical criteria but mask inconsistent physical assumptions through compensating errors. This study examines how stepwise integration of measured occupant behavior alters calibration outcomes, exposing compensating errors in conventionally calibrated models. Indoor temperature, humidity, presence, lighting, window-opening behavior, and weather were collected during a one-year monitoring campaign in an educational building. An iterative calibration framework was implemented using EDSL Tas (Environmental Design Solutions Ltd. TAS), progressing from an uncalibrated baseline to a statistically calibrated reference model and subsequently to behavior-integrated iterations. Although the conventionally calibrated model achieved strong statistical agreement with measured heating energy (NMBE = 1.39%), replacing standard schedules with measured presence increased predicted heating demand (3.96–4.19%). This revealed compensating errors associated with overestimated internal gains and constrained envelope parameters. Adaptive window-opening behavior shifted disagreement to 7.05% due to ventilation losses. Measured lighting use acted as a corrective internal gain, partially offsetting heating demand. Findings demonstrate that compliance with calibration benchmarks alone is insufficient to ensure physically accurate models. Monitored occupant behavior is essential for identifying compensating errors and establishing physically consistent baseline models suitable for evaluating energy conservation measures in existing buildings.

ARTICLE HISTORY
Received 4 December 2025
Accepted 11 February 2026

KEYWORDS
Compensating error; building energy simulation; manual calibration; occupant behavior

1. Introduction

Building energy performance (BEP) has been a key research area since the 1970s oil crisis, focusing on mitigating climate change, reducing energy consumption, and improving building efficiency (Copiello 2017; Kim, Kim, and Song 2021). The built environment accounts for 34–41% of total energy consumption and CO_2 emissions in advanced economies (Kim, Kim, and Song 2021; Uddin et al. 2021). BEP remains central in policy agendas, especially for the EU's 2050 net-zero objectives (Chen et al. 2024). Strategies for reducing energy consumption in existing buildings include retrofitting and passive measures (Buso et al. 2015; Khabir and Vakilinezhad 2023). However, a persistent challenge is the discrepancy between predicted and actual operational energy use, documented since the mid-1990s, and known as the building energy performance gap (BEPG) (Bahadori-Jahromi et al. 2022; Chiang and Calautit 2025; Esmaeilzadeh and Hamdy 2025; Menezes et al. 2012; Zare et al. 2023). To address the BEPG, which reaches 60–70% in some cases (Kim, Kim, and Song 2021), researchers used advanced modeling approaches, including data-driven techniques, machine learning (ML), and dynamic simulation (Borodinecs et al. 2024; Norouziaslouziasl and Jafari 2023; Chiang and Calautit 2025). Building energy simulation (BES) is a common method for estimating energy use intensity, incorporating factors such as climate, envelope, service systems, operational schedules, and occupant behavior (OB) (IEA (2013); Hong et al. 2016, Esmaeilzadeh and Hamdy 2025). The accuracy of BES is evaluated using standards such as ASHRAE Guideline 14 (2014) and the International Measurement and Verification Protocol (IPMVP 2012), which set error thresholds for validation (Chong et al. 2025; Kim, Kim, and Song 2021).

Occupant behavior is the most complex and stochastic factor affecting predictive uncertainty (Bahadori-Jahromi et al. 2022; Esmaeilzadeh and Hamdy 2025; Motuziene and Vilutiene 2013), driven by physical (e.g., temperature, air speed, and access to controls; Mitra, Chu, and Cetin 2021) and socio-psychological factors (e.g., comfort expectations, routines, age, gender, and cultural norms; Jia and Srinivasan 2020; Piselli and Pisello 2019). Since accurately incorporating OB into simulation is both

CONTACT Basak Gucyeter bgucyeter@ogu.edu.tr Department of Architecture, Eskisehir Osmangazi University, Bademlik Campus, Odunpazari, Eskisehir 26030, Turkey

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group on behalf of the Architectural Institute of Japan, Architectural Institute of Korea and Architectural Society of China.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

challenging and resource-intensive (Uddin et al. 2021), BES assessments typically depend on simplified, fixed, and deterministic schedules based on standards such as (ASHRAE Standard 90.1 2004; Carlucci et al. 2021; Chiang and Calautit 2025), which largely contribute to the BEPG (Gunay, O'Brien, and Beausoleil-Morrison 2013; Jia et al. 2021; Uddin et al. 2021). Energy consumption in office buildings can vary by 40% when using predetermined schedules (Kang et al. 2024), and prediction errors can reach 36% due to simplified presence data (Anandand, Cheong, and Sekhar 2022; Esmaeilzadeh and Hamdy 2025). To model adaptive OB, such as operating windows, blinds, or adjusting thermostats/HVAC (Gunay et al. 2014; Mitra, Chu, and Cetin 2021; Uddin et al. 2021), as well as non-adaptive behaviors such as presence or absence (Anandand, Cheong, and Sekhar 2022; Mitra, Chu, and Cetin 2021), stochastic and probabilistic methods (e.g., Markov Chains and Agent-Based Models), and machine learning techniques (e.g., ANNs, SVMs; Fabi et al. 2013; Gilani and O'Brien 2019; Norouziaslouziasl and Jafari 2023) are integrated into BES software through direct input or co-simulation (Chen et al. 2024; Kim, Kim, and Song 2021).

Recent efforts aim to address the limitations of simplified models by integrating measurement-based data, such as post-occupancy and field monitoring, into simulation and calibration (Bahadori-Jahromi et al. 2022; Royapoor and Ve Roskilly 2015). Calibration, a systematic adjustment of uncertain parameters to match measured data, improves model reliability by correcting initial assumptions (Chong et al. 2025; Esmaeilzadeh and Hamdy 2025). Uncalibrated models can misrepresent performance by more than 30% (Bahadori-Jahromi et al. 2022; Zare et al. 2023), whereas calibrated models serve as reliable references for retrofits and predictive control (Astete et al. 2022; Gucyeter 2018; Raftery, Keane, and Costa 2009).

Manual iterative calibration is a common approach, representing about 74% of calibration methods in the literature (Coakley, Raftery, and Ve Keane 2014), and is used to obtain reliable BES models (Angelottigelotti et al. 2023). It involves iteratively adjusting uncertain input parameters until outputs, such as energy consumption or temperature trends, match measured data (Cornaro et al. 2019). Key aspects include the following: (1) Reliance on expertise, where the process benefits from the experience and skill of the modeler (Angelottigelotti et al. 2023; Chaudhary et al. 2016). (2) Data collection through audits and monitoring provides detailed insights into the physical and operational characteristics of the building (Sun et al. 2016). (3) Parameter adjustments one at a time, which aim to minimize discrepancies between simulated variables and actual data (Cornaro et al. 2019). (4) Graphical interpretation to visually compare simulated and measured results, such as temperature plots to assess the impact of the iterative process (Sun et al. 2016). (5) Sensitivity analyses to identify and refine influential parameters such as lighting, equipment, air conditioning, and occupancy schedules until statistical criteria are met (Angelottigelotti et al. 2023; Dakyen, Dagbasi, and Özdenefe 2021; Yang et al. 2016). These aspects ensure a thorough understanding and control of the calibration process, avoiding blind tuning of unknown parameters to match the data (Guyot et al. 2020).

Calibration is essential because baseline simulation models often don't match measured data. It minimizes discrepancies between BES predictions and actual energy performance by adjusting uncertain parameters (Angelottigelotti et al. 2023), which enhances confidence and reliability in performance assessment. This is vital for accurately predicting the effects of retrofit measures (Heo, Choudhary, and Augenbroe 2012). Manual calibration must adhere to standards such as ASHRAE Guideline 14 (2014) and IPMVP (2012), which set statistical validation criteria, including the Coefficient of Variation of the Root Mean Square Error (CV(RMSE)) and the Normalized Mean Bias Error (NMBE).

Despite its widespread use, manual calibration lacks standardization (Angelottigelotti et al. 2023; Chaudhary et al. 2016). The complexity of the involved parameters, along with their wide range of variations, often leads to nondisclosed manual calibration steps (Chong, Gu, and Jia 2021). Furthermore, due to subjective expertise, the time spent on modeling, or poor data quality (Chaudhary et al. 2016; Cornaro et al. 2019), the credibility of the resulting calibrated models is sometimes considered highly dependent on the modeler's judgment (Sun et al. 2016). Manual calibration is frequently contrasted with automatic calibration, which employs mathematical and statistical methods, such as optimization algorithms, to minimize errors (Cornaro et al. 2019). Although automatic methods are usually faster computationally, they can act as "black boxes" that produce statistically good fits with physically unreasonable parameter combinations. Conversely, manual calibration involves human reasoning, expertise, and experience, potentially yielding more reliable and physically consistent building models (Chaudhary et al. 2016). Therefore, it is crucial to maintain a certain level of transparency in calibrated BES models to facilitate flexible adjustments in parameter settings.

Limited reporting and reproducibility are especially problematic when calibrating for OB, which significantly influences the BEPG (Cipriano et al. 2015). Uncertainty related to OB is even more critical than that of building envelope properties in highly energy-efficient buildings (Angelottigelotti et al. 2023). Therefore, there is a need for an approach that explicitly records and discloses OB parameter adjustments during calibration to improve reproducibility and reduce bias. This study aims to develop a systematic, evidence-based manual calibration

framework focusing on occupant presence, window operation, and lighting use behavior, with detailed documentation of iterative adjustments. This approach turns subjective expert judgments into flexible, combined steps and utilizes detailed monitoring data (e.g., window open/close states) to enhance model consistency and accurately reflect thermal dynamics. The study investigates whether simplified OB assumptions affect BEPG, how monitored datasets improve calibration, and the effectiveness of adaptive versus non-adaptive behaviors in reducing errors. The present study involved a one-year monitoring campaign at the Eskisehir Osmangazi University Department of Architecture M1 Building to develop an iterative calibration workflow in EDSL Tas. Section 2 covers the case study, monitoring setup, data collection, and calibration procedures. Section 3 presents findings from the iterative calibration based on ASHRAE Guideline 14 (2014) and the IPMVP (2012), while Section 4 synthesizes these findings and explores generalization.

2. Methodology

This study uses a hybrid method combining (i) a year-long in situ monitoring campaign, (ii) OB data extraction, (iii) developing a multi-zone BES model in EDSL Tas, and (iv) an iterative calibration incorporating deterministic assumptions and behavioral data. The framework is summarized in Figure 1.

2.1. Energy performance monitoring

2.1.1. Performance audit of the case building

Before energy performance monitoring, a preliminary audit of the M1 Building at the Architecture Department of Eskisehir Osmangazi University, Turkey, was conducted. This initial phase identified sample areas for monitoring, documented daily schedules, assessed construction discrepancies, and evaluated practical constraints to prevent disruption of academic activities (Ascionecione et al. 2019; Balaras 2022; Poel, van Cruchten, and Balaras 2007). The building's entrance façade faces 12° west of true north. Completed in the 1960s, the exact construction date and as-built drawings were inaccessible. On-site measurements and material inspections were needed to determine envelope features. The building lacks a cooling system, and heating is supplied by a boiler serving three separate buildings located in a neighboring building. Table 1 shows the building's features: a compact structure with a façade-to-volume ratio of 0.24, a window-to-wall ratio (excluding the atrium) of 23%, increasing to 40% with the atrium. It houses faculty offices, design studios, a semi-glazed workshop, an atrium, a graduate classroom, administrative offices, and a cafeteria.

2.1.2. Monitoring campaign

Following the audit, energy performance monitoring began on 1 January 2019, collecting a high-resolution year-long dataset on indoor and outdoor conditions and OB through the end of 2021. This study focuses on 2019, as it provides a complete annual dataset of normal operations before COVID-19 disruptions, reflecting stable schedules and typical occupancy. Later years had irregularities that could have affected the occupant's presence and behavior assessment, and thus the generalization of findings. Hence, 2019 is the best baseline for calibration development and testing.

Representative spaces were selected based on orientation, occupancy, function, and schedules to extrapolate monitored conditions. 39 Testo 174 H temperature and humidity loggers (18 ground floor, 21 first floor), 18 HOBO UX90-005x occupancy and lighting loggers (6 ground floor, 12 first floor), and 8 window loggers (HOBO UX90-001x) in first-floor offices were installed. Table 2 presents the monitoring equipment and measurement intervals. Indoor temperature and humidity were measured at various locations in occupied zones (Figure 2). Sensors, installed per ISO 7726 guidelines at appropriate heights and distances, recorded data every 10 minutes, capturing short-term fluctuations and long-term cycles.

U-value measurements of envelope elements (walls, columns, glazing, and frames) were carried out using Testo 435–2 with a radio-frequency external temperature probe and a three-wire thermocouple surface probe (Table 2). The method involved recording outdoor, indoor, and wall surface temperatures when the indoor–outdoor temperature difference (ΔT) exceeded 15°C, as recommended for accurate thermal transmittance assessment. A summary of measurement dates, locations, points, and data is in Table 3. The measured U-values listed in Table 3 served as inputs for the Baseline (V0.0) and initial calibration iterations. A −5.00% correction factor, described later (V1.9), was applied to these measurements as a calibration adjustment to account for thermal bridging not captured by the point probes.

Outdoor climatic data, including temperature (°C), relative humidity (%), wind speed (m/s), wind direction (°), and global horizontal solar radiation (W/m^2), were recorded at 10-minute intervals, using an on-site Davis Vantage Pro2 microclimatic weather station (Table 2). Hourly cloudiness data were obtained from the national meteorological service. These parameters were later used to develop the simulation weather file and to verify the accuracy of boundary conditions in the building energy simulation model.

Occupant presence and lighting use behavior were recorded at 1-minute intervals. Window operation was also monitored to record open/close states

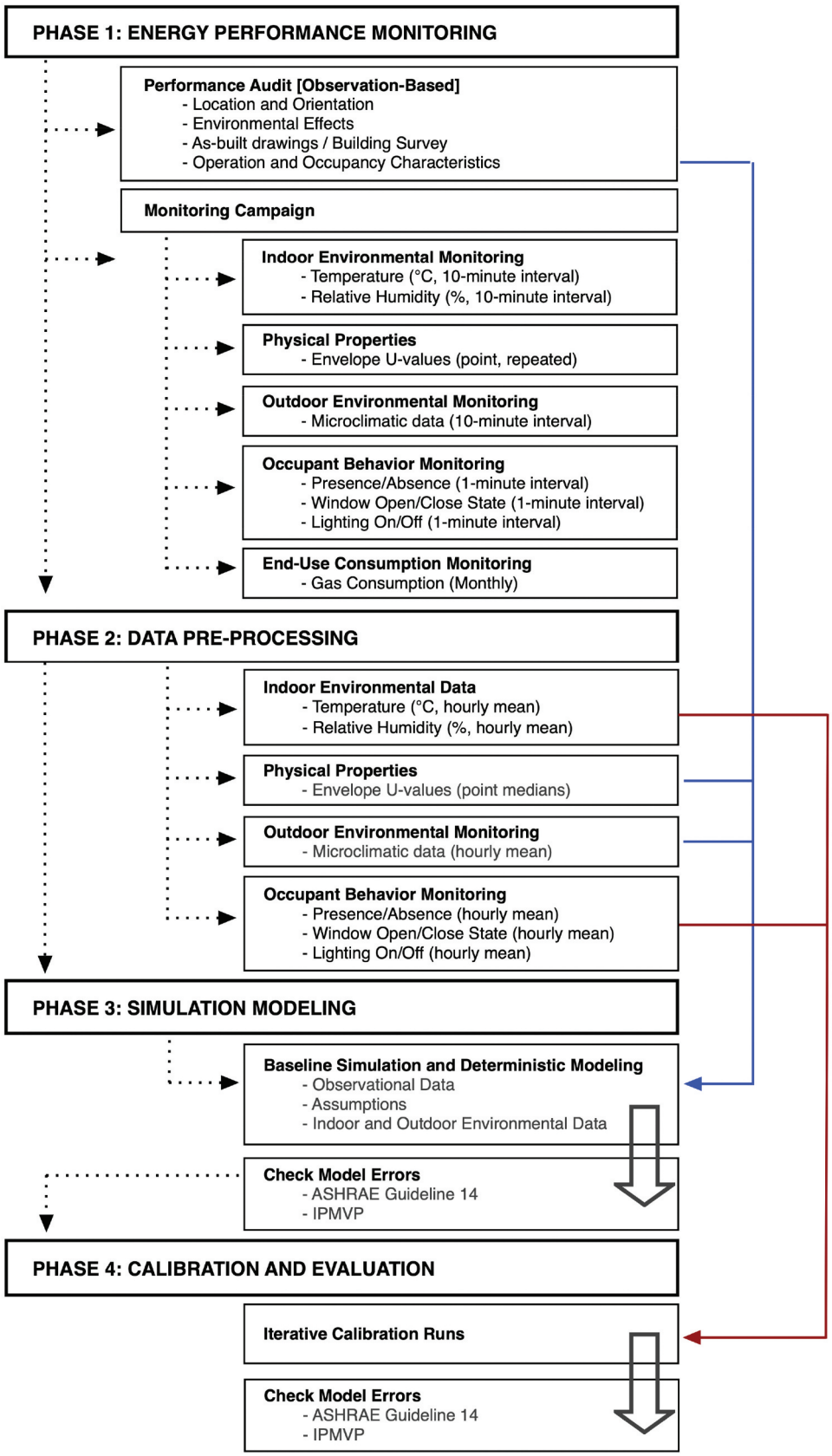


Figure 1. Framework of the study.

Table 1. Case building characteristics.

Characteristic	Data
Location and View of the Case Building	 
Total Floor Area	1486.00 m^2
Heated Floor Area	1336.73 m^2
Floor Height	3.30 m
Total Volume	4903.80 m^3
Total Façade Area	1192.32 m^2
Total Window Area	278.02 m^2
WWR	23.32% (~40% with atrium)
Compactness (A_f/V_{tot})	0.24

(Table 2). No thermostat or HVAC setpoint adjustments were observed, as the case building uses a centrally controlled heating system with no user-side interfaces. No interviews or qualitative surveys were conducted, as the study focused exclusively on objective, measurement-based behavioral characterization. These datasets served as inputs to the behavior-integrated calibration stages described in Section 2.4.

2.2. Data pre-processing

All monitoring data from the 12-month campaign were processed in Microsoft Excel. The workflow involved transferring raw data from all loggers, then systematically cleaning, structuring, and aggregating them. Timestamps were checked for continuity, with missing entries removed. Implausible outliers were filtered, and short gaps (<30 min) filled via linear extrapolation. Data at 10- and 1-minute intervals were converted to hourly averages, giving 8760 hourly values per sensor for 2019. In rooms with multiple sensors, data were averaged to create a single profile, capturing spatial differences, such as interior vs. exterior walls. For example, three sensors in Studio A3 (Figure 2) were processed separately, then averaged to produce a unified temperature and humidity profile. Unmonitored spaces received profiles based on similarities to known rooms. Gas consumption data at the building level were allocated to the M1 Block proportionally, based on its share of the total heated volume. The 2019 heating energy distribution for M1 is shown in Table 4.

2.3. Simulation modeling

Aligned with the framework in Figure 1, simulations of the M1 Building were conducted using EDSL Tas (Environmental Design Solutions Ltd., Tas Engineering, version 9.4.4), a dynamic building energy simulation tool used to predict transient thermal behavior and space-conditioning energy demand. The software employs a zone-based heat balance approach, in which each thermal zone is represented by coupled energy balance equations that account for heat transfer through the building envelope, internal heat gains, ventilation and infiltration, and thermal storage. Transient calculations are performed using time-step methods that resolve the interactions among external boundary conditions, internal loads, and zone air temperatures over the simulation period (EDSL 2026).

EDSL Tas consists of four integrated modules: (1) 3D Modeler, used for defining building geometry and thermal zoning, (2) Building Simulator, where construction properties, internal gains, operational schedules, and system parameters are assigned, (3) Results Viewer, used for post-processing and analysis of simulation outputs; and (4) Tas Systems, used for defining HVAC system configurations. In this study, the building geometry and zoning were first developed in the 3D Modeler and later transferred to the Building Simulator, where envelope properties, internal conditions, and calibration parameters were specified (EDSL 2026). This modular structure enabled a transparent, reproducible implementation within the iterative calibration workflow.

Heating and cooling loads are calculated at the zone level based on the imbalance between internal gains, envelope heat transfer, and air exchange losses. HVAC system performance is represented in the TAS Systems module through user-defined system configurations and control assumptions, enabling the simulation of centralized heating operation and zone-level thermal response under specified setpoints and

Table 2. Monitored parameters and monitoring intervals.

Parameters	Interval	Device	Accuracy
Indoor Temperature (°C)	10-minute	Testo 174 H Temperature and Humidity Datalogger	±0.5 °C (−20 to 70 °C)
Indoor Relative Humidity (%)			±3 %RH (2 to 98%)
Occupant Presence (1/0)	1-minute	HOBO Occupancy/Light	94° (±47°) Horizontal; 82° (±41°) Vertical
Lighting State (on:1/off:0)		Data Logger (UX90-005x)	>65 lux
Window State (open:1/ closed:0)	State	HOBO State Data Logger (UX90-001x)	±1 minute per month at 25°C
Gas Consumption	Monthly	Monthly Bills	
U-value (W/m^2K)	1-second Repeated	Testo 435–2 + Radio-frequency external temperature probe + three-wire thermocouple surface probe	Radio probe: ±0.5 °C + 0.3 % of remaining range (−40 to 500 °C) Surface probe: ±0.1 °C + 0.2 % of remaining range (−20 to 70 °C), when used with a radio probe.
Outdoor temperature (°C)	10-minute	Davis Vantage Pro2 Microclimatic Weather Station	±0.5 °C
Outdoor Relative Humidity (%)			±3 %
Global Horizontal Solar Radiation (W/m^2)			±5 %
Wind Speed (m/s)			±5 %
Wind Direction (°)			
Cloudiness (0–1)	1-hour	Meteorological Data	NA

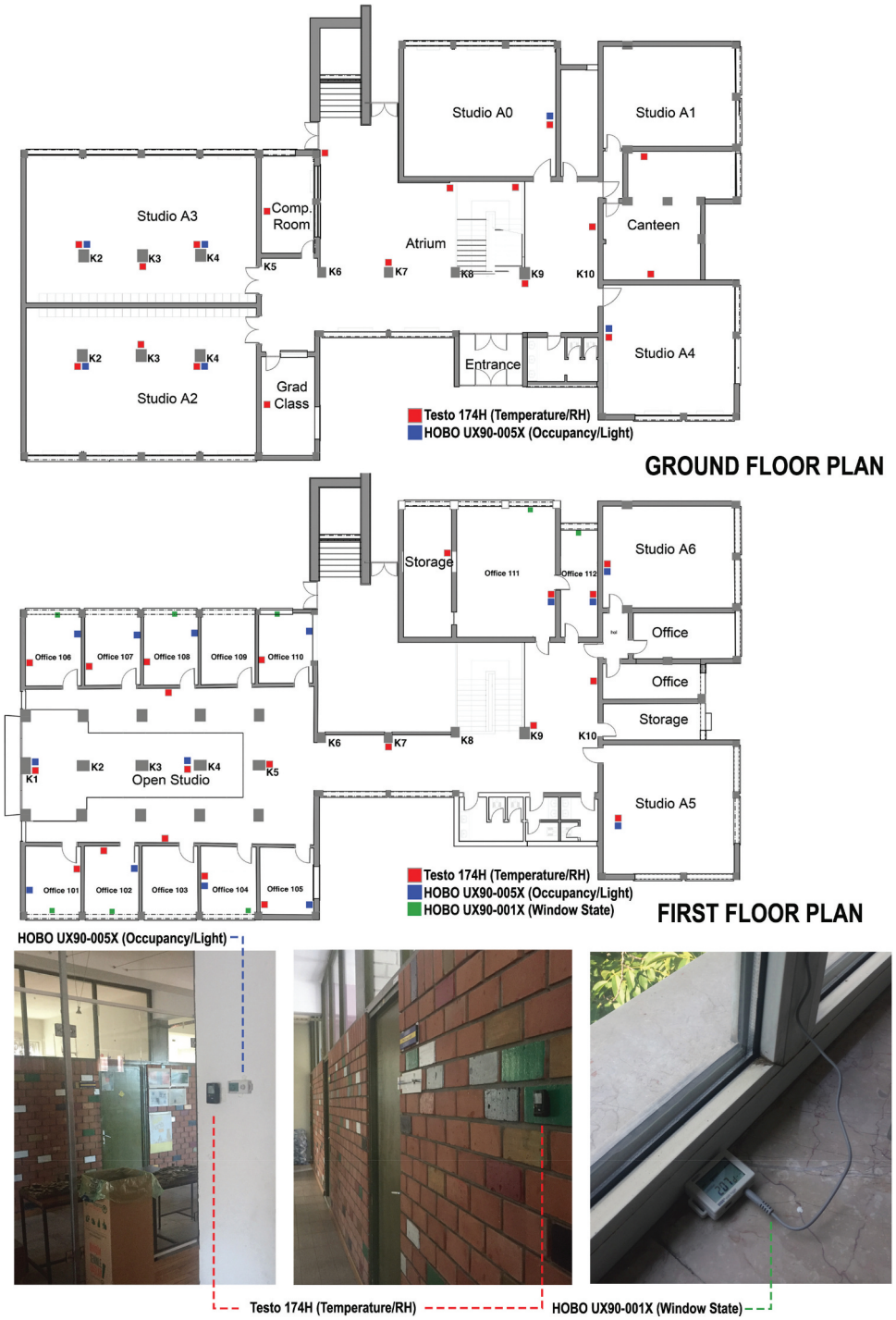


Figure 2. Dataloggers installed in the building.

schedules (EDSL 2026). While system modeling is simpler than component-level plant simulations, this approach is well suited for assessing space-conditioning demand and for calibration against measured energy use in existing buildings.

EDSL Tas complies with the ASHRAE Standard 140 test suite, which provides a standardized framework for assessing the performance of building energy simulation programs under controlled thermal scenarios. Hence, the software produces results consistent with other established simulation tools for zone heat balance calculations and space-conditioning load prediction, supporting its suitability for the present study (EDSL 2026).

Therefore, a multizone simulation model was developed in EDSL Tas to accurately represent the building's spatial and functional divisions. All physical and operational features were first constructed in the 3D Modeler and then transferred to the Building Simulator module. Figure 3 illustrates the resulting building geometry showing the north and south facades. The model captures the building's spatial configuration, thermal

Table 3. Measured U-values for envelope components.

								Accepted U-value		
Building Envelope Element	Space	Measurement Date	Number of Measurement Points	Max	Min	Mean	Std. Dev	Mean	Façade Ratio	Normalized
Reinforced Concrete Column										**1.678**
M01	Office 102	11 March 2021	2300	1.989	1.836	1.978	0.069	**1.746**	%33	
M02	Office 102	16 March 2021	1690	2.173	2.041	2.103	0.025			
M03	Kitchen	18 March 2021	2906	1.886	1.764	1.834	0.025			
M04	Studio A5	18 March 2021	3602	1.741	1.289	1.481	0.119			
M05	Studio A6	23 March 2021	3857	1.554	1.245	1.335	0.062			
Brick Exterior Wall										
M06	Studio A3	23 March 2021	445	1.598	1.542	1.566	0.110	**1.645**	%67	
M07	Office 102	9 March 2021	1657	1.412	1.157	1.279	0.066			
M08	Office 102	9 March 2021	2410	1.588	1.267	1.347	0.055			
M09	Office 102	9 March 2021	1499	1.935	1.842	1.889	0.018			
M10	Office 102	10 March 2021	3246	1.366	1.271	1.313	0.027			
M11	Office 102	11 March 2021	3128	1.998	1.746	1.875	0.062			
M12	Office 102	11 March 2021	2508	1.629	1.499	1.557	0.026			
M13	Circul. F02	16 March 2021	2624	1.725	1.622	1.668	0.019			
M14	Circul. F02	16 March 2021	2951	2.238	1.957	2.113	0.052			
M15	Studio A5	18 March 2021	2096	2.073	1.856	1.949	0.044			
M16	Studio A5	18 March 2021	2128	1.702	1.421	1.522	0.058			
M17	Studio A6	23 March 2021	2444	1.725	1.622	1.668	0.019			
Glazing										
M18	Studio A3	17 December 2021	961	2.010	1.837	1.947	0.045		**2.054**	
M19	Office 102	9 March 2021	1800	2.292	1.858	2.105	0.070			
M20	Office 102	9 March 2021	2267	2.266	2.089	2.165	0.046			
M21	Studio A6	11 March 2021	2143	2.118	1.918	2.000	0.037			
Frame										
M22	Office 102	9 March 2021	1477	2.568	2.411	2.490	0.037		**2.104**	
M23	Office 102	11 March 2021	3736	2.398	2.238	2.303	0.027			
M24	Kitchen	18 March 2021	2601	1.686	1.582	1.634	0.020			
M25	Studio A6	23 March 2021	2242	2.148	1.828	1.992	0.071			

Table 4. Monthly gas consumption shares for M1 building (2019).

Month	Gas Consumption (kWh)
January	49883,91
February	43720,39
March	30349,61
April	11574,24
May	1511,48
June	17,34
October	2257,02
November	30694,70
December	56656,58
Total	**226665,28**

zoning, orientation, and envelope articulation, which form the basis for assigning construction properties, internal gains, and operational schedules within the simulation environment.

The weather data interface of the Building Simulator allows on-site measured weather data to be imported in an hourly tabular format. In this study, the hourly microclimatic dataset obtained from on-site monitoring was formatted accordingly and integrated into the simulation to define boundary conditions for the annual analysis. Measured U-values for walls, columns, glazing, and frames were incorporated into the model as listed in Table 3 minimizing envelope-related uncertainty and improving alignment between the physical building and the simulated representation.

Internal conditions and system-related parameters were specified based on a combination of field observations, measured data, and standard references. Heating setpoints and controls were defined according to observations of the centrally managed heating system. Boiler efficiency was assumed to be 79% under typical winter non-condensing operation, consistent with experimentally reported efficiencies for traditional boilers (Bălănescu and Homutescu 2018). Individual thermostat control was not modeled, as occupants did not have local temperature control. Infiltration and ventilation rates were estimated using ASHRAE (2017) standard values, as direct measurements were unavailable. Occupancy, lighting, equipment, and metabolic gains were assigned using a combination of schedule-based monitoring data (including occupant presence, lighting use, and window operation) and standard assumptions from ASHRAE (2017) and ASHRAE 55 2010.

Internal gains and air exchange parameters were assigned at the zone level through the Internal Conditions module, using hourly schedules or functional factors: infiltration (ach), ventilation (ach), lighting gains (W/m^2), occupant sensible gains (W/m^2), occupant latent gains (W/m^2), equipment sensible gains (W/m^2), and equipment latent gains (W/m^2). By combining measured microclimatic data, empirically derived U-values, field observations of heating system operation, and standard internal-gain assumptions, the baseline BES model for the monitored building was developed. As shown in Figure 4, the baseline model served as the starting point for the calibration process described in Section 2.4. The baseline was deterministic, relying on 2004 occupancy schedules, default profiles for equipment and lighting use, reported HVAC operation schedules, measured envelope properties, and the 2019 weather dataset obtained from on-site monitoring. It represented typical practice, with deterministic occupant presence and behavior. The difference between baseline predictions and measured energy use provided the initial estimate of the building energy performance gap (BEPG).

2.4. Calibration and evaluation

The third step in the methodological framework demonstrates that a BES model based on assumed or deterministic inputs, especially when parameters are limited, produces energy predictions that deviate from real building behavior. Therefore, phase four of the workflow (Figure 1) focuses on developing a structured

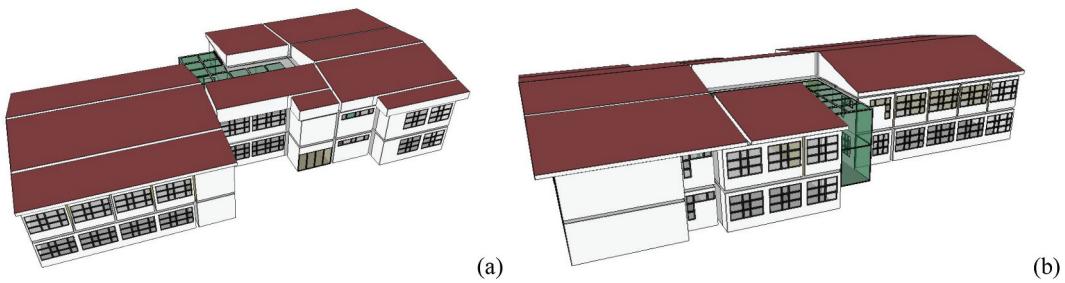


Figure 3. EDSL Tas model (a) north and (b) south facades.

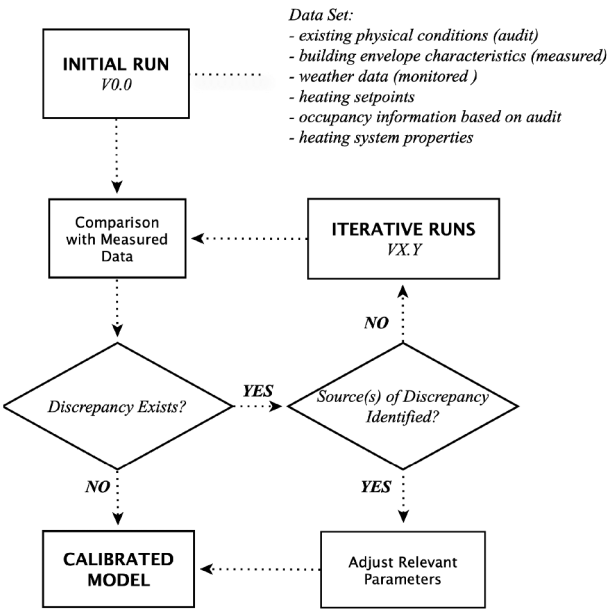


Figure 4. Iterativeerative calibration approach (Gucyeter 2018; Raftery, Keane, and Costa 2009).

calibration procedure to systematically reduce this discrepancy and assess the contribution of occupant-related behavioral parameters to model accuracy. As discussed in the Introduction, calibrated models reduce the BEPG and improve predictive reliability. This section describes the calibration methodology in detail, which allowed the assessment of the main research questions: (i) whether simplified OB models contribute to BEPG, (ii) how measured occupant data affect calibration precision, and (iii) how adaptive and non-adaptive behaviors impact the calibrated model's performance. The calibration approach adopted in the present study followed the iterative strategy previously proposed by Gucyeter (2018), adapted from Raftery, Keane, and Costa (2009), and informed by the requirements of ASHRAE Guideline 14 (2014) and IPMVP (2012).

2.4.1. Iterative calibration workflow

When deviations between the model and measurements exceeded thresholds in Table 5, calibration continued by adjusting the most probable source of discrepancy. If a specific parameter, such as infiltration, setpoint, or thermal transmittance, was identified, it was modified, and the simulation was rerun. If no direct source was found, parameters were adjusted sequentially following the workflow in Figure 4 until the simulated and measured data converged and the CV (RMSE) and NMBE decreased.

Table 5. Calibration benchmarks.

		Calibration Type	
Guidelines		Hourly	Monthly
ASHRAE 14	NMBE	±10%	±5%
	CV(RMSE)	30%	15%
IPMVP	NMBE	±5%	±20%
	CV(RMSE)	20%	NA

A four-step hierarchical framework was used to calibrate the building energy model (Figure 5): (1) a conventionally calibrated baseline: V1, (2) presence-based calibration: V2, (3) presence and window operation behavior: V3, and (4) calibration combining presence, window operation, and lighting use behavior: V4. The framework systematically evaluated how behavioral parameters affected discrepancies between the model and monitoring data. The uncalibrated baseline model (V0.0) was created through five runs using the physical, operational, and climatic data described in Section 2.3. It was then calibrated using standard parameters, including infiltration coefficients, envelope U-values, heating setpoints, internal gains, and assumed indoor temperature profiles. V1 produced models calibrated with conventional parameter adjustments until the lowest errors were achieved, per ASHRAE Guideline 14 (2014). After V1, the workflow incorporated empirical occupant presence data (V2). Two iterations assessed the impact of refining presence data: V2.1 used only measured occupant presence (from HOBO UX90-005x). V2.2 combined measured presence with extrapolated schedules for unmonitored zones. Both began from the best-performing model (V1.9). This phase evaluated how replacing default schedules with empirical patterns improved calibration accuracy.

Next, ventilation-related behavior for window operation events (V3) was integrated into two model types. Single-parameter iterations (S) isolated the effect of the window operation behaviors. V3.1S used the measured window state, while V3.2S incorporated inferred (measured and extrapolated) window state behavior for

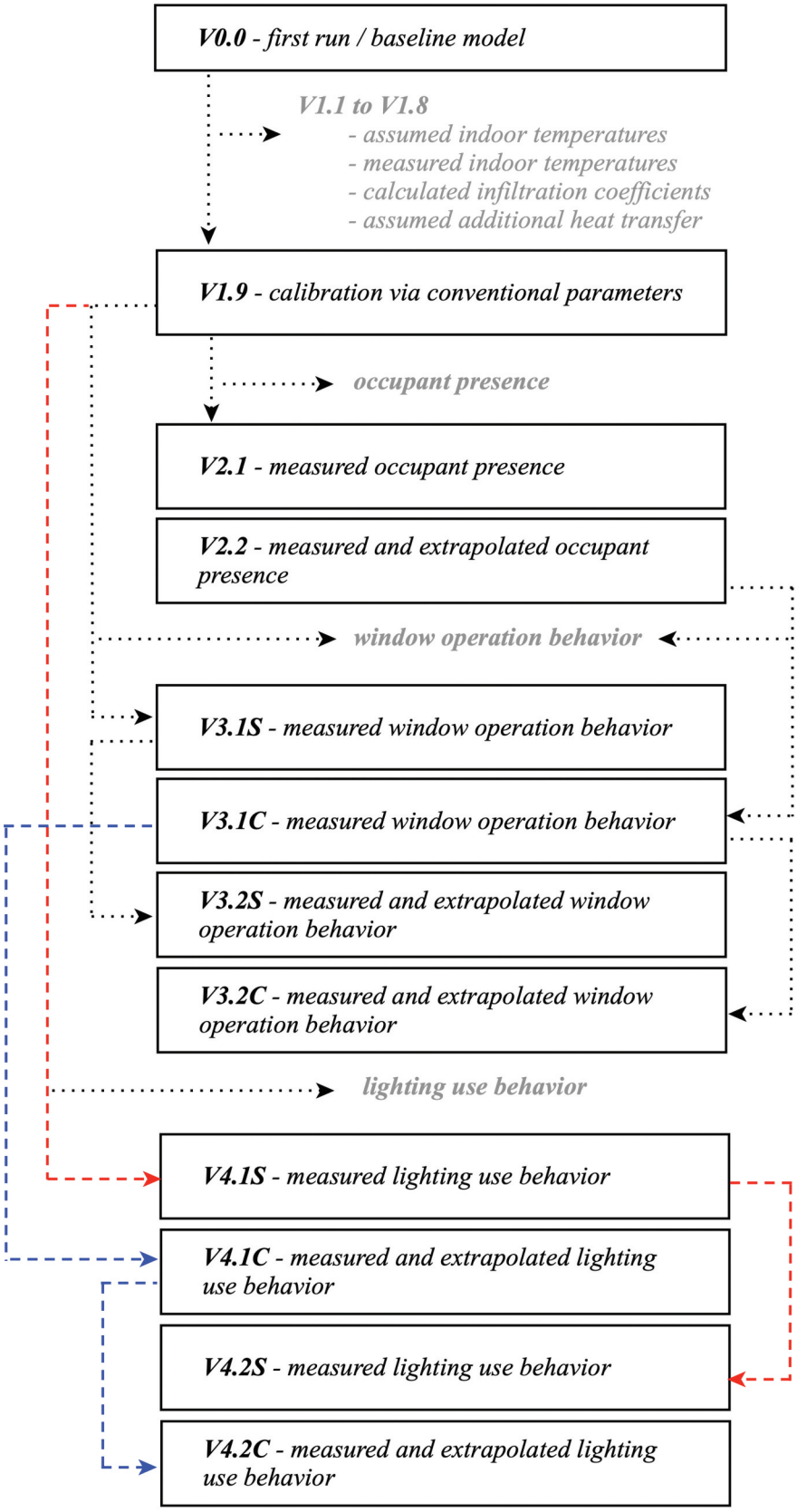


Figure 5. Iterativeerative calibration framework using single and combined parameter sets.

unmonitored zones using the V1.9 model. Combined-parameter iterations (C) evaluated window operation behavior along with empirical presence data. V3.1C integrated measured window operation with measured and derived presence schedules using V2.2, whereas V3.2C added extrapolated window operation behavior to the V3.1C model. This stage examined how adaptive window operation affected heat gains, losses, and air-change rates when applied alone or in combination with refined presence schedules.

The last step, V4, incorporated lighting use behavior using the same single-parameter and combined-parameter logic. V4.1S applied measured lighting use behavior directly to V1.9, while V4.2S expanded this with extrapolated lighting schedules. Combined-parameter iterations V4.1C and V4.2C were constructed based on the presence-and-window-operation data model (V3.1C), enabling the evaluation of lighting use as part of a fully integrated behavioral profile. This final stage completed the calibration hierarchy, enabling systematic comparison of how each behavioral component, presence, window operation, and lighting use contributed to improving model accuracy.

2.4.2. Evaluation metrics and analysis

Model predictive accuracy was assessed by comparing hourly indoor temperatures and monthly heating consumption with monitoring data across three stages: (1) deterministic baseline model, (2) conventionally calibrated model, and (3) behavior-integrated models. Performance was measured using NMBE and CV (RMSE), calculated as follows:

$$NMBE(\%) = \frac{\Sigma_{i=1}^{n}(M_i - S_i)}{(n-p)\ \bar{M}} \quad (1)$$

$$CV(RMSE)(\%) = \frac{\sqrt{\frac{1}{n-p}\Sigma_{i=1}^{n}(M_i - S_i)^2}}{\bar{M}} \times 100 \quad (2)$$

Where:

- M_i is the measured value at time i,
- S_i is the simulated value at time i,
- n is the number of data points,
- p is the number of adjusted calibration parameters,
- $\bar{M}$ is the mean of the measured values.

Results were benchmarked against ASHRAE Guideline 14 (2014) and the IPMVP (2012) compliance limits (Table 5), including secondary metrics such as mean hourly correlation (r^2) and temperature difference (ΔT). Following the hierarchical workflow of modifying only one parameter per iteration with prior inputs fixed, the independent parameter count was $p = 1$ for calibrated simulations.

3. Findings and discussion

3.1. Uncalibrated baseline (V0)

The baseline model (V0.0) was created using measured envelope U-values, on-site weather data, and deterministic internal gains and schedules based on 2004. The simulation covered 2019, with hourly outputs compared to monitored data from 32 zones. In terms of thermal behavior, the baseline model captured the building's indoor temperature trends but exhibited significant local deviations. The mean hourly correlation coefficient (r^2) across all zones was 0.69 (Table 6), with individual zones ranging from 0.90 (atrium and workshop) to 0.28 (computer lab).

The mean temperature difference between measured and simulated values was 1.22°C (Table 6). To move beyond aggregate error metrics and examine the distributional structure of model error, Figure 6 presents the frequency distributions of hourly indoor temperature differences ($\Delta T = T_m - T_s$) for the uncalibrated baseline (V0) and the conventional calibration (V1), based on 8760 hours of data across 32 zones, totaling 280,320 data points. In Figure 6, the blue histogram bars represent the empirical frequency of ΔT values, while the fitted trendline (dotted red) highlights the systematic bias and distributional tendencies within each iteration. In V0.0, errors exhibited a platykurtic distribution with a flattened peak at 28,950 hours (mode = 1.00°C) and a wide error extending from approximately −6.00°C to +10.00°C. This broad dispersion indicated that, despite a low mean error, the uncalibrated model showed substantial deviations from actual hourly temperature fluctuations, suggesting model deficiencies that are not captured by mean-based performance indicators alone.

The NMBE and CV (RMSE) for indoor temperature were 5.24% and 16.70%, respectively. These metrics were acceptable for hourly calibration, but the distribution of temperature differences (Figure 6) indicated that the model failed to capture specific thermal dynamics and did not follow a normal error distribution. Heating energy prediction revealed a significant gap: simulated annual consumption was 143.86 kWh/m^2, while actual consumption was 169.57 kWh/m^2, resulting in an NMBE of −15.16%, a CV(RMSE) of 26.42%, and a failure to meet ASHRAE or IPMVP standards, confirming the V0.0's limitations. Although thermally plausible ($r^2 = 0.69$), V0.0 inaccurately estimated consumption, due to inadequate representation of infiltration, setpoint schedules, and internal gains. The findings are consistent with prior research showing that ungrounded assumptions about infiltration and schedules are major contributors to performance gaps (Menezes et al. 2012). The baseline model, therefore, provided a clear motivation for a structured calibration.

3.2. Conventional parameter calibration (V1)

This phase addressed uncertainties related to the envelope and building operation using standard calibration parameters, specifically indoor temperature setpoints, infiltration rates, and envelope thermal properties, before considering the effects of stochastic OB on model calibration.

Table 6. Summary of calibration iterations for the conventional parameter phase (*V0.0 – V1.9.*).

Model Code	Modified Parameters	Infiltration (ACH)	U-Values (W/m^2K)	Internal Gains	Heating Setpoint	Mean Hourly Correlation (T, r^2)	Mean Hourly Temperature Difference (ΔT, °C)	Hourly Indoor Temperatures		Monthly Heating Consumption		ASHRAE 14 Compliance
								NMBE (%)	CV(RMSE) (%)	NMBE (%)	CV(RMSE) (%)	
V0.0	Uncalibrated baseline	Default (0.30)	Measured	Default	- Upper limit 22.00°C - Lower limit 18.00°C	0.69	1.22	5.24	16.70	-15.16	26.42	No
V1.1	Monitored hourly indoor temperatures (with correction)	Default (0.30)	Measured	Default	- Upper limit +0.50°C - Lower limit -1.00°C - 10% setpoint reduction weekend	0.85	0.99	4.23	13.62	-13.06	21.27	No
V1.2	Extrapolated infiltration assumptions	Between 0,26 (circulation F02) and 5,00 (attic)	Measured	Default	- Upper limit +0.50°C - Lower limit -1.00°C - 10% setpoint reduction weekend	0.87	1.67	7.18	12.71	25.63	32.79	No
V1.3	ATTMA-based infiltration estimations (F=0.40)	Between 0,16 (circulation F02) and 3,10 (attic)	Measured	Default	- Upper limit +0.50°C - Lower limit -1.00°C - 10% setpoint reduction weekend	0.86	1.42	5.97	12.93	5.04	15.26	Yes
V1.4	ATTMA-based infiltration estimations (F=0.30)	Between 0,12 (circulation F02) and 2,33 (attic)	Measured	Default	- Upper limit +0.50°C - Lower limit -1.00°C - 10% setpoint reduction weekend	0.85	1.25	5.56	13.13	-3.08	14.78	Yes
V1.5	Correction of weekend setback assumptions	Between 0,12 (circulation F02) and 2,33 (attic)	Measured	Default	- Upper limit +0.50°C - Lower limit -1.00°C	0.87	0.96	4.13	11.50	2.47	14.19	Yes
V1.6	Monitored hourly indoor temperatures (without correction)	Between 0,12 (circulation F02) and 2,33 (attic)	Measured	Default	- 8760 h/zone monitored hourly indoor temperatures	0.88	0.36	1.52	9.67	12.13	18.68	No
V1.7	Monitored hourly indoor temperatures (with correction)	Between 0,12 (circulation F02) and 2,33 (attic)	Measured	Default	- Upper limit +0.50°C - Lower limit -0.50°C	0.88	0.66	2.81	10.48	7.29	15.79	Yes
V1.8	Monitored hourly indoor temperatures (with correction)	Between 0,12 (circulation F02) and 2,33 (attic)	Measured	Default	- Upper limit +0.50°C - Lower limit -1.50°C	0.86	1.24	5.21	12.80	-2.34	14.23	Yes
V1.9	U-Value Corrections	Between 0,12 (circulation F02) and 2,33 (attic)	-5.0% correction	Default	- Upper limit +0.50°C - Lower limit -0.50°C	0.87	0.53	2.26	10.82	1.39	14.21	**Yes (reference model)**

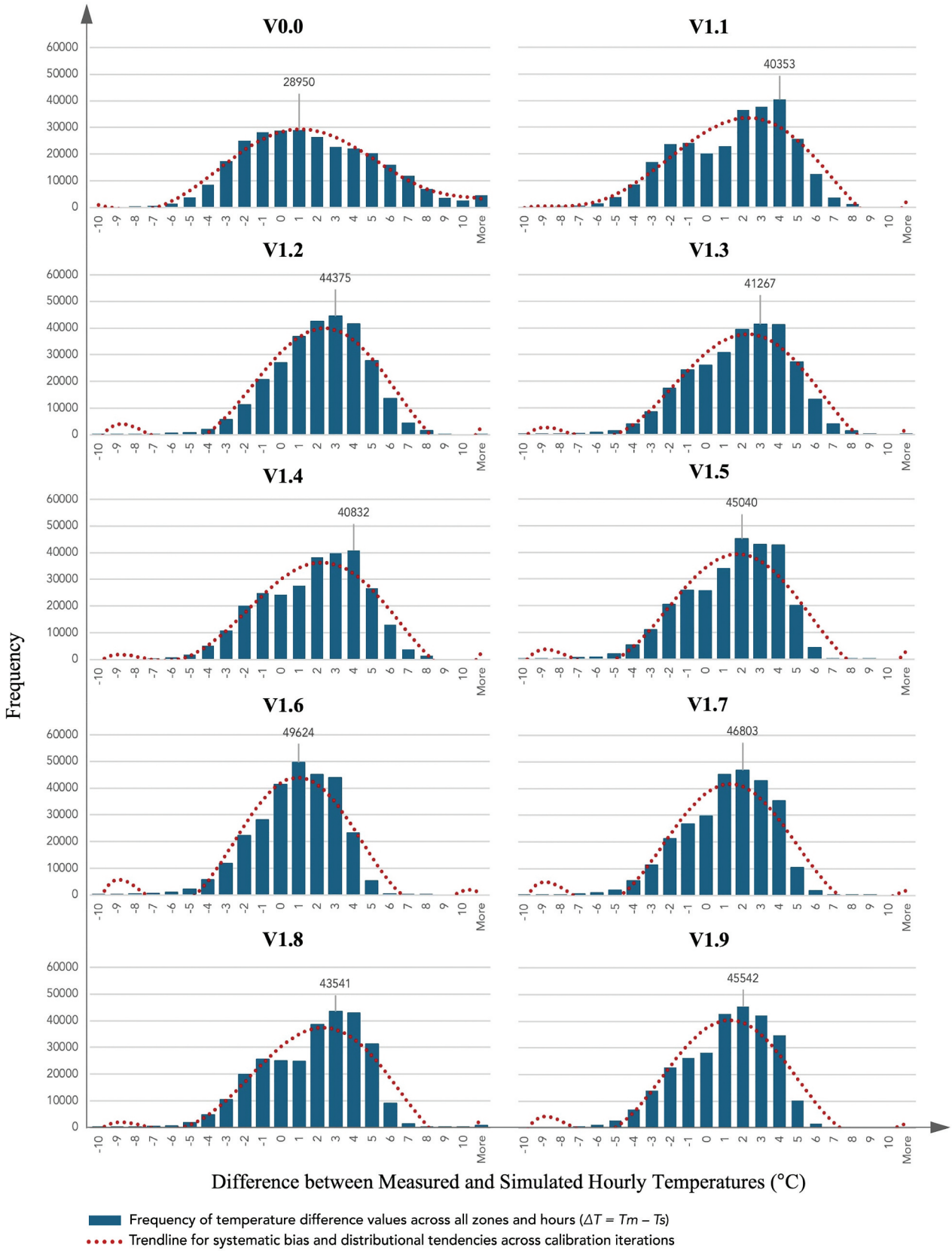


Figure 6. Frequency distribution of hourly temperature differences ($\Delta T = Tm - Ts$) across iterative calibration runs V0.0 to V1.9.

3.2.1. Integration of monitored hourly temperatures (V1.1)

The first calibration iteration (V1.1) replaced fixed heating schedules, previously based on indoor temperature margins, with annual hourly temperature profiles from monitoring data. A + 0.50°C adjustment was applied to the hourly-averaged temperatures as the upper temperature setpoint limit, and a −1.00°C adjustment was applied to the lower heating setpoint. Data showed that the indoor heating setpoint was lowered on weekends; therefore, a median 10% reduction in the heating setpoint was also implemented in V1.1. These adjustments improved the accuracy of the temperature profile. The mean correlation coefficient (r^2) increased to 0.85, and the mean temperature difference decreased to 0.99°C, with an NMBE and CV(RMSE) of 4.23% and 13.62%, respectively. However, improvements in predicting heating energy were minor: NMBE remained at −13.06%, and CV(RMSE) decreased slightly to 21.27%, still exceeding ASHRAE Guideline 14 and IPMVP limits.

The error histograms show a clear difference between the uncalibrated baseline (V0.0) and the first calibration iteration (V1.1). The baseline had a wide, platykurtic error distribution (Section 3.1; Figure 6), indicating high random variation. V1.1 replaced fixed schedules with annual hourly profiles constrained by setpoint limits and a median 10% weekend reduction. These changes created a more peaked, compressed distribution with 40,353 hours, improving indoor temperature calibration accuracy ($r^2 = 0.85$). However, this consistency reveals a systematic bias, with a skewed distribution shifting the mode from about 1.00°C to 4.00°C. The simulation underpredicts indoor temperatures, and despite strict setpoint constraints, minimal improvement in heating energy prediction suggests physics issues, such as uncalibrated infiltration or U-values, or operational factors, such as weekend setbacks, cause the model to run colder than actual. For V1.1, hourly temperature schedules alone cannot fully address the underestimation of heating energy; envelope or infiltration parameters account for most of the discrepancies. Indoor temperature accuracy can improve without proportional heating prediction improvements for older buildings with continuously operating heating systems that obscure underlying envelope inefficiencies.

3.2.2. Extrapolated infiltration assumptions (V1.2)

To address the underestimation of heat loss observed in V0.0 and V1.1, airtightness values (ACH50) measured in a neighboring building of similar age and construction on the same campus were applied to the model (ranging from 0.26 to 5.00 ACH) to test the transferability of envelope infiltration data. V1.2 resulted in a significant overestimation of energy use. The heating consumption NMBE shifted from a −13.06% underestimation to a + 25.63% overestimation, and the CV (RMSE) increased to 32.79% (Table 6). Results showed airtightness is building-specific and cannot be reliably transferred, even between similar structures. Error analysis (Figure 6) revealed a leptokurtic distribution with a peak at 44,375 hours and a positive temperature bias, where simulated indoor temperatures were lower than actual temperatures. Excessive infiltration coefficients led to an overestimate of heating consumption (+25.63% NMBE), failing to account for actual use fully. Indoor temperature predictions were low, confirming that airtightness is a highly building-specific factor influenced by construction defects, which is critical for accurate heating energy modeling.

3.2.3. Calculation-based infiltration estimations (V1.3 and V1.4)

Airtightness coefficients, based on the 2010 standard, were recalculated by estimating infiltration from envelope geometry instead of surrogate measurements. Two scaling factors tested were a normal baseline (F = 0.40 in V1.3) and a tighter envelope assumption (F = 0.30 in V1.4). Both iterations brought the model within the benchmarks (Table 6), confirming that the estimates were more reliable than those from surrogate data. V1.3 achieved an NMBE of +5.04% and a CV (RMSE) of 15.26%. Tighter assumptions in V1.4 (0.12–2.33 ACH) improved consumption prediction, with NMBE shifting to −3.08% and CV (RMSE) to 14.78%. This suggests that the actual envelope was less permeable than the standard assumption. Despite better consumption estimates, the error distribution remained similar (Figure 6). V1.3's mode was 3.00°C (41,267 hours), while V1.4 shifted to 4.00°C (40,832 hours), indicating the model underpredicted indoor temperatures. The discrepancy between high consumption accuracy and biased temperature predictions indicates that the model ran cooler than the actual building, an issue to be addressed in future calibration through evaluating the dynamic thermal response and thermal mass.

3.2.4. Correction of weekend setback assumptions (V1.5)

Previous iterations assumed a 10% reduction in weekend heating regimes. Iteration V1.5 tested continuous heating without weekend setback, resulting in a noticeable change in energy predictions. The heating consumption NMBE shifted from underestimation in the previous step (−3.08%) to slight overestimation (+2.47%), while the CV (RMSE) improved to 14.19% (Table 6). This confirms that the building's heating system operates as on workdays due to technical factors such as freeze protection and the high thermal inertia of the hydronic system, which discourage frequent shutdowns. Error distribution for V1.5 (Figure 6)

supports this improvement, as the mode in V1.4 was at 4.00°C, shifting significantly to 2.00°C in V1.5. At the same time, the peak frequency of hours around the mode increased to 45,040 hours, indicating greater temporal consistency. Additionally, improved indoor temperature calibration ($r^2 = 0.87$) and reduced mean hourly temperature difference ($\Delta T = 0.96$°C) suggest that the weekend setback assumption in earlier models contributed to the higher positive temperature differences (Table 6).

3.2.5. Direct use of monitored hourly temperatures (V1.6)

Iteration V1.6 removed temperature control limits and used monitored temperatures as hourly setpoints to see whether exact empirical data would improve calibration. This adjustment reduced prediction accuracy, increasing NMBE to 12.13% and CV (RMSE) to 18.68%, failing ASHRAE Guideline 14 benchmarks (Table 6). Although the model closely tracked indoor environmental conditions ($r^2 = 0.88$; $\Delta T = 0.36$°C), the energy consumption levels required to do so were high, indicating that the simulation was micro-cycling or responding to noise in the raw data that actual heating systems with hysteresis would ignore. The error distribution for V1.6 (Figure 6) highlights this, showing peaking at 49,624 hours and clustering around 1.00°C. The discrepancy between error frequency, mean temperature difference, and unrealistic consumption shifts (Table 6). suggests a need to smooth thermostat deadband data to simulate control inertia. Removing the deadband caused the model to overreact to tiny temperature changes, increasing energy use despite a seemingly ideal error distribution.

3.2.6. Use of monitored hourly temperatures with control deadband (V1.7 and V1.8)

V1.7 and V1.8 tested reintroducing a control deadband around hourly indoor temperature data. This aimed to balance the thermal accuracy of V1.6 with the more realistic consumption patterns of earlier iterations. In V1.7, an upper limit of +0.50°C and a lower limit of −0.50°C were applied to monitored hourly data, whereas V1.8 widened the deadband asymmetrically to an upper limit of +0.50°C and a lower limit of −1.50°C. In V1.8, the NMBE and CV(RMSE) decreased to 7.29% and 15.79%, meeting ASHRAE guidelines, with an indoor temperature correlation of 0.88 and a mean difference of 0.66°C. However, widening the band in V1.8 caused the model to under-predict consumption (NMBE = −2.34%; CV(RMSE) = 14.23%) and to shift the temperature calibration ($r^2 = 0.86$; $\Delta T = 1.24$°C), indicating that V1.7's range provided better calibration. The error distribution for V1.7 peaked at 2.00°C, whereas V1.8 had a flatter distribution. V1.7 balanced good temperature fit with accurate energy consumption reflection.

3.2.7. U-Value adjustments (V1.9)

By iteration V1.9, further thermostat deadband testing was unnecessary, as a calibrated model had been established in V1.7. To reduce the gap between simulated and measured heating, a uniform −5.00% correction was applied to envelope U-values, aligning energy use with actual data (NMBE = 1.39%; CV (RMSE) = 14.21%). This calibration, meeting ASHRAE 14 and IPMVP standards, confirmed model stability (Figure 6) with a leptokurtic error distribution, peak at 45,542 hours at 2.00°C. U-value reductions primarily affected energy needs rather than temperature distribution, achieving balanced thermal profiles and consumption patterns. The resulting reliable model minimized errors and served as a reference for evaluating OB adjustments.

3.3. Sensitivity analysis of behavior-integrated calibration (V2 to V4)

The conventional calibration created a reference model (V1.9) that met the benchmarks (NMBE = 1.39%) but relied on deterministic assumptions, using the default 2004 schedules and the internal gains listed in Table 6. The subsequent integration of measured behavioral datasets in phases V2 through V4 led to

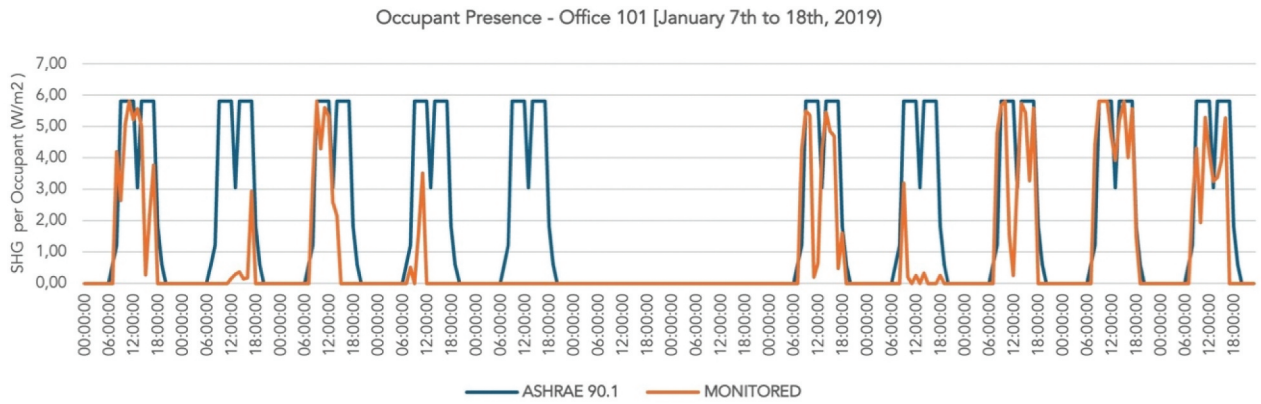


Figure 7. Internalternal gains based on occupant presence for office 101 – ASHRAE 90.1 vs. Monitored.

shifts in energy prediction accuracy. Instead of viewing these shifts solely as changes in reliability, this section examines them as sensitivity indices that measure the impact of specific adaptive and non-adaptive behaviors relative to the calibrated reference model.

3.3.1. Sensitivity to non-adaptive occupant presence (V2.1 and V2.2)

The first iteration, V2.1, replaced 2004 occupancy schedules with high-resolution presence data from 17 monitored zones. V2.2 used data extrapolation for

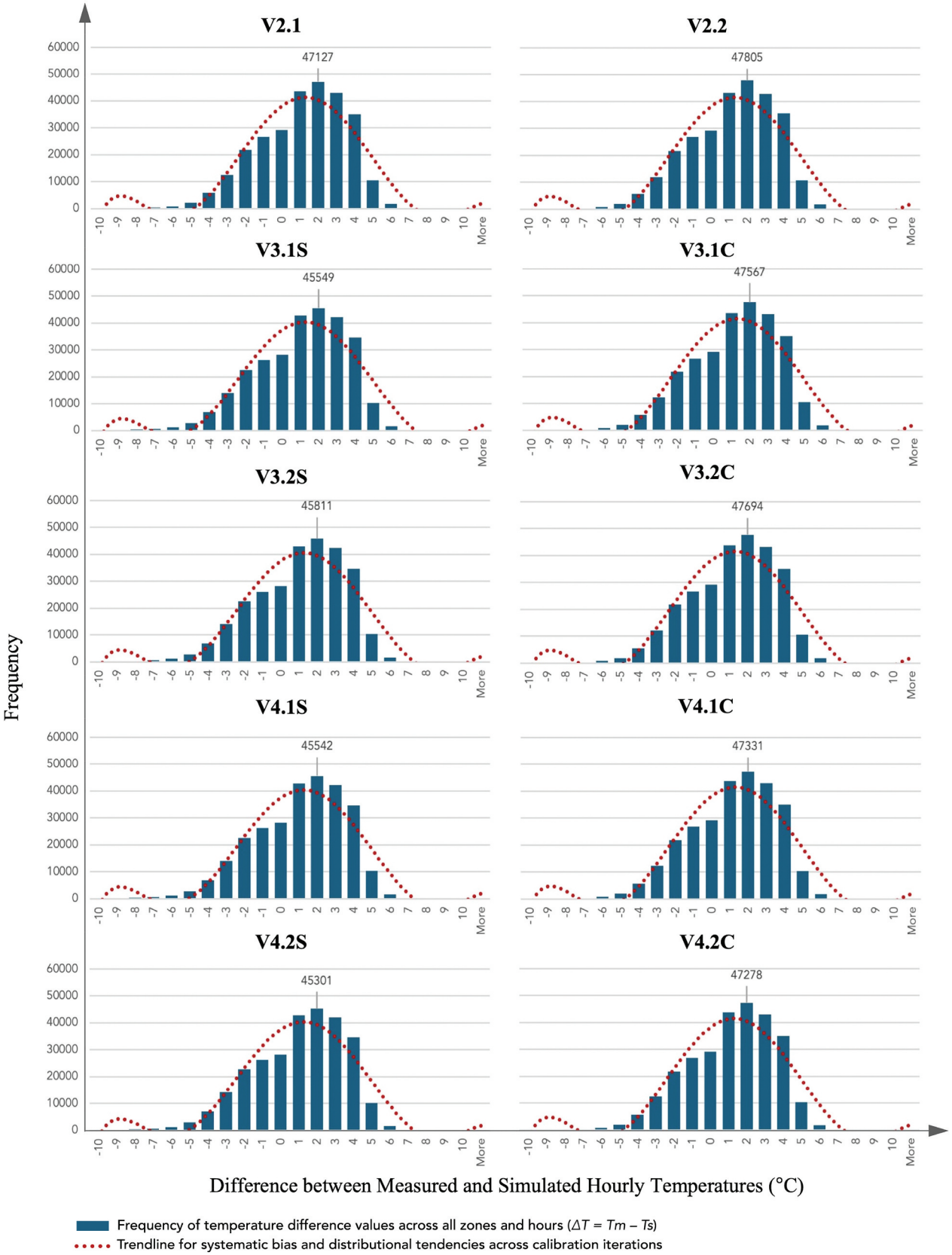


Figure 8. Frequency distribution of hourly temperature differences ($\Delta T = Tm - Ts$) across iterative calibration runs V2.1 to V4.2C.

Table 7. Sensitivity analysis of behavior-integrated calibration iterations (*V2.1 – V4.2C.*).

Model Code		Modified Parameters	Reference	Mean Hourly Correlation (T, r^2)	Mean Hourly Temperature Difference (ΔT, °C)	Hourly Indoor Temperatures		Monthly Heating Consumption			
						NMBE (%)	CV(RMSE) (%)	NMBE (%)	CV(RMSE) (%)	Sensitivity (%)	
V0	**V0.0**	Measured U-values / on-site weather data / thermostat setpoints / default internal gains / heating system properties	None	0.69	1.22	5.24	16.70	-15.16	-26.42		
V1	**V1.9**	Calculated infiltration coefficients/ -5.0% correction to U-values / default internal gains / ±0.50°C upper and lower limits to monitored hourly temperatures	V1.7	0.87	0.53	2.26	10.82	1.39	14.21		**In compliance with ASHRAE 14**
V2	**V2.1**	Measured occupant presence	V1.9	0.88	0.65	2.73	10.51	3.96	14.46	+2,57 to V1.9	
	V2.2	Measured and extrapolated occupant presence	V1.9	0.88	0.65	2.80	10.47	4.19	14.52	+2,80 to V1.9	
V3	**V3.1S**	Measured window operation behavior	V1.9	0.87	0.53	2.27	10.82	1.76	14.17	+0,37 to V1.9	
	V3.1C		V2.1	0.88	0.65	2.80	10.46	4.55	14.58	+3,16 to V1.9 +0,59 to V2.1	
	V3.2S	Measured and extrapolated window operation behavior	V1.9	0.87	0.54	2.41	10.80	4.24	13.52	+2.85 to V1.9	
	V3.2C		V2.2	0.88	0.66	2.83	10.45	7.05	14.52	+5.66 to V1.9 +2.86 to V2.2	
V4	**V4.1S**	Measured lighting use behavior	V1.9	0.88	0.53	2.23	10.82	0.72	14.21	-0.67 to V1.9	
	V4.1C		V3.1C	0.88	0.65	2.68	10.46	3.90	14.41	+2.51 to V1.9 -0,65 to V3.1C	
	V4.2S	Measured and extrapolated lighting use behavior	V1.9	0.87	0.52	2.40	10.83	0.54	14.25	-0.85 to V1.9	
	V4.2C		V3.1C	0.87	0.52	2.75	10.47	3.72	14.40	-0.83 to V3.1C	

unmonitored spaces. The baseline assumed continuous occupancy during work hours, resulting in constant heat gains. Data showed actual occupancy was more stochastic and lower than standard assumptions (Figure 7).

The error distribution for V2.1 (Figure 8) shows the impact of reduced internal gains compared to the reference model. It remains stable and leptokurtic, peaking at 47,127 hours at 2.00°C, reflecting V1.9's thermal profile. Adding measured presence shifted energy predictions, with NMBE rising from 1.39% (V1.9) to 3.96%, a + 2.57% sensitivity (Table 7). This shift and positive ΔT bias suggest V1.9's calibration compensated for envelope heat losses using overestimated internal gains from 2004 schedules, replacing continuous occupant presence with intermittent presence, which increased heating system compensation and predicted energy use. Model V2.2 expanded the measured data by applying extrapolated monitored occupancy to unmonitored zones based on functional and density similarities to test the sensitivity of the whole-building occupant density. The error distribution for V2.2 (Figure 8) was similar to V2.1, peaking at 47,805 hours at 2.00°C. Extrapolating to unmonitored zones further reduced internal gains, increasing the V2.1 consumption discrepancy to an NMBE of 4.19% (Table 7), a total increase of 2.80% or 4.74 kWh/m^2. Using realistic presence schedules revealed a higher heating demand that had previously been obscured by the inflated internal gains assumed in ASHRAE 90.1. In similar studies, replacing static occupancy schedules with high-resolution, monitored presence typically reduced hourly CV (RMSE) by 5–20% in office and educational buildings (Gunay, O'Brien, and Beausoleil-Morrison 2015; Sayfikar and Jenkins 2021; Tahmasebi and Mahdavi 2015).

3.3.2. Sensitivity to adaptive window operation (V3.1S to V3.2C)

Adaptive window operation was evaluated to assess the model's sensitivity to ventilation-driven heat loss. Window operation was monitored as a binary state (0 = closed, 1 = open) using HOBO UX90-001x state change dataloggers installed on operable windows. The devices recorded time-stamped opening and closing events, allowing the window opening duration to be derived from the elapsed time between successive state changes.

For visualization and for compatibility with hourly simulation inputs, the event-based binary series were converted into hourly fractional values. For each hour, the total duration the window remained open was calculated as a time fraction ranging from 0 (window fully closed per hour) to 1 (window fully open per hour). Intermediate values represent partial opening durations; for example, a value of 0.2 corresponds to a 12-minute window opening (Figure 9). When an opening event spanned multiple hours, the open duration was allocated proportionally to the corresponding hours. Figure 9 presents these hourly fractions, reflecting measured open-time proportions rather than instantaneous window states.

In V3 iterations, the model's sensitivity to ventilation-driven heat loss was assessed. Monitoring data from eight offices revealed pronounced inter-occupant variability in window operation behavior. For example, the occupant in Office 101 frequently kept the 0.60 × 0.60 m operable window open (Figure 9), whereas the occupant in Office 106 rarely opened it. Window operation behavior was evaluated as a single parameter with the suffix S (as a behavioral version of V1.9) and as a combined parameter with the suffix C (as a behavioral version of the presence-calibrated models V2.1 and V2.2).

Models V3.1S and V3.1C used monitored opening patterns in eight offices (Figure 2) to assess the effects of natural ventilation on thermal performance. The error distribution for V3.1S resembled V1.9, with a peak at 2.00°C and 45,549 hours. V3.1C, based on V2.1, had a similar peak at 2.00°C and 47,567 hours (Figure 8). Error distributions were stable, but the consumption data were sensitive to ventilation losses. In V3.1S, the monthly consumption NMBE changed by +0.37% (from 1.39% to 1.76%; Table 7). In V3.1C, reduced internal gains caused the error to increase by +0.59% compared to V2.1 (3.96% to 4.55%; Table 7). The limited magnitude of this shift is due to the building's physical characteristics. The

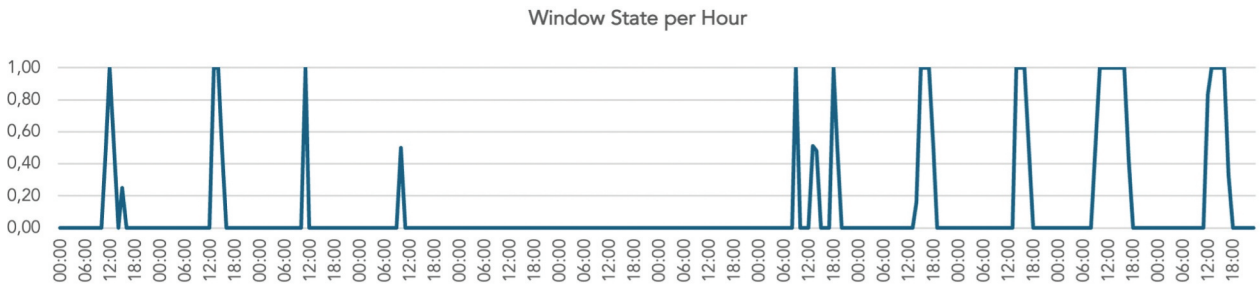


Figure 9. Hourly adaptive window operation fractions derived from state-change monitoring for the occupant in office 101 (Jan 7th to 18th, 2019).

operable window area in the monitored zones was small (0.36 m^2), which slightly increased heating system consumption.

Models V3.2S and V3.2C extrapolated monitored window state probabilities to other building zones based on OB to assess the sensitivity of the whole-building energy balance to adaptive ventilation. The error distribution for these extrapolated models showed that the heating system effectively offset the thermal effects of whole-building adaptive ventilation. V3.2S had a slight peak increase to 45,811 hours at 2.00°C versus V1.9, while V3.2C maintained 47,694 hours (Figures 5 and 7). While the histograms indicated indoor temperature stability between iterations, window operation influenced consumption profiles. Extrapolating monitored window operation in V3.2S increased the heating consumption NMBE to 4.24%, with a sensitivity of +2.85% compared to the baseline. The combined effect of lower occupant sensible gains and higher ventilation losses (V3.2C) shifted the NMBE to 7.05%, reflecting a substantial net sensitivity of +2.86% compared to V2.2 (Table 7).

For V3 iterations, error distributions showed that the iterations reliably tracked indoor temperature setpoints and indicated an increase in end consumption to offset the ingress of outdoor air temperature caused by natural ventilation. This reveals the building's actual thermal behavior, which the conventionally calibrated reference model (V1.9) lacked. Ignoring adaptive ventilation behaviors may lead to a systematic underestimation of consumption. As the number of monitored zones with active window operation increases, the discrepancy may become even more significant (Penna et al. 2015).

3.3.3. Sensitivity to adaptive lighting use (V4.1S to V4.2C)

Monitored lighting use schedules were incorporated to assess additional internal gains, a key factor in buildings with dated lighting systems. Lighting use was monitored as a binary state variable (1 = on, 0 = off) using one-minute interval loggers, installed in each zone as reported in Table 2. The devices recorded the lighting state at each minute, allowing the total duration of lighting use behavior to be derived for each hour. For visualization and compatibility with hourly simulation inputs, the one-minute binary series was aggregated into hourly fractional values. For each hour, the number of minutes the lights were on was summed and divided by 60, yielding an hourly lighting-use fraction ranging from 0 (lights off for the entire hour) to 1 (lights on for the entire hour). Intermediate values represent partial-hour usage (e.g., 0.3 corresponds to approximately 18 minutes of lighting use). Figure 10 presents these hourly fractions, which reflect the proportion of lighting on-time rather than the instantaneous lighting state.

Monitoring data from 17 zones indicated that lighting use in active spaces, such as offices and studios, frequently extended beyond the duration or intensity predicted by baseline schedules. For example, Figure 10 illustrates lighting use in the open studio during and after class hours when students were present. Lighting use behavior (0 = on, 1 = off) was tested as a single parameter (suffix S) as a behavioral iteration of V1.9, and as a combined parameter (suffix C) as a behavioral iteration of the presence-and-window-operation-calibrated models (V31.C).

Models V4.1S and V4.1C incorporate monitored lighting data, isolating the direct impact of lighting sensible heat gains. Error distributions for both the single (V4.1S) and combined (V4.1C) iterations showed that lighting gains did not affect overall indoor temperature fluctuations (Figure 8). Unlike presence and window operation, which increased predicted heating consumption (positive sensitivity), lighting use behavior demonstrated a negative sensitivity. In V4.1S, incorporating measured lighting decreased the heating consumption NMBE by 0.67%, reducing it from 1.39% (in V1.9) to 0.72% (Table 7). In the combined model (V4.1C), lighting use behavior lowered the total error from 4.55% (in V3.1C) to 3.90%, a reduction of 0.65%. This consistent negative sensitivity confirms that lighting acts as a corrective internal heat source. The measured data indicated that actual lighting use

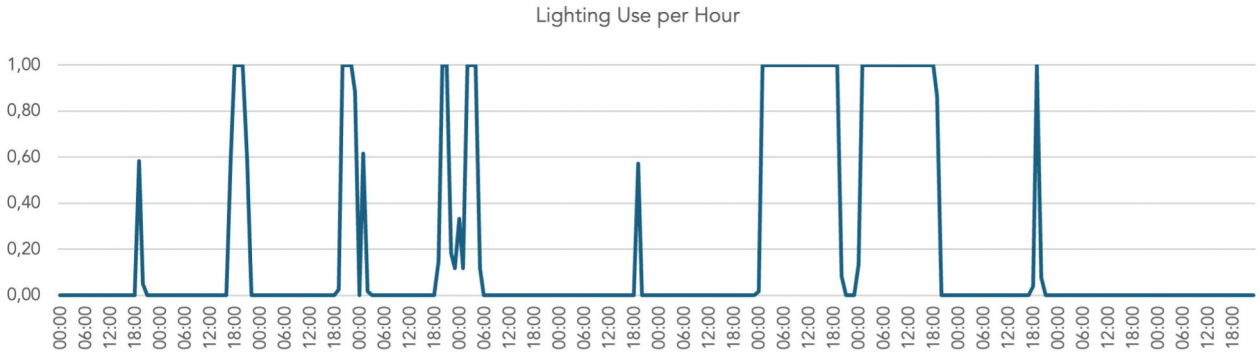


Figure 10. Hourly lighting use behavior fractions in the open studio (1–12 April 2019), derived from one-minute binary monitoring (1 = on, 0 = off).

exceeded baseline assumptions, providing internal gains that offset part of the heating demand caused by occupant density and ventilation losses.

Models V4.2S and V4.2C extrapolated the monitored lighting use behavior to unmonitored zones, testing the sensitivity of the whole-building energy balance to widespread lighting gains. The error distributions confirmed that the building's thermal mass and HVAC system effectively buffered the increase in internal gains, maintaining indoor temperature profiles in V4.2S and V4.2C (Figure 8). These iterations showed the highest sensitivity to lighting gains and further reduced the predicted heating consumption. Extrapolating measured lighting use as a single parameter (V4.2S) decreased the heating consumption NMBE to 0.54%, a net reduction of 0.85% compared to the conventional reference model (Table 7). In the final fully integrated model, which included measured and extrapolated presence, window operation, and lighting use behavior (V4.2C), the heating consumption error decreased by 0.83% relative to V3.1C, resulting in a final NMBE of 3.72%. This highlights the importance of lighting use as a significant, distributed heat source in educational buildings. Accurately modeling these gains reduces predicted heating needs, partially offsetting load increases caused by realistic occupancy and ventilation modeling, resulting in a more physically coherent energy balance.

4. Synthesis and generalization

The iterative calibration workflow demonstrated that meeting the statistical standards of ASHRAE Guideline 14 and IPMVP does not automatically ensure the physical accuracy of a building energy simulation (BES) model. By carefully separating behavioral inputs from physical parameters, this study uncovered a key phenomenon of compensating errors that often appear during manual calibration.

The conventional calibration phase ended with V1.9, which reached the highest statistical accuracy (NMBE = 1.39%, CV (RMSE) = 14.21%; Table 6). However, adding measured occupant presence in V2.1 increased the energy prediction error to 3.96%. This change showed that the conventional model had achieved accuracy by balancing two opposing errors.

(1) Overestimated Internal Gains: The baseline ASHRAE 90.1 schedules assumed continuous occupancy and higher density, artificially inflating the internal heat generation.
(2) Artificially Tight Envelope: To counterbalance these high gains and align with utility bills, the calibration process required tightening the infiltration rates (F = 0.30) and reducing U-values by 5% in V1.9 (Table 6).

When realistic, lower density occupancy data were incorporated in V2, the artificial surplus of internal heat was eliminated. This revealed the underlying thermal behavior: the building envelope was probably more permeable than the V1.9 calibration assumed. The further integration of window operation behavior (V3.2C) acted as a load multiplier, increasing ventilation heat losses and raising the NMBE to 7.05% (Table 7). This demonstrates

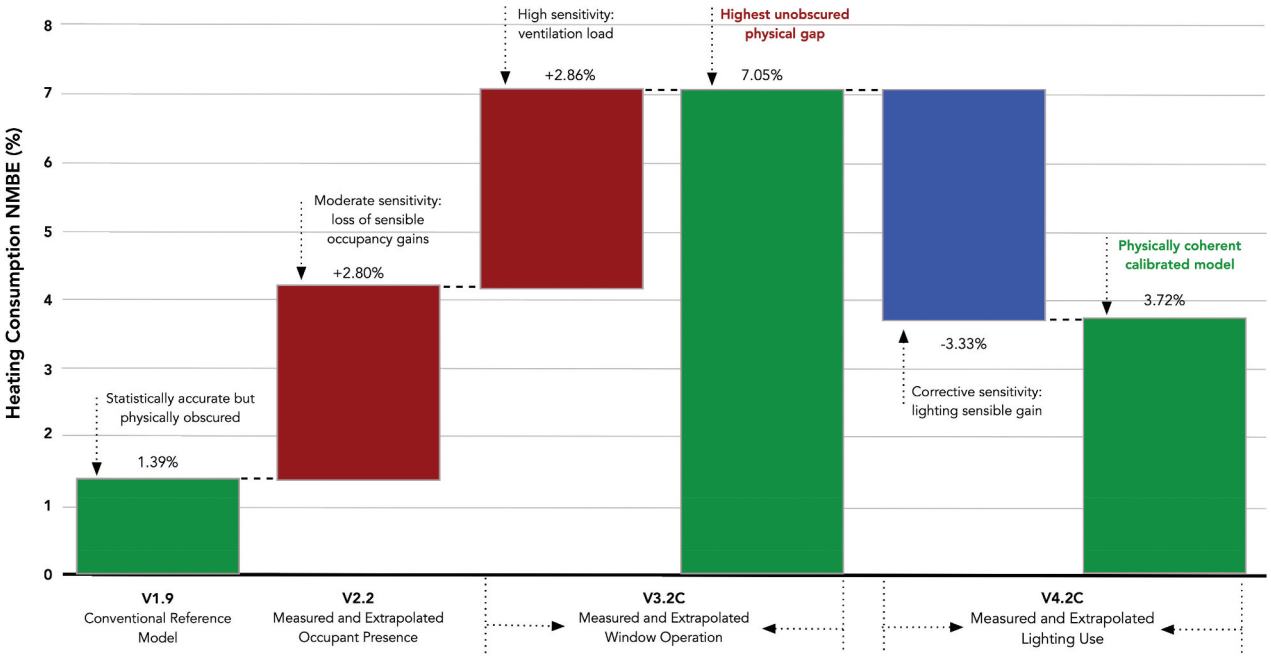


Figure 11. Deconstruction of compensating errors: the impact of behavior integration on heating consumption calibration.

that the conventionally calibrated model masked the building's actual ventilation load behind unrealistic occupant density.

The study also establishes a hierarchy of sensitivity for educational buildings with central heating systems. The variation in NMBE across the behavior-integrated models (0.54% to 7.05%; Table 7, Figure 11) serves as a sensitivity index:

- High Sensitivity: The interaction between reduced internal gains and adaptive ventilation (ventilation-occupancy interaction, V3.2C) caused the most significant deviation (+5.66% in sensitivity compared to V1.9). This highlights adaptive window operation as the primary stochastic variable influencing heating consumption in older, naturally ventilated buildings.
- Moderate Sensitivity: Non-adaptive occupant presence alone contributed a sensitivity of +2.57% to +2.80%. Therefore, accurate occupancy modeling is essential to determine the actual thermal behavior baseline of the simulated building.
- Corrective Sensitivity: Adaptive lighting use behavior showed a negative sensitivity (approximately −0.65% to −0.85%), acting as a distributed internal gains source that partially offset the increased ventilation loads (Table 7, Figure 11).

The difference between the conventionally calibrated model (V1.9) and the behavior-integrated model (V4.2C) indicates a need to improve conventional calibration procedures. A model that incorporates measured behavior (V4.2C) offers a more physically consistent foundation for retrofitting, even though its initial error metrics are higher (NMBE = 3.72%) than those of the conventional model. For accurate retrofitting evaluations, the calibration process must focus on behavior-based parameter adjustments. The results suggest that a final tuning of physical parameters, particularly for the infiltration coefficients and the original measured U-values, would be required after V4.2C to balance the energy equation without relying on compensating errors.

5. Conclusion

This study presented a systematic, evidence-based manual calibration framework applied to the M1 Building at the Architecture Department of Eskisehir Osmangazi University, Turkey. By shifting from a deterministic baseline to a behavior-integrated simulation, the research quantified the separate effects of adaptive and non-adaptive occupant behaviors on the model's sensitivity.

The uncalibrated baseline significantly underestimated the building's heating consumption (NMBE = −15.16%), confirming that relying only on design-stage assumptions and standard operating schedules is inadequate for evaluating operational performance. Although traditional calibration with infiltration and U-value adjustments achieved high statistical accuracy (NMBE = 1.39%), incorporating monitored data showed that the accuracy was partly due to inflated occupancy gains to compensate for envelope heat losses.

The stepwise integration of behavioral datasets revealed the following key insights:

(1) Occupancy as a Compensating Factor: Using measured data instead of standard ASHRAE presence profiles increased the predicted heating consumption by about 2.80%. This shows that standard schedules often overestimate internal gains, hiding the real insulation and airtightness issues in existing buildings.
(2) Adaptive Ventilation as a Load Multiplier: Window operation behavior, when applied to the entire building, acted as a significant load multiplier, raising energy use by over 2.8% compared to the occupancy-calibrated model. Ignoring this adaptive behavior leads to a systematic underestimation of the heating required to condition outdoor air entering the building.
(3) Lighting use as a Corrective Gain: Conversely, measured lighting use was higher than assumed, providing additional heat that reduced the heating load error by 0.83% in the fully integrated model.

The present study concludes that meeting statistical benchmarks (ASHRAE Guideline 14 or IPMVP) is a necessary but not sufficient condition for model validity. A model that complies statistically can still be physically inconsistent if behavioral inputs are standardized rather than directly measured. To ensure simulation models serve as a reliable baseline for energy conservation measures (ECMs), calibration must explicitly incorporate monitored behavioral data to prevent compensating errors. A limitation of this study was its reliance on calculated infiltration rates (ATTMA standards) instead of physical blower-door testing, which was not feasible due to the building's educational use. Future research should verify these calculated air exchange rates through empirical methods such as tracer gas or pressurization tests. Additionally, this framework can be expanded by integrating agent-based modeling to dynamically simulate stochastic occupant interactions, reducing dependence on deterministic schedules.

Acknowledgements

The author would like to thank Orkun Alptekin and Hasan Unver for their assistance during the monitoring phase. As they were not involved in the data analysis, simulation, calibration, or article preparation stages of the study, they have formally waived authorship rights for this manuscript. The author also

acknowledges the use of OpenAI's ChatGPT 5.1 to enhance the clarity, structure, and language of the text, under the author's supervision, who maintains full responsibility for the final content of the paper.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The present study was funded by the Eskisehir Osmangazi University Scientific Research Projects Coordination Unit, with Project ID [308] and Project Code [201915015], titled *"Calibration of Building Energy Models Used in Energy Performance Retrofit: An Education Building Case Study."*

References

Anandand, P., D. Cheong, and C. Sekhar. 2022. "A Review of Occupancy-Based Building Energy and IEQ Controls and its Future Post-COVID." *Science of the Total Environment* 804:150249. https://doi.org/10.1016/j.scitotenv.2021.150249.

Angelottigelotti, A., L. Mazzarella, C. Cornaro, F. Frasca, A. Prada, P. Baggio, I. Ballarini, G. De Luca, and V. Corrado. 2023. "Calibrating the Dynamic Energy Simulation Model for an Existing Building: Lessons Learned from a Collective Exercise." *Energies* 16 (7): 2979. https://doi.org/10.3390/en16072979.

Ascionecione, F., M. Borrelli, R. F. De Masi, F. de' Rossi, and G. P. Vanoli. 2019. "Energy Refurbishment of a University Building in Cold italian Backcountry. Part 2: Sensitivity Studies and Optimization." *Energy Procedia* 159:10–15. https://doi.org/10.1016/j.egypro.2018.12.010.

ASHRAE. 2017. *ASHRAE handbook: Fundamentals*. Atlanta, Georgia: American Society of Heating, Refrigerating and Air-Conditioning Engineers.

ASHRAE 55. 2010. *Thermalrmal Environmental Conditions for Human Occupancy (ANSI/ASHRAE Standard 55-2010)*. Atlanta, Georgia: American Society of Heating, Refrigerating and Air-Conditioning Engineers.

ASHRAE Guideline 14. 2014. *Measurement of Energy and Demand Savings*. Atlanta: American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc.

ASHRAE Standard 90.1. 2004. *Energy Standard for Buildings Except Low-Rise Residential Buildings*.

Astete, N., R. S. Adhikari, M. Buzzetti, C. Del pero, H. E. Huerto-Cardenas, F. Leonforte, and A. Miglioli. 2022. "nZEB: Bridging the Gap Between Design Forecast and Actual Performance Data." *Energy and Built Environment* 3 (1): 16–29. https://doi.org/10.1016/j.enbenv.2020.10.001.

ATTMA TSL2. 2010. *Technical Standard L2: MEASURING AIR PERMEABILITY OF BUILDING ENVELOPES (Non-Dwellings)*. 1 Spencer Parade, Northampton: The Air Tightness Testing and Measurement Association c/o the British Institute of Non-Destructive Testing.

Bahadori-Jahromi, A., R. Salem, A. Mylona, A. U. Hasan, and H. Zhang. 2022. "The Effect of Occupants' Behaviour on the Building Performance Gap: UK Residential Case Studies." *Sustainability* 14 (3): 1362. https://doi.org/10.3390/su14031362.

Bălănescu, D. T., and V. M. Homutescu. 2018. "Experimental Investigation on Performance of a Condensing Boiler and Economic Evaluation in Real Operating Conditions." *Applied thermal Engineering* 143:48–58. https://doi.org/10.1016/j.applthermaleng.2018.07.082.

Balaras, C. A. 2022. "Building Energy Audits—Diagnosis and Retrofitting Towards Decarbonization and Sustainable Cities." *Energies* 15 (6): 2039. https://doi.org/10.3390/en15062039.

Borodinecs, A., A. Palcikovskis, A. Krumins, D. Zajecs, and K. Lebedeva. 2024. "Assessments of HVAC performance and Savings in Office Buildings Using Data-Driven Method." *Clean Technologies* 6 (2): 802–813. https://doi.org/10.3390/cleantechnol6020041.

Buso, T., V. Fabi, R. K. Andersendersen, and S. P. Corgnati. 2015. "Occupant Behaviour and Robustness of Building Design." *Building & Environment* 94:694–703. https://doi.org/10.1016/j.buildenv.2015.11.003.

Carlucci, S., F. Causone, S. Biandrate, M. Ferrando, A. Moazami, and S. Erba. 2021. "On the Impact of Stochastic Modeling of Occupant Behavior on the Energy Use of Office Buildings." *Energy & Buildings* 246:111049. https://doi.org/10.1016/j.enbuild.2021.111049.

Chaudhary, G., J. New, J. Sanyal, P. Im, Z. O'Neill, and V. Garg. 2016. "Evaluation of "Autotune" Calibration Against Manual Calibration of Building Energy Models." *Applied Energy* 182:115–134. https://doi.org/10.1016/j.apenergy.2016.08.073.

Chen, W. A., Y. H. Wang, H. J. Chang, and R. L. Hwang. 2024. "Development of a Rapid Assessment Tool for Integrating Thermal Comfort in Early Design Stage of Energy-Efficient Office Buildings." *Applied Energy* 363:123072. https://doi.org/10.1016/j.apenergy.2024.123072.

Chiang, C.-L., and J. Calautit. 2025. "Reducing Building Energy Performance Gap: Integrating Agent-Based Modelling and Building Performance Simulation." *Buildings* 15 (10): 1728. https://doi.org/10.3390/buildings15101728.

Chong, A., Y. Gu, and H. Jia. 2021. "Calibrating Building Energy Simulation Models: A Review of the Basics to Guide Future Work." *Energy & Buildings* 253:111533. https://doi.org/10.1016/j.enbuild.2021.111533.

Chong, A., D. Yan, K. Sun, S. Zhan, S. Cheng, Y. Wu, Y. Chen, and T. Hong. 2025. "Ten Questions Concerning Calibrating Building Energy Simulation Models." *Building & Environment* 284:113404. https://doi.org/10.1016/j.buildenv.2025.113404.

Cipriano, J., G. Mor, D. Chemisana, D. Pérez, G. Gamboa, and X. Cipriano. 2015. "Evaluation of a Multi-Stage Guided Search Approach for the Calibration of Building Energy Simulation Models." *Energy & Buildings* 87:370–385. https://doi.org/10.1016/j.enbuild.2014.08.052.

Coakley, D., P. Raftery, and M. Ve Keane. 2014. "A Review of Methods to Match Building Energy Simulation Models to Measured Data." *Renewable and Sustainable Energy Reviews* 37:123–141. https://doi.org/10.1016/j.rser.2014.05.007.

Copiello, S. 2017. "Building Energy Efficiency: A Research Branch Made of Paradoxes." *Renewable and Sustainable Energy Reviews* 69:1064–1076. https://doi.org/10.1016/j.rser.2016.09.094.

Cornaro, C., F. Bosco, M. Lauria, V. A. Puggioni, and L. De Santoli. 2019. "Effectiveness of Automatic and Manual Calibration of an Office Building Energy Model." *Applied Sciences* 9 (10): 1985. https://doi.org/10.3390/app9101985.

Dakyen, M. M., M. Dagbasi, and M. Özdenefe. 2021. "Energy Models for Cost-Optimal analysis: Development and Calibration of Residential Reference Building Models for northern Cyprus." *Indoor and Built Environment* 31 (3): 657–681. https://doi.org/10.1177/1420326X211013076.

EDSL. 2026. "Tas Engineering - Environmental Design Solutions Limited." https://www.edsltas.com/tas-engineering/.

Esmaeilzadeh, A., and M. Hamdy. 2025. "Dynamic Building performance Gap Related to Occupant Behaviors: A Framework to Predict the Impacts of Possible Causal Factors." *Energy & Buildings* 347 (A): 116340. https://doi.org/10.1016/j.enbuild.2025.116340.

Fabi, V., R. V. Andersendersen, S. P. Corgnati, and B. W. Olesen. 2013. "A Methodology for Modelling Energy-Related Human Behaviour: Application to Window Opening Behaviour in Residential Buildings." *Building Simulation* 6 (4): 415–427. https://doi.org/10.1007/s12273-013-0119-6.

Gilani, S., and W. O'Brien. 2019. "Exploring the Impact of Office Building Users' Modeling Approaches on Energy Use Across Canadian Climates." *Energy & Buildings* 197:68–86. https://doi.org/10.1016/j.enbuild.2019.05.042.

Gucyeter, B. 2018. "Calibration of a Building Energy performance Simulation Model via Monitoring Data." In *Proceedings of the Building Performance Analysis Conference and SimBuild*, 542–549. Chicago, IL, USA: American Society for Heating, Refrigerating and Air Conditioning Engineers. https://publications.ibpsa.org/conference/paper/?id=simbuild2018_C074.

Gunay, H. B., W. O'Brien, and I. Beausoleil-Morrison. 2013. "A Critical Review of Observation Studies, Modeling, and Simulation of Adaptive Occupant Behaviors in Offices." *Building & Environment* 70:31–47. https://doi.org/10.1016/j.buildenv.2013.07.020.

Gunay, H. B., W. O'Brien, and I. Beausoleil-Morrison. 2015. "Development of an Occupancy Learning Algorithm for Terminal Heating and Cooling Units." *Building & Environment* 93 (2): 71–85. https://doi.org/10.1016/j.buildenv.2015.06.009.

Gunay, H. B., W. O'Brien, I. Beausoleil-Morrison, and B. Huchuk. 2014. "On Adaptive Occupant-Learning Window Blind and Lighting Controls." *Building Research & Information* 42 (6): 739–756. https://doi.org/10.1080/09613218.2014.895248.

Guyot, D., F. Giraud, F. Simon, D. Corgier, C. Marvillet, and B. Tremeac. 2020. "Building Energy Model Calibration: A Detailed Case Study Using Sub-Hourly Measured Data." *Energy & Buildings* 223:110189. https://doi.org/10.1016/j.enbuild.2020.110189.

Heo, Y., R. Choudhary, and G. A. Augenbroe. 2012. "Calibration of Building Energy Models for Retrofit Analysis Under Uncertainty." *Energy & Buildings* 47:550–560. https://doi.org/10.1016/j.enbuild.2011.12.029.

Hong, T., S. C. Taylor-Lange, S. D'Oca, D. Yan, and S. P. Ve Corgnati. 2016. "Advances in Research and Applications of Energy-Related Occupant Behavior in Buildings." *Energy & Buildings* 116:694–702. https://doi.org/10.1016/j.enbuild.2015.11.052.

IEA. 2013. *IEA ECBCS annex 53: Annex 53 total Energy Use in Buildings: Analysis and Evaluation Methods*. International Energy Agency.

IPMVP. 2012. "International Performance Measurement and Verification Protocol." In *Concepts and Options for Determining Energy and Water Savings*. Vol. 1. Toronto, Canada: Efficiency Valuation Organization. https://www.eeperformance.org/uploads/8/6/5/0/8650231/ipmvp_volume_i__2012.pdf.

Jia, M., and R. Srinivasan. 2020. "Building Performance Evaluation Using Coupled Simulation of EnergyPlus™ and an Occupant Behavior Model." *Sustainability* 12 (10): 4086. https://doi.org/10.3390/su12104086.

Jia, M., R. Srinivasan, R. J. Ries, G. Bharathy, and N. Weyer. 2021. "Investigatingvestigating the Impact of Actual and Modeled Occupant Behavior information Input to Building Performance Simulation." *Buildings* 11 (1): 32. https://doi.org/10.3390/buildings11010032.

Kang, E., H. Lee, J. Yoon, H. Cho, C. Chaichana, and D. Kim. 2024. "Investigating the Influence of Uncertainty on Office Building Energy Simulation Through Occupant-Centric Control and Thermal Comfort Integration." *Energy & Buildings* 322:114741. https://doi.org/10.1016/j.enbuild.2024.114741.

Khabir, S., and R. Vakilinezhad. 2023. "Energy and Thermal Analysis of DSF in the Retrofit Design of Office Buildings in Hot Climates." *Architectural Engineering and Design Management* 19 (6): 642–664. https://doi.org/10.1080/17452007.2022.2147898.

Kim, C. H., M. Kim, and Y. J. Song. 2021. "Sequence-To-Sequence Deep Learning Model for Building Energy Consumption Prediction with Dynamic Simulation Modeling." *Journal of Building Engineering* 43:102577. https://doi.org/10.1016/j.jobe.2021.102577.

Kim, Y. K., L. Bande, K. A. Tabet Aoul, and H. Altan. 2021. "Dynamic Energy Performance Gap Analysis of a University Building: Case Studies at UAE University Campus, UAE." *Sustainability* 13:120. https://doi.org/10.3390/su13010120.

Menezes, A. C., A. Cripps, D. Bouchlaghem, and R. Buswell. 2012. "Predicted Vs. Actual Energy Performance of Non-Domestic Buildings: Using Post-Occupancy Evaluation Data to Reduce the Performance Gap." *Applied Energy* 97:355–364. https://doi.org/10.1016/j.apenergy.2011.11.075.

Mitra, D., Y. Chu, and K. Cetin. 2021. "Cluster Analysis of Occupancy Schedules in Residential Buildings in the United States." *Energy & Buildings* 236:110791. https://doi.org/10.1016/j.enbuild.2021.110791.

Motuziene, V., and T. Vilutiene. 2013. "Modelling the Effect of the Domestic Occupancy Profiles on Predicted Energy Demand of the Energy Efficient House." *Procedia Engineering* 57:798–807. https://doi.org/10.1016/j.proeng.2013.04.101.

Norouziaslouziasl, S., and A. Jafari. 2023. "Identifying the Most influential Parameters in Predicting Lighting Energy Consumption in Office Buildings Using Data-Driven Method." *Journal of Building Engineering* 72:106590. https://doi.org/10.1016/j.jobe.2023.106590.

Penna, P., S. Schmiedhofer, F. Cappelletti, A. Gasparella, F. Tahmasebi, and A. Mahdavi. 2015. A Multi-Stage Approach to the Calibration of A School Building's Simulation Model." In Building Simulation 2015: 14th Conference of IBPSA, Hyderabad, India , edited by J. Mathur and V. Garg, 2825–2832. Wisconsin, US: International Building Performance Simulation Association. https://doi.org/10.26868/25222708.2015.2298.

Piselli, C., and A. L. Pisello. 2019. "Occupant Behavior Long-Term Continuous Monitoring Integrated to Prediction Models: Impact on Office Building Energy Performance." *Energy* 176:667–681. https://doi.org/10.1016/j.energy.2019.04.005.

Poel, B., G. van Cruchten, and C. A. Balaras. 2007. "Energy Performance Assessment of Existing Dwellings." *Energy & Buildings* 39 (4): 393–403. https://doi.org/10.1016/j.enbuild.2006.08.008.

Raftery, P., M. Keane, and A. Costa. 2009. "Calibration of a Detailed Simulation Model to Energy Monitoring System Data: A Methodology and Case Study." In *Proceedings of Building Simulation 2009, Eleventh IBPSA Conference*, Glasgow, Scotland. Wisconsin, USA: International Building Performance Simulation Association. https://doi.org/10.26868/25222708.2009.1199-1206.

Royapoor, M., and T. Ve Roskilly. 2015. "Building Model Calibration Using Energy and Environmental Data."

Energy & Buildings 94:109–120. https://doi.org/10.1016/j.enbuild.2015.02.050.

Sayfikar, M., and D. P. Jenkins. 2021. "The Impact of Using Actual Occupancy Data on Building performance Simulation Results for a Multi-Purpose University Building." In *Proceedings of the 17th IBPSA Conference*, edited by D. Saelens, J. Laverge, W. Boydens and L. Helsen. Vol. 3409. Bruges, Belgium: International Building Performance Association https://doi.org/10.26868/25222708.2021.30187 .

Sun, K., T. Hong, S. C. Taylor-Lange, and M. A. Piette. 2016. "A Pattern-Based Automated Approach to Building Energy Model Calibration." *Applied Energy* 165:214–224. https://doi.org/10.1016/j.apenergy.2015.12.026.

Tahmasebi, F., and A. Mahdavi. 2015. "The Sensitivity of Building Performance Simulation Results to the Choice of occupants' Presence Models: A Case Study." *Journal of Building performance Simulation* 10 (5–6): 625–635. https://doi.org/10.1080/19401493.2015.1117528.

Uddin, M. N., H.-H. Wei, H. L. Chi, and M. Ni. 2021. "Influencefluence of Occupant Behavior for Building Energy Conservation: A Systematic Review Study of Diverse Modeling and Simulation Approach." *Buildings* 11 (2): 41. https://doi.org/10.3390/buildings11020041.

Yang, T., Y. Pan, J. Mao, Y. Wang, and Z. Huang. 2016. "An Automated Optimization Method for Calibrating Building Energy Simulation Models with Measured Data: Orientation and a Case Study." *Applied Energy* 179:1220–1231. https://doi.org/10.1016/j.apenergy.2016.07.084.

Zare, N., S. M. E. Saryazdi, A. M. Bahman, A. Shafaat, and M. Sartipipour. 2023. "Investigationvestigation of Heating Energy performance Gap (EPG) in Design and Operation Stages of Residential Buildings." *Energy & Buildings* 301:113747. https://doi.org/10.1016/j.enbuild.2023.113747.