

Calibrated models for effective clustering: Discriminating operation schedules in occupied buildings

Karla Guerrero Ramírez^{1,§} (✉), Cristina Nuevo-Gallardo^{2,§}, Jesús Miguel Santamaría Ulecia¹, Beatriz Montalbán Pozas³, Carlos Fernández Bandera³

1. Instituto de Biodiversidad y Medioambiente BIOMA, Universidad de Navarra, Irunlarrea St. 1, Pamplona, 31008, Navarra, Spain

2. School of Architecture, Universidad de Navarra, Universidad St., Pamplona, 31009, Navarra, Spain

3. School of Technology of Cáceres, Universidad de Extremadura, Universidad Av., Cáceres, 10003, Extremadura, Spain

Abstract

European directives advocate for end-users to be aware of their energy consumption. However, individual energy monitoring tools, such as energy meters or cost allocators, are not always affordable or technically feasible to install. Therefore, the development of virtual tools that enable the study of energy consumption in existing buildings is necessary. Virtual sensors, particularly based on white-box models, offer the opportunity to recreate these behaviours. When calibrated with measured data, white-box models, which incorporate detailed building physics, become increasingly valuable for designing energy-efficient buildings. This research explores a novel approach to identifying building's load period directly from energy data generated by these calibrated models. The volume of data generated by white-box models can be overwhelming for visual analysis, but the hypothesis here is that analysing this data through clustering techniques can reveal patterns related to occupant behaviour and operational schedules. By feeding indoor temperature data into the calibrated model and analysing the resulting energy outputs, the research proposes a method to identify the heating, ventilation and air conditioning (HVAC) system operation schedule, free oscillation periods and non-recurrent events. Validation is achieved by comparing the identified periods with actual measured data. This methodology enables the development of a virtual sensor for cost allocation, which minimises the need for physical sensor deployment while complying with European Union directives. The research not only demonstrates high accuracy but also the potential to outperform measured schedule. This suggests the ability of the method to identify missing sensor data or other factors affecting temperature curves, enabling fault detection and diagnostics (FDD). Consequently, this opens doors for setting optimised operation schedules that balance energy efficiency with occupant comfort.

1 Introduction

The building sector is increasingly demanding more efficient solutions to combat climate change and the growing demand for energy resources. Improving the energy efficiency of buildings not only reduces energy consumption and greenhouse gas emissions, but also improves the comfort and quality of life of their occupants (EU 2023). European

Union Directive 2012/27/EU (EU 2012) introduced the requirement for end-user involvement in energy consumption. Recast by Directive 2023/2002/EU (EU 2023), it mandates the installation of individual energy meters for heating, cooling, and domestic hot water consumption in multi-apartment and multi-purpose buildings with central heating or cooling systems. While recognising the potential technical and cost-related challenges of measuring heat consumption,

Keywords

calibrated models
white-box
cluster
schedule
free oscillation
load

Article History

Received: 31 July 2024
Revised: 20 September 2024
Accepted: 08 October 2024

© The Author(s) 2024

E-mail: kguerrera@unav.es

[§] Karla Guerrero Ramírez and Cristina Nuevo-Gallardo contributed equally to this work.

the directive permits the use of heat cost allocators or alternative methods. In all instances, the devices must be remotely accessible.

The practical challenges of installing and maintaining energy meters or cost allocators in existing buildings, including factors such as pipework, sensor costs, and data interpretation, often render these solutions impractical. Consequently, exploring alternative energy allocation methodologies is essential to ensure widespread application while maintaining technical and economic viability. Building performance simulation (BPS) can contribute to the development of cost allocation methodologies, potentially reducing sensor costs while establishing a platform to enhance end-user energy consumption.

BPS has become a powerful technique in building design (Hensen and Lamberts 2019), enabling the creation of highly energy-efficient structures (Salvalai et al. 2024) that comply with the latest regulations and standards. Moreover, BPS offers various modelling approaches categorised as white-box, grey-box, or black-box, with the key distinction being the level of physical principles incorporated into the models (Arendt et al. 2018). White-box models fully represent building physics, yet simulations often exhibit discrepancies between simulated and measured performance (de Wilde 2014). The Post-occupancy Review of Buildings and their Engineering studies (PROBE) concluded that this discrepancy, denominated performance gap, can be twice the predicted energy use (Menezes et al. 2012).

The Chartered Institution of Building Services Engineers (CIBSE) has extensively studied and classified the performance gap in buildings, focusing on deviations arising from operational conditions and technical issues (CIBSE 2020). Nevertheless, even when building technical specifications are meticulously modelled and operational conditions are aligned with actual schedules, discrepancies can still occur. These discrepancies often stem from parameter uncertainty caused by inaccurate descriptions of building characteristics or overly simplified building physics (Jeong and Byon 2024).

Therefore, in order to obtain feasible performance data from a white-box model, it needs to undergo a calibration process to bridge the gap. The American Society of Heating, Refrigerating and Air Conditioning Engineers (ASHRAE) defined the calibration process as optimising model accuracy by comparing simulated outputs to actual measured data under identical conditions (ASHRAE 2014). A model is considered calibrated when the uncertainty between its predicted performance and the corresponding measured data falls below a threshold established by an international standard (e.g., ASHRAE, IPMVP, or CIBSE) (DOE 2015; CIBSE 2020; Herbringer et al. 2023).

The explicit representation of the physical building, its systems, and the environment allows calibrated white-box

models to simulate the building's thermodynamics and generate several outputs for analysis (Coakley et al. 2014). This capability goes beyond monitoring parameters, providing a deeper understanding of a building's performance.

However, the sheer volume of data generated by white-box model simulations can be overwhelming for visual analysis. Data-driven techniques like data mining are integral to analysing sensor data and enhancing fault detection and diagnostics (FDD) (Zhang et al. 2014b). By applying data mining techniques, valuable information can be extracted from the rich data produced by calibrated white-box models, even when historical data from sensors might be incomplete due to malfunctions or network complexities (Jeong et al. 2021).

Understanding and managing the data generated by a calibrated model is crucial, as it holds the potential to significantly optimise energy consumption. Data mining is a useful process to manage data, which involves identifying intriguing patterns, models, and various forms of knowledge within extensive datasets (Han et al. 2011). Research in this area has demonstrated the efficacy of data mining technologies in uncovering hidden patterns and knowledge within these vast datasets (Zhang et al. 2021a). Unsupervised data mining-based methods are the most suitable for this purpose, among which clustering-based and association rule mining-based methods have been extensively used in the building industry over the past two decades (Zhao et al. 2020).

Regarding clustering-based methods, these have been used, among other purposes, to identify building performance patterns. Panapakidis et al. (2014), Miller et al. (2015), and Nepal et al. (2020) approached this task through electrical load patterns, while Pieri et al. (2015) and Choksi et al. (2020) used energy consumption patterns. In the same way, Wu and Clements-Croome (2007) and Nikolaou et al. (2012) made use of heating load patterns and indoor environment distribution patterns, respectively. Clustering has also been used for FDD in building operation (Zhao et al. 2020). Among the clustering methods that can be employed, the most widely used is the *k*-means method as it allows working with unlabeled data and processing large amounts of data (Ikotun et al. 2023).

White-box models, combined with data clustering methods, also enable the development of virtual sensors (Yoon 2022). As described by Martin et al. (2021), virtual sensors are software-based tools that generate signals by analysing data from other sensors. These sensors are particularly useful in situations where physical measurement is impractical due to challenges in measurability, long-term monitoring, or economic cost (Yoon 2022).

Recent research has explored the application of virtual sensors in intelligent buildings, primarily focusing on control

and automation of heating, ventilation and air conditioning (HVAC) systems, FDD, and comfort analysis. Kim and Yoon (2023) proposed a virtual sensor for predictive mean vote (PMV) to enhance physical thermostat effectiveness for indoor thermal comfort, based on manually recorded operational data from two cassette air conditioners and mechanical ventilation systems. Verbet et al. (2017) developed a system-level HVAC fault diagnosis system using historical data and virtual sensors. Lee and Kim (2024) proposed a virtual sensor for FDD in variable refrigerant flow (VRF) systems, enabling the assessment of performance degradation ratios.

Based on the above context, this research aims to bridge the concepts of calibrated models with virtual sensing using data clustering techniques. The core hypothesis is that analysing indoor temperature measurements with a calibrated model, combined with clustering techniques applied to energy consumption data, can reveal valuable insights into energy usage patterns. This data can be used to create a virtual sensor for cost allocation, offering a practical solution for buildings where traditional sensors are either too expensive or technically impractical.

The methodology is based on the calibration of a white-box model with short periods of free oscillation monitored indoor temperature data. Clustering methods, in particular *k*-means, are then used to classify energy levels and differentiate between HVAC loads and free oscillation periods, as well as to identify anomalies caused by non-recurrent events such as window openings or individual radiator usage. Annual simulations are used to validate this research, although the simulation period could be modified and adjusted to align with typical energy billing cycles.

The application of this suggested methodology relies on indoor temperature sensors to be installed in rooms equipped with HVAC systems. With the rise of the Internet of Things (IoT), these sensors are now more affordable,

compact, remotely accessible and battery-operated, making them easy to deploy in large buildings (Yu et al. 2023). This setup complies with the remote access requirement set by Directive 2023/2002 /EU.

This approach is not only economical but also contributes to better energy management and efficiency strategies, ensuring that energy costs are accurately attributed to different zones or users within the building. By using calibrated models, data mining techniques and virtual sensors, the methodology offers a scalable solution that aligns with current regulatory frameworks while minimising the need for extensive sensor deployment. Ultimately, this results in more effective energy management and a pathway towards achieving the energy efficiency goals set by the European Union.

This research unfolds in distinct sections: Section 2 dives right in with a brief overview of the project description. It then delves deeper into the two crucial methodologies used in this research: calibration and clustering. The following section (Section 3) explores the outcomes of the clustering process. Here, the paper will compare these findings with the actual measured data. The analysis will particularly focus on how the clustering technique identifies HVAC system operation schedule, free oscillation periods and potentially uncovers other sources of load on the building. Finally, the last section (Section 4) wraps up the paper by presenting the main findings drawn from the analysis.

2 Methodology

This section is devoted to the methodology followed in this work, which is illustrated in Figure 1. The methodology consists of the following steps:

- Building energy model (BEM), developed in DesignBuilder, which is fed with the monitoring data.
- Calibration process, performed within EnergyPlus environment through genetic algorithm.

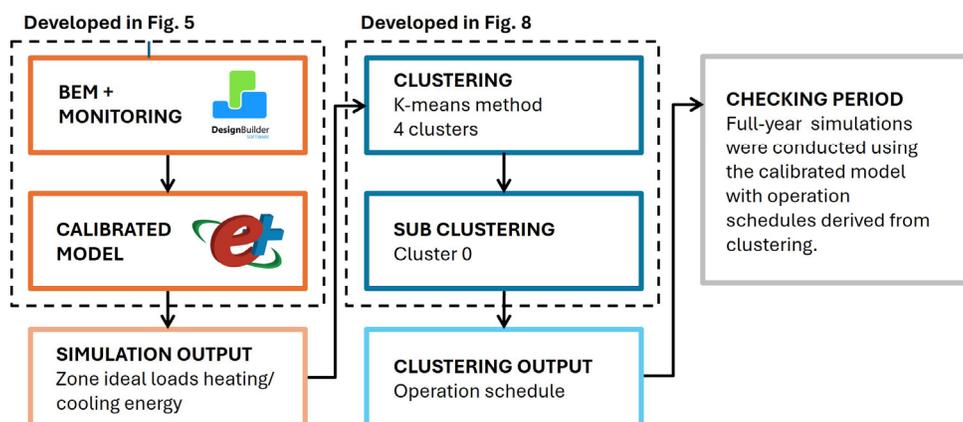


Fig. 1 Step-by-step methodology followed for this research

- c. Simulation of calibrated model obtaining energy outputs.
- d. Clustering of energy data generated by the calibrated model.
- e. Subclustering of the cluster with the lowest energy (cluster 0) to determine pure free oscillation.
- f. Full-year simulation, employing the calibrated model and derived operation schedules, was conducted to assess deviations from measured temperatures.

The calibration process and the clustering methodology are explained in more detail below.

2.1 Project description

The School of Technology of Cáceres (EPCC, from its Spanish abbreviation) is made up of six buildings, three of them being similar in architecture and structure, with this work focusing on Computer Science Pavilion (shown in Figure 2). This pavilion consists of a central hall open to the ground floor and the first floor, with classrooms, laboratories, and offices on both sides. Building envelope has been described by Montalbán Pozas et al. (2022). Due to its construction year, the pavilion adhered to the Spanish current regulation, Norma Básica de Edificación (NBE CT 79) (BOE 1979), which imposed less stringent energy efficiency standards, resulting in thinner insulation layers and increased infiltration. Concerning the heating and cooling systems, the HVAC system utilises two boilers to meet heating demands through radiators. On the other hand, only a limited number of rooms have individual air conditioning units for cooling.

Computer Science Pavilion is monitored with environmental, window opening, electricity, and gas consumption sensors, and actuators on radiators and

blinds. The monitoring of this building started with the SmartPolitech (Ministerio de Economía y Competitividad 2013) project in 2013 and the different types of sensors were installed progressively (Montalbán Pozas et al. 2022, 2023). For the environmental sensors, data is available from 2016 onwards, while electricity consumption sensors were in operation in the period between June 2017 and December 2020. The remaining sensors and actuators were installed from 2022 onwards.

With regard to the room on which this work is focused, room PB:033 is a computer laboratory located on the ground floor of the Computer Science Pavilion of 54.67 m². As illustrated in Figure 3, room PB:033 is situated on the western façade of the building, with only one exterior wall. The room's typical occupancy hours are from 9:00 to 17:00. Despite being designated as a computer laboratory, it is primarily used as a general classroom. The room is only equipped with seven computers, which is a relatively small number for a room of its size. Room PB:033 was chosen specifically due to its unique HVAC system. Unlike other rooms, it lacks a radiator and relies on an individual unit for both heating and cooling. The HVAC system operates on demand, responding to occupant comfort needs. This results in an irregular operation schedule with intermittent on/off cycles, short usage periods, and extended periods of inactivity. Comfort needs are often met through alternative methods like window opening or individual radiator use.

For environment monitoring, PB:033 is equipped with an environmental sensor that measures indoor temperature, relative humidity, and window opening (EcoWin, self-made), and an electricity consumption sensor connected to the air conditioning machine (Circutor Wibeec). Their technical specifications can be found in Table 1 together with the



Fig. 2 Overview of EPCC (Computer Science Pavilion marked in red box. Image downloaded from Google Earth (2024))

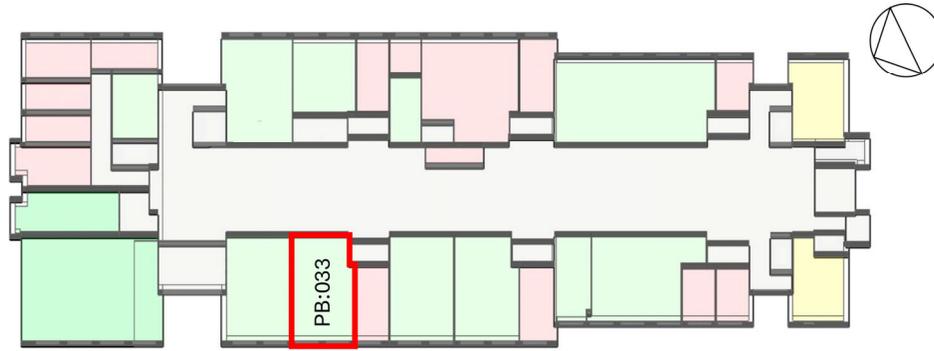


Fig. 3 EPCC ground floor. (PB:033 marked in red box. legend: laboratories and classrooms in green colour; offices or workrooms in orange; and bathrooms in yellow)

Table 1 Technical specifications of the sensors/actuators installed in the EPCC and used in this work (based on Montalbán Pozas et al. (2022, 2023))

Trademark/brand	Type of sensor	Communication/sampling interval	Parameter	Resolution	Accuracy
EcoWin (self-made)	Environmental sensor	WiFi/10'	Temperature	0.1 °C	±0.5 °C
			Humidity	0.1%	±3%
			Window opening	—	On/Off
Circutor Wibeec	Electricity consumption	WiFi/1'	Active power	2%	±4%
			Cumulative energy consumption	2%	±4%
MClimate Vicki LoRaWAN	Thermostatic valve	LoRaWAN/15'	Temperature	0.18 °C	±0.2 °C
			Humidity	0.39%	2%

radiator valve actuator information, MClimate sensor, which was used for the initial model calibration as detailed in Section 2.2.

Weather data was obtained from Agencia Estatal de Meteorología (AEMET) weather station located within 500 meters of EPCC. The station provided hourly monitoring data measuring various parameters, as detailed in Table 2. This data was employed to generate representative weather files in EPW format for the years 2017, 2018, 2019, and 2020.

2.2 Calibration methodology

An existent “as-built” Building Information Model (BIM)

was used to define the building’s architecture, materials and constructions. The load schedule was set according to the occupant schedule of the building. The location and height of surrounding buildings were obtained from an existent 3D mapping of the EPCC campus. DesignBuilder software was used to create the BEM’s volume (shown in Figure 4). The HVAC system was modeled as a simplified version with ideal loads, without equipment capacity limitations. Therefore, the model prioritises achieving measured temperatures, disregarding the required energy consumption. While DesignBuilder provides a user-friendly platform, modifying certain parameters for calibration can be challenging and time-consuming. To address this, the

Table 2 Weather station measured parameters

Sensor	Unit
Dry bulb temperature	°C
Dew point	°C
Relative humidity	%
Horizontal radiation	W/m ²
Diffuse radiation	W/m ²
Wind speed	m/s
Wind direction	deg
Precipitation	mm
Atmospheric pressure	MPa

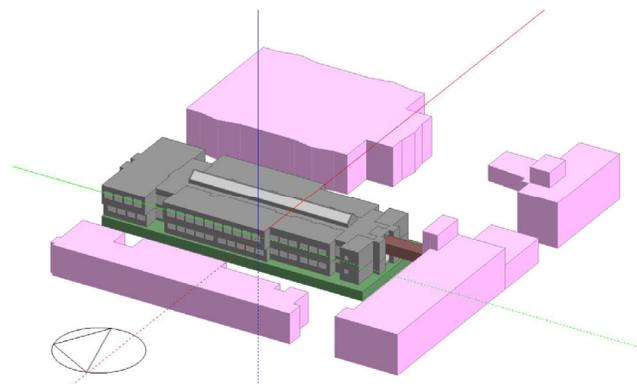


Fig. 4 Visualisation of the BEM, including surrounding buildings

BEM was exported to the EnergyPlus format. To isolate the effect of other loads on the building envelope and HVAC system calibration, the variables for People, Lights, and Other Equipment were removed in the EnergyPlus Editor.

The BEM underwent calibration through the methodology detailed in Figure 5. This involved training and adjusting the model using measured temperatures. The methodology employed was developed and validated by Fernández Bandera and Ramos Ruiz (2017) and Gutiérrez González et al. (2020). Since the ideal load settings lacked limits, the load periods were not suitable for model training, but were used for model warm-up. Consequently, the free oscillation was set as boundary condition, with measured temperature as the control parameter. By doing so, we reduced the number of unknown parameters, ensuring that the building envelope was solely influenced by external conditions (Ramos Ruiz et al. 2016). This methodology has proven effective for accurately determining parameters related to the building envelope (Fernández Bandera and Ramos Ruiz 2017). With a well-calibrated envelope, the energy consumption patterns can be reliably identified and classified, free from the distortions caused by envelope misinterpretation.

Envelope parameters were optimised through a combined process using a Genetic Algorithm NSGA-II (Ramos Ruiz et al. 2016) within JePlus+EA software (Zhang and Korolija 2010), with optimisation continuing until the simulation results closely matched the monitored data. The program was configured to simulate a population of 30 models per generation, iterating through 200 generations (6000 models total). A tournament selection operator was used, comparing three models at a time. The objective function was set to

minimise the combined values of MAE, RMSE, and the inverse of R^2 .

Despite building monitoring with EcoWin sensors since 2016, data processing and validation revealed significant quality issues. Sensor registration data instability led to extensive periods of missing values, hindering model performance and potentially compromising calibration effectiveness (Lillstrang et al. 2022). Consequently, MClimate sensor data was chosen for whole-building calibration. These sensors provided an accurate representation of building temperature, making them ideal for envelope calibration.

The calibration using MClimate sensor data involved a 19-day continuous free oscillation period from December 23, 2023, to January 10, 2024 (comprising 1506 timesteps). During this period, called Training 1, the envelope’s primary materials were calibrated, including air chamber, ceramic beam filling, and EPS in the slab; double-perforated brick, 1-foot perforated brick, ceramic brick, 1/2-foot perforated brick, and mineral fiber in the walls; and foundation material and window glazing. Additionally, individual room zone capacitance, internal mass, and effective leakage area were calibrated. The specific parameters calibrated are detailed in Table 3.

As previously mentioned, room PB:033 lacked a radiator and, thus, a MClimate sensor, rendering it incompatible with the initial calibration process focused on radiator sensors. Moreover, EcoWin sensors were unavailable during this Training 1 period. Hence, this thermal zone remained partially calibrated, with three missing parameters (zone capacitance, internal mass, and effective leakage area). Therefore, PB:033 required separate calibration for these parameters during August 2017 using data collected by EcoWin sensors. This second calibration period, named Training 2, involved a 14-day continuous free oscillation period from August 1, 2017, to August 14, 2017 (comprising 2016 timesteps).

The BEM was considered calibrated when it achieved compliance with CIBSE benchmarks (CIBSE 2020). To assess model performance, simulations were conducted during two distinct periods: Checking 1 (Spring, 08/04/2017–17/04/2017) and Checking 2 (Winter, 23/12/2017–31/12/2017). Table 4 summarises the periods used for model training and checking.

Table 5 and Figure 6 present the temperature model performance, comparing simulated and measured temperature (M_i) for each calibration stage, named as Base Model (BM, prior calibration model), Partially Calibrated Model (PCM, Training 1 model), and Fully Calibrated Model (FCM, Training 2 model). Although the Training 1 period was excluded from the analysis due to the lack of data for PB:033, the resulting model PCM was included to compare

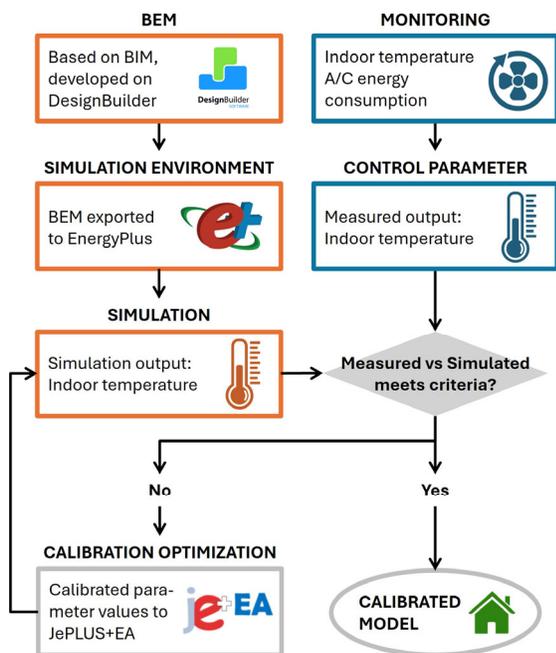


Fig. 5 Calibration environment

Table 3 Parameters modified during calibration

Parameter	Unit
Zone capacitance multiplier	
Temperature capacity multiplier	—
Internal mass	
Surface area	m ²
Zone infiltration: effective leakage area	
Effective air leakage area	cm ²
Foundation: Kiva settings	
Soil conductivity	W/(m·K)
Soil density	kg/m ³
Soil specific heat	J/(kg·K)
Ground solar absorptivity	—
Ground thermal absorptivity	—
Ground surface roughness	m
Far-field width	m
Material	
Conductivity	W/(m·K)
Density	kg/m ³
Specific heat	J/(kg·K)
Solar absorptance	—
Thermal absorptance	—
Window material: glazing	
Solar transmittance	—
Visible transmittance	—
Conductivity	W/(m·K)

Table 4 Training and checking periods for calibration

	Training 1	Training 2
Date	23/12/2023–10/01/2024	01/08/2017–14/08/2017
Timesteps	1506	2016
Sensor data	MClimate	EcoWin
	Checking 1	Checking 2
Date	08/04/2017–17/04/2017	23/12/2017–31/12/2017
Timesteps	1440	1236
Sensor data	EcoWin	EcoWin

Table 5 Models' uncertainty values

	Checking 1			Training 2			Checking 2		
	MAE	RMSE	R ²	MAE	RMSE	R ²	MAE	RMSE	R ²
CIBSE				1.00	±1.50	>75.00%			
BM	0.37	0.44	74.96%	1.99	2.17	23.87%	0.59	0.71	68.70%
PCM	0.40	0.47	70.63%	0.65	0.78	56.16%	0.98	1.08	75.47%
FCM	0.64	0.71	93.32%	0.22	0.28	92.82%	0.56	0.59	95.17%

Note: CIBSE: Chartered Institution of Building Services Engineers - International standard.

BM: Base Model, referred to the model prior to calibration process.

PCM: Partially Calibrated Model, referring to the envelope calibration during Training 1.

FCM: Fully Calibrated Model, i.e., the calibrated model for the missing parameters during Training 2.

with the FCM and to highlight the significant improvements achieved.

During Training 2, the FCM significantly outperformed BM and PCM, closely approximating measured temperatures as demonstrated in Figure 6(b). This superiority is reflected in the lower MAE and RMSE values and a substantially higher R² percentage in Table 5. While BM and PCM exhibited closer temperature predictions during Checking 1 (lower MAE and RMSE), their R² values fell below 75 %, potentially due to a peak of temperature around midday shown in Figure 6(a). In contrast, the FCM maintained a high R² percentage, aligning closely with the measured temperature curve. During Checking 2, the FCM once again demonstrated superior performance compared to BM and PCM showing a curve performance closer to Mi in Figure 6(c).

Following the calibration of PB:033, the model was configured to continuously activate the ideal loads to achieve the measured temperature. Under this configuration, the model was simulated for the full years 2017 to 2020. Model simulations for the years 2017 to 2020 generated output data for “Zone Ideal Loads Zone Total Heating Energy” and “Zone Ideal Loads Zone Total Cooling Energy”, which were subsequently used for clustering analysis.

2.3 Clustering methodology

The clustering methodology employed in this study is fundamental to understanding the obtained results. This subsection provides a detailed explanation of the underlying theory and specific steps involved in grouping the data, as well as a thorough description of the process used in this work, justifying its selection and detailing its application to the data.

2.3.1 Foundations

Clustering is a data classification technique that consists of dividing a population of data into a number of sets (called clusters), where the objects in each set are similar to each



Fig. 6 Graphical representation temperature models performance: (a) Checking 1, (b) Training 2, and (c) Checking 2

other. In this way, it is possible to divide large populations of data into more homogeneous groups, and to analyse each cluster separately to obtain more reliable statistical results for each group (Backhaus et al. 2021).

The main steps in applying a clustering technique are: (1) selecting the variables to be clustered; (2) determining the similarities between objects within a cluster, or how to calculate the distances between them to determine whether or not they belong to the same cluster; (3) choosing the most appropriate clustering method for the data; (4) selecting the number of clusters (if the method requires it); and (5) interpreting the results obtained (Backhaus et al. 2021).

Regarding the clustering method, there are several types: hierarchical, density based, grid based, model based, fuzzy (soft), and partition methods. These methods differ in aspects such as the pre-selection of the number of clusters, the calculation to determine the objects belonging to each cluster, and if an object can belong to one or more clusters. Without going into detail about each of the methods (more information can be found in Acito (2023) and Oyewole and Thopil (2023)), the requirements that the method chosen for this work should have are as follows:

- The number of clusters must be specific and pre-selected.

- Clusters shall be exclusive, i.e., each object shall belong to only one cluster.
- The method must be able to cluster a large amount of data.

Based on the aforementioned requirements, partition group methods, such as the k -means algorithm, are best suited for this application. This clustering method groups the data into k clusters according to the distance of the points from the cluster centroid and is used in various fields. For example, it can identify potential locations for ocean renewable energy development by analysing relevant data points (Uti et al. 2023) or to correlate wave heights with data used to predict annual coastal bed evolution (Papadimitriou and Tsoukala 2024). The k -means method works as follows (Backhaus et al. 2021; Acito 2023):

- Initial k centroids are determined, either randomly or from the first k data of the population.
- Each object is assigned to the cluster for which the Euclidean distance to the centroid is the smallest.
- From the k clusters formed, the centroids are recalculated from the mean of the data in each cluster.
- The objects are reassigned to the cluster whose centroid is closest.

e. Steps c and d are repeated until the maximum number of iterations previously established is reached or until the distance between the previous centroids and the recalculated ones is less than a threshold.

Finally, to select the number of clusters to be considered, the silhouette coefficient (SC) calculation is used, which provides a quantitative way to assess the quality of the clustering results (its calculation can be found in Acito (2023)). The quality of a clustering can be classified according to SC ranges:

- $SC \leq 0.5$ indicates poor results.
- $0.50 < SC < 0.7$ suggests medium quality results.
- $SC \geq 0.7$ implies good clustering and strong classification.

2.3.2 Applied methodology

The aim of clustering in this work is to determine, from the simulation outputs generated by the calibrated model, those energy levels that represent the HVAC system schedule, free oscillation, and non-recurrent events. Within the scope of this research, non-recurrent events refer to anomalous loads introduced into the energy consumption pattern, primarily caused by occupant behaviours like window opening or individual radiator use. For this purpose, clustering is a useful technique because it classifies the data according to the similarity between them, so it is possible to perform a classification of the energy consumed according to different energy levels. Based on the above, the variable used for clustering was the energy required by the model to maintain the temperature of the room, as it reflects the energy consumption that would occur at each simulated time point.

In addition, the method used for clustering was the k -means described above, as it meets the requirements established for this application. This method requires a prior selection of the number of clusters and classifies the data according to them, with each object belonging to a single cluster. Finally, the data to be clustered corresponded to sensor data generated during a calendar year in 10-minute intervals (52560 timesteps), so it is necessary to cluster large amounts of data, which can be done using the k -means method.

The decision regarding the number of clusters to be used was influenced by two factors: the SC and the energy values within the lowest energy level (cluster 0). The objective was to attempt to classify intervals of free oscillation (minimum energy consumption) within cluster 0, necessitating exceptionally low energy values in this cluster. To identify the optimal number of clusters, several clusterings were conducted, only varying the number of clusters used. From the results, the mean and maximum energy values of cluster 0 alongside the SC were evaluated for each year. The resulting SCs, shown in Figure 7(a), show a strong

classification according to the above ranges, regardless of the year and the number of clusters. The lower the number of clusters used, the higher the SC. However, an analysis of cluster 0's energy values reveals a clear correlation: as the number of clusters decreases, the average and maximum energy within cluster 0 increases (see Figures 7(b) and (c)). This indicates that the clustering classification has inadvertently included energy consumption values associated with energy sources like cooling or heating, which are not representative of free oscillation. It is evident that starting with 4 groups, the average and maximum energy levels significantly decrease and remain relatively consistent for 4, 5, and 6 clusters. Given that 4 clusters mark a turning point in terms of energy values and the SC continues to decline beyond this point, a number of 4 clusters was ultimately selected.

The methodology of this work consists of carrying out two clustering processes with the aim of detecting the clusters that most accurately reflect the free oscillation schedule of the room. For both processes, clusters are classified from low energy values (cluster 0) to high energy values (cluster 3). The clustering methodology is summarised in Figure 8 and detailed below:

- a. A first clustering was performed to identify the energy levels that most closely match the real schedule of the HVAC system, thus providing a first approximation of the free oscillation schedule of the room. In this first clustering, the cluster that best identifies the free oscillation was considered to be cluster 0, i.e., the one with the lowest energy. The simulated free oscillation schedule obtained from this clustering will be called FO-1C.
- b. A second clustering (subclustering, from now on) was performed on cluster 0 of the previous step (lowest energy) in order to detect more faithfully the energy levels corresponding to the real free oscillation and to distinguish them from those corresponding to non-recurrent events. In this way, the levels corresponding to the energy input from individual radiators, people, open windows, etc., are eliminated, and the approximation of the free oscillation schedule of the room is more accurate. In this second clustering, the subclusters that best reflect the free oscillation were considered to be subclusters 0, 1 and 2, leaving subcluster 3 out of this analysis. The simulated free oscillation schedule obtained from this clustering will be named as FO-2C.

Figure 8(a) outlines the clustering process and how the clusters were grouped to generate the free oscillation schedules. These schedules are composed of values 0 and 1, where 0 is free oscillation and 1 is the presence of energy consumption. Figure 8(b) shows a flowchart of the process of assigning the free oscillation schedule according to the clusters/subclusters and the clustering process. In both the

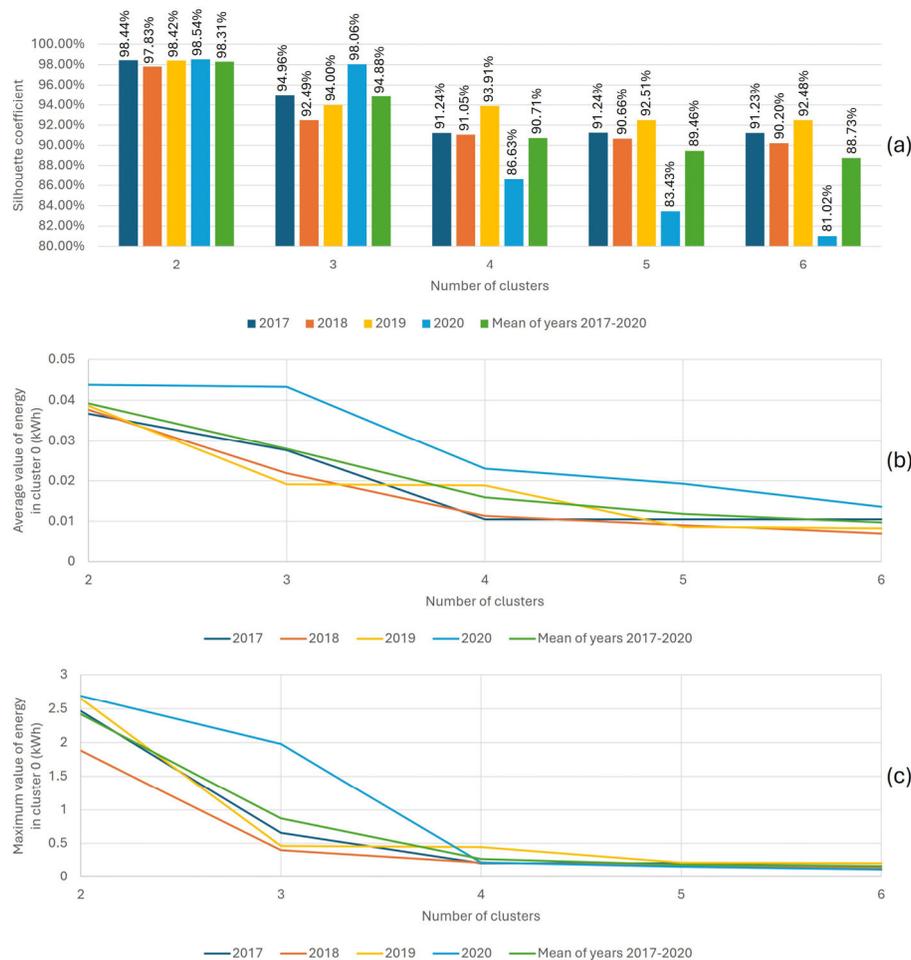


Fig. 7 Results used for selecting the number of clusters: (a) silhouette coefficient, (b) mean energy values in the lowest energy cluster, (c) maximum energy values in the lowest energy cluster

first and second clusterings, all data timesteps were evaluated one by one, analysing to which cluster/subcluster the timestep belonged to. In the case of the first clustering (on the left side of Figure 8(b)), the membership in cluster 0 was evaluated and, if positive, a 0 was assigned in the free oscillation schedule (since cluster 0 was considered as FO-1C). In the case of the second clustering (on the right side of Figure 8(b)), it was analysed whether the corresponding timestep belonged to one of the subclusters 0, 1 or 2 and, if so, a 0 was assigned in the free oscillation schedule (since subclusters 0, 1 and 2 were considered to be FO-2C).

Regarding the real HVAC operation schedule, which was necessary to test the effectiveness of clustering, it was obtained from the data measured by the Circutor Wibeec sensor during the years 2017–2020 (January to May 2017 is not included in this work as the sensor was not yet installed.). The machine operating hours were recorded both in summer and in winter, as room PB:033 is not heated by radiators (as the rest of the building is). In order to assign the free oscillation schedule, it was determined that those instants when the machine was not switched

on corresponded to free oscillation. This free oscillation schedule from the HVAC system will be referred to as FO-HVAC, hereafter.

Finally, for clustering, the Python library sklearn (now scikit-learn), version 1.4.2, has been used in this work, together with Python version 3.11.9.

3 Results and discussion

This section is dedicated to the results of the simulations and clustering. First, the real HVAC operation schedule of PB:033 of Computer Science Pavilion is compared with the results of the two clusterings, and then the model calibration results are analysed, both steps by means of the free oscillation schedules.

FO-1C and FO-2C are compared with FO-HVAC in Table 6. It can be seen that the schedules of both clusterings coincide with those of the real HVAC system in very high percentages, exceeding 84% in all cases except July 2020, where a result of 75% match was obtained. The year in which the coincidence is best is 2019, although in all years

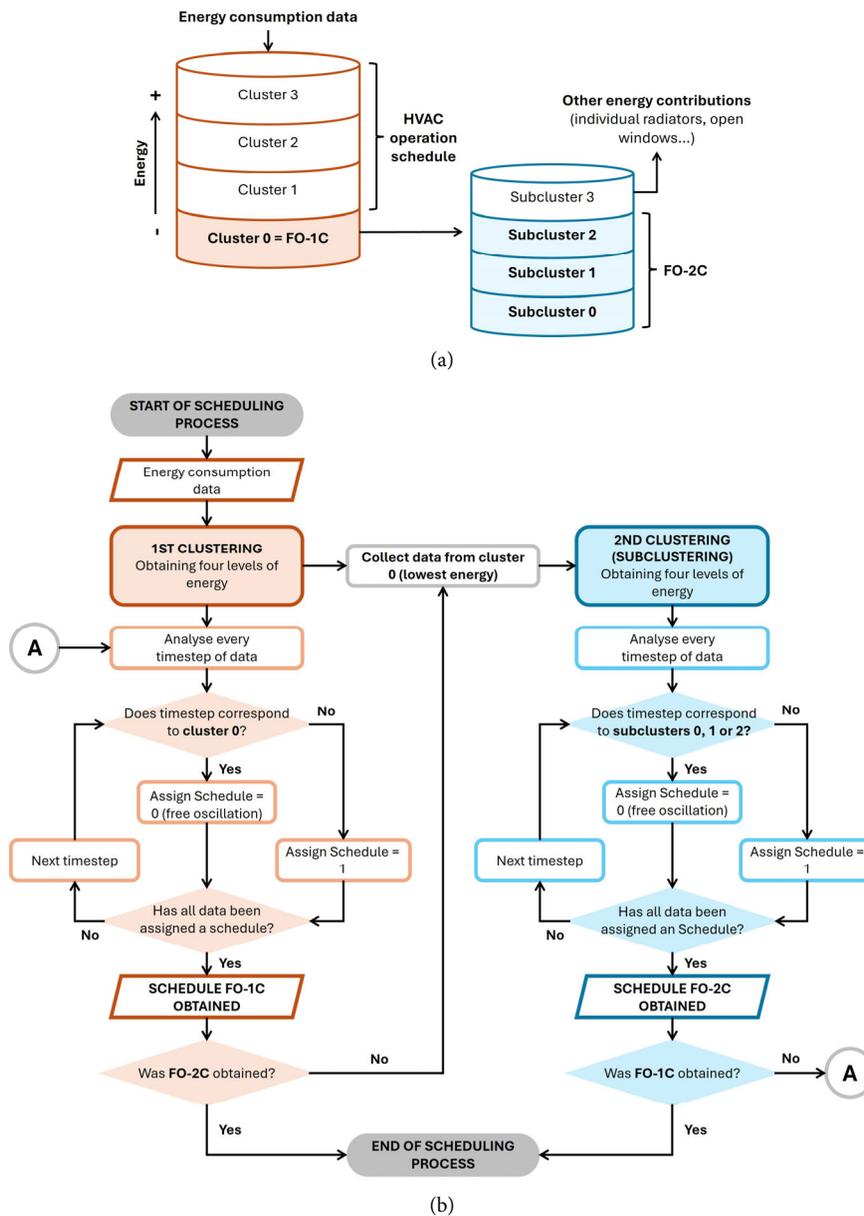


Fig. 8 Methodology for obtaining the free oscillation schedule from clustering: (a) clustering methodology, (b) free oscillation scheduling methodology

the overall match rate always exceeds 90%, showing that the clustering has been effective.

Comparing the percentages by month, it can be seen that, in 2017, the highest number of matches were obtained between July and October; in 2018, from August to October; in 2019, in April, May, June and October; and, finally, in 2020 in April and from August to October. It is observed that the month of October is a good matching month in all years, and furthermore, high percentages are observed from August to October in 2017, 2018 and 2020. In general, the best coincidences are achieved in the spring, summer and early autumn, when the weather in Cáceres is warmer and it is necessary to use the air conditioner. This may also

be due to the fact that, at this time of year, the only source of cooling in the room is the air conditioner, so this machine provides all the energy supplied to the room. However, in winter, the temperature increases caused by both occupancy and individual radiators used for heating, which are not controlled by the Circutor Wibeee sensor, consequently not being taken into account into the hours of operation of the HVAC system. In addition, the room has a large amount of computer equipment, as it is a laboratory, which raises the temperature inside the room.

By comparing the results of the first and second clustering, it can be extracted that the coincidence is higher between FO-HVAC and FO-1C in almost all cases, which implies

Table 6 Coincidence between free oscillation schedule obtained from the air conditioner operation schedule and clustering results (%). Column named “1st” refers to free oscillation schedule obtained from first clustering (FO-1C), while column “2nd” corresponds to that schedule obtained from second clustering (FO-2C)

Year	2017		2018		2019		2020		
	Clustering	1st	2nd	1st	2nd	1st	2nd	1st	2nd
June–December		95.9	94.69	96.49	94.84	96.59	96.35	96.02	90.18
Whole year		—	—	94.58	92.50	96.56	95.22	95.00	90.70
January		—	—	91.64	88.01	91.46	90.95	88.91	87.99
February		—	—	89.16	85.74	93.60	89.24	88.57	86.11
March		—	—	91.17	89.11	98.70	92.59	94.09	92.47
April		—	—	91.60	90.00	99.26	96.67	98.98	97.87
May		—	—	95.54	92.81	99.35	98.36	96.93	92.45
June		92.92	90.76	95.95	94.88	99.81	99.35	95.37	91.46
July		98.10	97.49	95.88	94.62	96.62*	97.56*	89.87	75.49
August		99.26	99.06	99.19	98.70	98.03*	98.14*	98.03	91.96
September		98.50	97.80	97.85	96.81	97.80*	97.82*	98.31	96.85
October		98.61	97.78	98.30	96.44	99.17	98.34	98.48	97.54
November		91.00	88.98	93.01	88.38	92.80	91.23	97.36	94.14
December		92.79	90.73	95.18	93.88	91.94*	92.00*	94.98	84.32

* In these cases, coincidence of real schedule is higher with 2nd clustering results rather than 1st clustering.

that cluster 0 of the first clustering more closely reflects the shutdown schedule of the HVAC system. However, as it is not yet 100% consistent with the actual machine schedule and since the percentage of agreement with FO-2C is even lower, it is a reasonable assumption that energy levels not corresponding to HVAC were included in other clusters. The results indicate that both cluster 1 and subcluster 3, which correspond to intermediate levels in terms of energy as previously shown in Figure 8(a), include energy consumption that does not correspond to the machine.

Although in most cases the overlap with FO-1C is greater than with FO-2C, there are some exceptions that have been marked in Table 6. These coincidences occur in July–September and December, both in 2019. In these cases, it is likely that the opposite is true: subcluster 3 still includes some energy consumption levels of the real machine and, therefore, the coincidence percentages are slightly higher (although the difference is minimal).

On the basis of the above results, it can be seen that the consumption levels of the machine are included in clusters 1 to 3, so that the machine schedule is clearly reflected by these clusters. As for the free oscillation, it is clear that it should be found in cluster 0 (cluster opposite to the machine schedule), although according to the above analysis, it is possible that other energy levels have also been included in this cluster. For this reason, the second clustering was carried out within cluster 0 in order to obtain a more detailed classification of the low energy consumption levels and thus to determine the free oscillation schedule more precisely.

To validate the methodology and determine which of the two clusters was more effective in determining the free oscillation, full-year simulations for 2017 to 2020 were conducted to assess deviations between simulated and measured temperatures using the calibrated model and derived operation schedules from the two clusterings and the measured data. The results are tabulated in Table 7 and show a high percentage of free oscillation periods that align with the typical operating patterns of the HVAC system. Although FO-1C demonstrates a higher success rate in matching the actual equipment schedule, FO-2C identifies fewer free oscillation timesteps. Cluster FO-2C effectively recognises occupant behaviour independent of the HVAC system, such as opening windows or individual radiators. These occupant-driven loads introduce variations in temperature patterns, leading to higher uncertainty values in FO-1C and FO-HVAC simulations.

The cluster analysis reveals key differences between FO-1C and FO-2C. In FO-1C, cluster 0 appears to contain anomalous load data, while FO-2C, on the other hand, groups these potential anomalies within subcluster 3. This distinction allows FO-2C to isolate pure free oscillation periods in subclusters 0, 1, and 2, unlike FO-1C which may have these free oscillation periods mixed with anomalous loads. Overall, the ability of FO-2C to identify non-HVAC loads, including potential anomalies, leads to a more accurate representation of free oscillation periods.

The year 2020 exemplifies this challenge. Due to the COVID-19 pandemic, occupancy rates and occupant behaviour significantly changed. This variation, however, was not

Table 7 Calibrated model's uncertainty values for full year simulation according free oscillation schedule obtained from the air conditioner operation schedule and clustering results

Operation schedule	Year	Free oscillation timesteps	Free oscillation timesteps	MAE [°C]	RMSE [°C]	R ²
FO-HVAC	2017*	26205	85.03%	0.79	1.17	96.82%
	2018	51112	97.25%	1.06	1.43	94.35%
	2019	50189	95.49%	0.91	1.26	93.56%
	2020	51800	98.55%	1.74	2.10	92.81%
FO-1C	2017*	29470	95.63%	0.40	0.46	99.47%
	2018	49327	93.85%	0.69	1.10	95.83%
	2019	51211	97.43%	0.72	1.33	92.56%
	2020	49754	94.66%	1.18	1.58	94.77%
FO-2C	2017*	29019	94.17%	0.36	0.41	99.58%
	2018	48154	91.62%	0.36	0.40	99.53%
	2019	49891	94.92%	0.38	0.45	99.01%
	2020	47398	90.18%	0.46	0.50	99.71%

* The 2017 simulation covered only the period from June to December, not the entire year.

accounted for by the FO-1C schedule, leading to failures in meeting simulation uncertainty thresholds (MAE and RMSE). This highlights the need for building performance models to consider occupant behaviour beyond just the operation of the HVAC system. The success of FO-2C, which explicitly identifies non-HVAC loads such as window opening, demonstrates the importance of incorporating a broader perspective in these models.

To assess cluster performance, the full-year simulations, previously elaborated for the years 2017 to 2020, were divided into summer, autumn, winter, and spring periods, and the temperature curve for one week of each season was graphically represented. The representative week for each season was selected based on load periods and the ability to exemplify specific situations. These are, in detail, the following:

- Figure 9: February 4–10. Year 2017 was excluded due to data availability starting on May 31st.
- Figure 10: April 3–9. Year 2017 was excluded due to data availability starting on May 31st.
- Figure 11: August 19–25.
- Figure 12: December 11–17. Year 2020 was excluded due to HVAC system downtime from late September to 2023.

Mi was included as a reference curve, meaning that deviations from the Mi curve indicate that the clusters struggled to capture the free oscillation periods. To identify non-occupancy days, typically weekends, the corresponding days have been highlighted in red. Moreover, to justify differences between simulated temperature curves, the graphs present the discrepancies between schedules. The shaded grey area represents free oscillation of FO-HVAC, while the remaining unshaded area (in white) corresponds to the HVAC system's operation schedule. The cyan and green

bars indicate anomalous loads identified by FO-1C and FO-2C during the FO-HVAC's free oscillation period.

Starting with the **winter period** plotted in Figure 9, the following can be extracted:

- Results from **2018 and 2019** (see Figures 9(a) and (b)) show that the temperatures coincide with the **HVAC operating schedule** shaded in white, noting temperature increases in these periods that coincide with the HVAC operation in heating mode. An exception is seen with FO-1C on days 5–6 February, where the model has not been able to reach the Mi curve. This is probably due to a lack of load detection by the cluster, thus preventing the model from raising the room temperature and increasing the actual temperature levels. Continuing with the FO-HVAC results, it exhibited minimal deviations in 2018 and 2019, although anomalous loads identified by both FO-1C and FO-2C on February 7, 2018, and February 7, 2019, introduced notable discrepancies between FO-HVAC and Mi curve. Concerning FO-1C, it struggled to identify HVAC operation during specific dates in 2018 and 2019, as said before, affecting its ability to follow the actual temperature curve. On the contrary, FO-2C consistently demonstrated strong performance, accurately identifying anomalous loads that contributed to energy consumption and to achieve temperature values similar to Mi curve.
- Regarding **2020** (depicted in Figure 9(c)), the results reveal a stark contrast to previous years. The HVAC system remained inactive throughout the week illustrated, with the exception of February 4. However, the Mi curve indicates that temperature increases likely occurred due to other heating sources, such as individual radiators. As evidenced, the consumption of these energy sources is

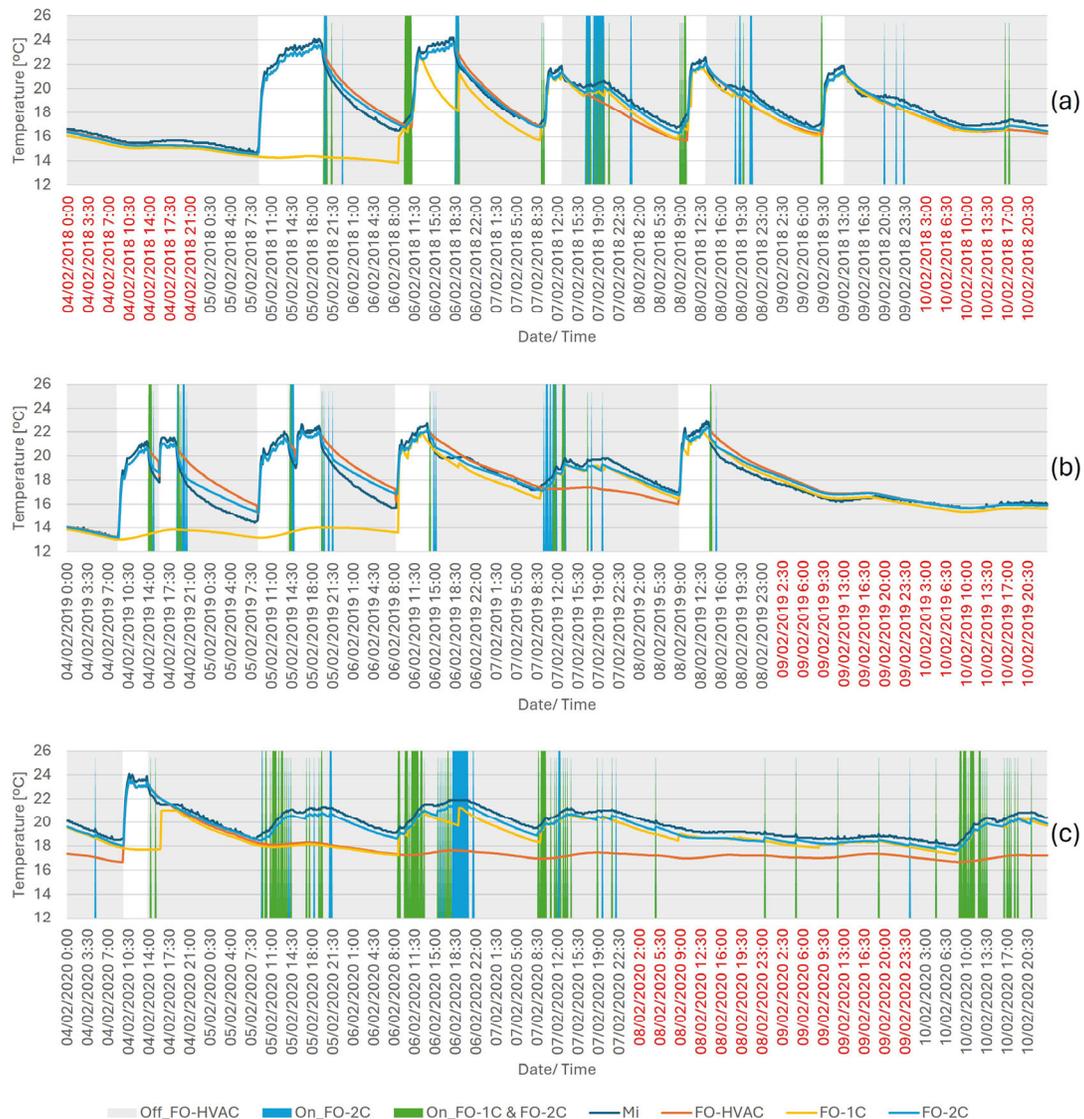


Fig. 9 Winter period: comparison between measured temperatures (Mi) and simulated temperatures from the calibrated model under free oscillation schedules (FO-HVAC, FO-1C, and FO-2C) during one week of the winter period (February 4–10): (a) 2018, (b) 2019, and (c) 2020. The shaded areas represent free oscillation periods: grey for FO-HVAC, cyan and green for anomalous loads detected by FO-1C and FO-2C during FO-HVAC free oscillation. Non-occupancy periods have been highlighted in red

not accounted for by the FO-HVAC schedule, rendering it incapable of accurately predicting actual temperatures.

- With respect to **anomalous loads**, most of those identified across all years coincide with the initiation or termination of machine startup sequences. These periods might overlap with times when the machine is operating at low power or when alternative energy sources are temporarily utilised. Notable examples include the afternoon of February 7, 2018, and the morning of February 7, 2019, which coincide with increased room temperatures. These anomalous loads confirm the use of additional heat sources, as they cannot be attributed to factors like opening windows, especially given the time of year analysed, which would typically lead to a temperature decrease.
- Finally, during the **non-occupancy periods** indicated in red colour, the temperatures reflect free oscillation patterns, as expected during weekends when the building is unoccupied and the HVAC system is inactive. Nevertheless, both FO-1C and FO-2C identified limited anomalous loads in 2018 and 2020, potentially due to clustering errors, as the temperature data may not accurately reflect the behaviour of additional energy sources.

Regarding **spring period** plotted in Figure 10, a discussion similar to the previous period can be drawn:

- Regarding the **HVAC operating hours**, only one use of the HVAC was observed during **2018** (Figure 10(a)), and more specifically on 3 and 9 April. The temperature curves of all models, except FO-1C, reflect the temperature



Fig. 10 Spring period: comparison between measured temperatures (Mi) and simulated temperatures from the calibrated model under free oscillation schedules (FO-HVAC, FO-1C, and FO-2C) during one week of the spring period (April 3–9): (a) 2018, (b) 2019, and (c) 2020. The shaded areas represent free oscillation periods: grey for FO-HVAC, cyan and green for anomalous loads detected by FO-1C and FO-2C during FO-HVAC free oscillation. Non-occupancy periods have been highlighted in red

rise caused by the use of this machine. FO-1C exhibited poor performance in April 9, 2018 and from April 8 to 9, 2019, failing to identify certain anomalous loads and, thus, to achieve Mi curve. In addition, the performance of FO-HVAC deteriorated compared to FO-1C and FO-2C, as well as in 2019 (depicted in Figure 10(b)), probably due to prolonged periods where the HVAC has not been used. This implies that periods of possible loads and temperature rises are not reflected in the FO-HVAC schedule, and that the simulated temperature curve is not able to reach the Mi curve. According to the results, the best performing model was FO-2C in first place, followed by FO-1C and FO-HVAC.

- Regarding 2020 (plotted in Figure 10(c)), this year might

be analysed separately due to its unique circumstances, as the building's low occupancy caused by the COVID-19 pandemic makes it a special case. This situation is reflected in the results, with all simulated temperatures very close to the Mi curve. It is clear that the building is in free oscillation as this period coincides with the confinement during the pandemic, that is why the whole period in this graph is marked as unoccupied. The observed discrepancies between FO-HVAC and FO-2C, as indicated by the lower coincidence rate in Table 6, suggest that the methodology might be susceptible to errors when dealing with limited operational data.

- With regard to **anomalous loads**, both FO-1C and FO-2C identified certain instances in 2018, while FO-2C

detected some in 2019. These anomalous loads coincided with minor temperature increases, similar to those observed during the winter period. These increases could be attributed to factors like open windows, as outside temperatures begin to rise, or the presence of people and IT equipment within the room, leading to slight indoor temperature elevations. Given the current season, the use of individual radiators is unlikely, so they were not considered in this analysis. Anomalous loads identified by both clusters, primarily between April 4 and 6, 2018, likely contributed to the deviations observed in the FO-HVAC schedule.

- Finally, the **non-occupancy periods**, marked in red, correspond to weekends in 2018 and 2019, and to the COVID-19 pandemic confinement in 2020. Examining the anomalous loads detected during these periods in 2018 and 2019 (given that 2020 has already been discussed), it is evident that fewer anomalies were identified compared to the winter period. Specifically, only two anomalous loads were detected by FO-2C, and one by both FO-1C and FO-2C. This suggests that the clustering algorithm performed more effectively in these non-occupancy periods.

Based on the results for the **summer period**, as illustrated in Figure 11, the following observations can be made:

- During this period, all years exhibited similar characteristics. In 2017 and 2019, the HVAC was activated for several days, albeit briefly. The activation is clearly discernible as a sharp temperature drop due to air conditioning usage. Regarding **temperature results**, all simulated temperature curves closely resembled the Mi curve, except for 2020, where the curves diverged more significantly. The FO-HVAC model exhibited the poorest performance in this year, experiencing a temperature drop of nearly 1.5 °C below Mi, similar to the spring confinement period. This discrepancy could be attributed to the prolonged lack of occupancy during both the confinement period and the summer holidays.
- Regarding **anomalous loads**, several instances were detected during the summer months. For example, on August 21 and 25, 2017; August 21 and 24, 2018; and August 23, 2019, anomalous loads coincided with sudden temperature drops. These anomalies could be attributed to errors in the HVAC consumption sensor, which might not have accurately detected its activation, or the opening of windows during cool morning hours. However, given the substantial temperature drops observed, it is more likely that the sensor errors were the primary cause. Additionally, both FO-1C and FO-2C schedules identified anomalous loads in 2017, 2018, and 2020, potentially due to clustering errors. Notably, the loads detected

by FO-2C in 2020 helped to correct deviations in its temperature curve.

- Finally, concerning **non-occupancy periods**, which in this case correspond to weekends, some anomalous loads were detected in 2017 and 2020, also likely due to clustering errors.

During the **autumn period** depicted in Figure 12, the following can be outlined:

- The HVAC system was used in heating mode in all years, with brief periods of operation in 2018 and 2019. However, in 2020, the system remained active for extended periods, even at night, suggesting a lack of user attention. The machine's usage is clearly reflected in increased temperatures, as observed during the winter period (Figure 9).
- Overall, FO-1C exhibited the poorest performance, failing to identify extended load periods in 2019. For specific periods, such as December 11th to 13th, 2019, the cluster incorrectly classified the period as non-load, likely due to the prolonged operation of the machine in a low-power state solely for temperature maintenance. This aligns with the observed temperature drop after the initial load detection, indicating a lack of subsequent load identification. However, for the period of December 16 to 17, 2019, a clear error occurred within the cluster, preventing it from identifying any portion of the load period. In contrast, FO-2C demonstrated exceptional performance, followed by FO-HVAC.
- The **anomalous loads** observed during autumn could be attributed to the use of individual radiators, similar to the winter period. These loads often coincide with slight temperature increases. Additionally, as in winter, some anomalous loads align with the start or end of HVAC activation periods, which may also overlap with the use of individual radiators. A particularly high concentration of anomalous loads occurred on December 15, 2017, and December 12, 2018, detected by both clusters on the former date and solely by FO-2C on the latter. Since these events coincide with minor temperature increases, they are likely indicative of individual radiator usage.
- Finally, during **non-occupancy periods** in 2018 and 2019, free oscillation patterns were similar across all schedules. However, in 2017, FO-1C and FO-2C identified several anomalies, which may be attributed to methodological errors.

The study's findings demonstrate that the models effectively predicted temperature changes aligned with HVAC operation in most scenarios. However, the models encountered challenges in accurately identifying anomalous loads, particularly during periods of low occupancy or when alternative heating sources were utilised. Clustering



Fig. 11 Summer period: comparison between measured temperatures (Mi) and simulated temperatures from the calibrated model under free oscillation schedules (FO-HVAC, FO-1C, and FO-2C) during one week of the summer period (August 19–25): (a) 2017, (b) 2018, (c) 2019, and (d) 2020. The shaded areas represent free oscillation periods: grey for FO-HVAC, cyan and green for anomalous loads detected by FO-1C and FO-2C during FO-HVAC free oscillation. Non-occupancy periods have been highlighted in red

errors also contributed to inaccurate predictions in certain instances. Furthermore, the performance of the models varied across different seasons and specific time periods.

In summary, the following assertions can be drawn from the analysis:

- 1) Regarding clustering, clusters 1 to 3 better represent the

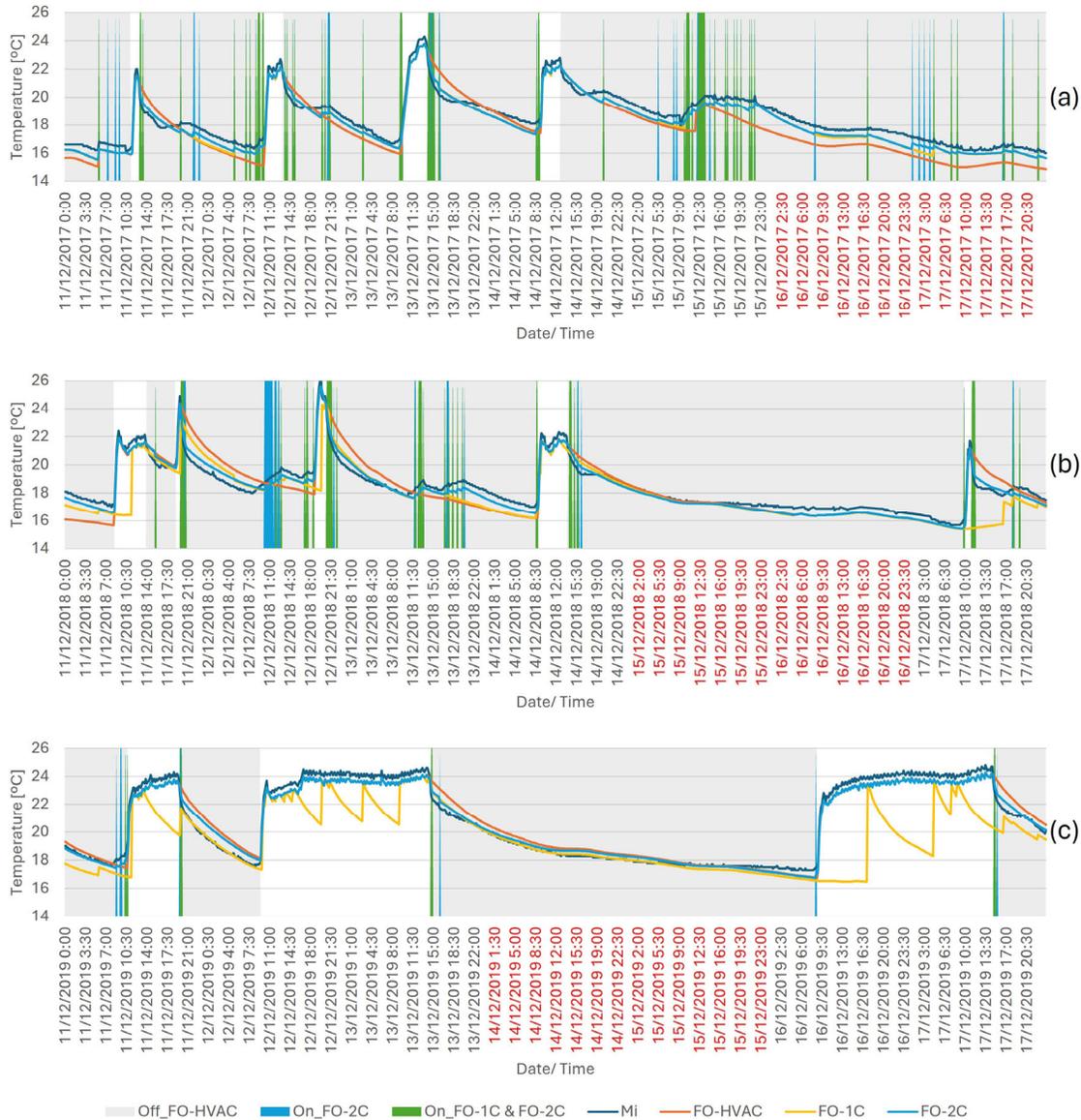


Fig. 12 Autumn period: comparison between measured temperatures (Mi) and simulated temperatures from the calibrated model under free oscillation schedules (FO-HVAC, FO-1C, and FO-2C) during one week of the autumn period (December 11–17): (a) 2017, (b) 2018, and (c) 2019. The shaded areas represent free oscillation periods: grey for FO-HVAC, cyan and green for anomalous loads detected by FO-1C and FO-2C during FO-HVAC free oscillation. Non-occupancy periods have been highlighted in red

HVAC operation schedule, while non-recurrent events appear to be included in subcluster 3 and the free oscillation periods are likely to be included in subclusters 0 to 2.

- 2) FO-1C struggled to identify extended periods of continuous HVAC operation, particularly during the autumn period (Figure 12).
- 3) During non-occupancy periods, FO-1C and FO-2C identified only a few anomalous loads, which could be attributed to methodological errors. In 2020, the limited HVAC operation data likely contributed to a higher incidence of these errors, resulting in suboptimal cluster performance.

- 4) Notably, FO-2C appears to capture the pure free oscillation behaviour more accurately compared to FO-1C regardless of the season, which is consistent with the first statement and with the inclusion of free oscillation in subclusters 0 to 2.

4 Conclusion

This paper has proposed a methodology combining white-box modeling and clustering techniques to identify free oscillation periods within a building based on energy data. A Building Energy Model (BEM) was initially developed and calibrated using DesignBuilder, EnergyPlus and JePlus+EA

softwares together with historical sensor data from the target building. This calibrated model was subsequently employed to calculate the simulated energy consumption of a specific room, incorporating actual temperature measurements. Clustering techniques were then applied to group the results into distinct energy levels, resulting in a cluster that encompassed the room's free oscillation periods, as determined by comparison with the actual heating, ventilation and air conditioning (HVAC) system operating hours.

To validate the clustering outcomes, annual simulations were conducted using EnergyPlus, and the simulated temperature results were compared with actual temperature measurements. The results showed that the method was effective and able to distinguish free oscillation periods from load periods, whether these were typical loads from HVAC systems or due to non-recurrent events such as windows being opened or individual radiators being used. Furthermore, it was shown that calibrated models are able to detect periods of load not detected by the sensors, even allowing to improve the performance of the installed HVAC sensors or to bridge periods when their performance has not been correct.

As part of the research findings, it was shown that a short period of free oscillation in monitored indoor temperature is sufficient for calibration. This allows the model to accurately reflect the building's energy dynamics, laying the foundation for clustering energy consumption patterns and enabling the creation of a virtual sensor. While annual simulations were used in this study, the methodology can be easily adapted for shorter evaluation periods, such as monthly or quarterly, to align with typical energy billing cycles.

This research demonstrates that integrating calibrated white-box models with data clustering techniques presents a transformative opportunity for addressing the challenges of energy management in buildings. Key advantages include:

- **Outperforming Real Sensors:** calibrated models can outperform real HVAC sensors in generating operation schedules, especially valuable when dealing with missing data.
- **Reduced Sensor Deployment:** identifying pure free oscillation periods using only indoor temperature data minimises sensor deployment, leading to cost savings and less intrusive monitoring.
- **Comprehensive Data Generation:** the calibrated model offers a comprehensive environment for generating building data, providing insights beyond individual sensors.

By leveraging data mining and clustering techniques, this approach enhances energy modeling accuracy and enables the identification of previously unnoticed patterns

and anomalies. This, in turn, facilitates real-time HVAC system monitoring, fault detection and diagnostics (FDD), and improved building performance, reduced energy waste, and enhanced occupant comfort. Furthermore, these findings contribute significantly to our understanding of HVAC operation and occupant behaviour, laying the groundwork for FDD and the development of strategies to optimise system performance and enhance comfort.

Furthermore, by identifying HVAC system operation schedules, this methodology functions as a virtual cost allocator, determining individual thermal zone energy consumption patterns. This approach is particularly valuable in buildings where users have direct control over HVAC configuration (including operating times and temperature setpoints), making extensive sensor deployment for detecting HVAC operation impractical.

The integration of IoT-enabled temperature sensors with calibrated white-box models offers a flexible and scalable solution for building energy management, adaptable to various building types and operational conditions. This approach minimises sensor deployment while leveraging advanced virtual sensor techniques, making it technically and economically viable for large-scale implementation. As sustainable energy practices become increasingly essential, virtual sensors will play a pivotal role in creating smarter, more efficient buildings. The growing demand for energy-efficient solutions, driven by climate change and regulations like the European Union's Directive 2023/2002/EU, necessitates innovative approaches like virtual sensors as a practical, scalable, and cost-effective alternative to traditional energy metering systems.

While effectively differentiating between HVAC system schedules, free oscillation periods and non-recurrent events, this methodology currently lacks energy consumption quantification. Future research will focus on quantifying energy consumption to develop individual energy meters. This virtual tool empowers end-users to make informed decisions about energy consumption without incurring high sensor deployment costs. This knowledge enables the development of targeted and effective energy-efficient management strategies.

Data availability

The data is available upon request.

Acknowledgements

The researcher Karla Guerrero Ramírez has been funded by the Catedra Sanitas de Salud y Medio Ambiente of Universidad de Navarra. The researcher Cristina Nuevo-Gallardo has been funded by the Agencia Estatal de

Investigación under the project “Gemelo Digital de Nueva Generación de Edificios Inteligentes” (DigiTwin) (ref. CPP2021-008909). We would like to thank School of Technology of Cáceres (Spain), for providing us with both the building documentation and the sensor data to perform the necessary tests for this paper. Additionally, the authors wish to acknowledge that the weather data was elaborated by the Agencia Estatal de Meteorología, Ministerio de Agricultura, Alimentación y Medio Ambiente.

Funding note: Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

Ethical approval

This study does not contain any studies with human or animal subjects performed by any of the authors.

Author contribution statement

Conceptualisation: K. Guerrero Ramírez, C. Nuevo-Gallardo and C. Fernández Bandera. Methodology: K. Guerrero Ramírez, C. Nuevo-Gallardo and C. Fernández Bandera. Software: K. Guerrero Ramírez and C. Nuevo-Gallardo. Validation: K. Guerrero Ramírez and C. Nuevo-Gallardo. Investigation: K. Guerrero Ramírez and C. Nuevo-Gallardo. Resources: C. Fernández Bandera. Building and sensor data provider: B. Montalbán Pozas. Writing—original draft preparation: K. Guerrero Ramírez and C. Nuevo-Gallardo. Writing—review and editing: K. Guerrero Ramírez, C. Nuevo-Gallardo and C. Fernández Bandera. Supervision: C. Fernández Bandera and J.M. Santamaría Ulecia. Project Administration: C. Fernández Bandera. Funding acquisition: C. Fernández Bandera.

Open Access: This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory

regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

References

- Acito F (2023). Predictive Analytics with KNIME. Analytics for Citizen Data Scientists. Cham, Switzerland: Springer.
- Arendt K, Jradi M, Shaker H, et al. (2018). Comparative analysis of white-, gray- and black-box models for thermal simulation of indoor environment: Teaching building case study. In: Proceedings of Building Performance Analysis Conference and SimBuild, ASHRAE and IBPSA-USA.
- ASHRAE (2014). ASHRAE Guideline 14-2014: Measurement of Energy and Demand Savings. Atlanta, GA, USA: American Society of Heating, Refrigerating and Air-Conditioning Engineers.
- Backhaus K, Erichson B, Gensler S, et al. (2021). Cluster analysis. In: Roscher B (Ed), Multivariate Analysis. Gabler, Germany: Springer.
- BOE (1979). Real Decreto 2429/1979, de 6 de julio, por el que se aprueba la norma básica de edificación NBE-CT-79, sobre condiciones térmicas en los edificios. BOE (Diario oficial Boletín Oficial del Estado). Available at <https://www.boe.es/eli/es/rd/1979/07/06/2429> (in Spanish)
- Choksi KA, Jain S, Pindoriya NM (2020). Feature based clustering technique for investigation of domestic load profiles and probabilistic variation assessment: smart meter dataset. *Sustainable Energy, Grids and Networks*, 22: 100346.
- CIBSE (2020). TM63 Operational performance: Building performance modelling. Technical Report. Chartered Institution of Building Services Engineers (CIBSE).
- Coakley D, Raftery P, Keane M (2014). A review of methods to match building energy simulation models to measured data. *Renewable and Sustainable Energy Reviews*, 37: 123–141.
- de Wilde P (2014). The gap between predicted and measured energy performance of buildings: A framework for investigation. *Automation in Construction*, 41: 40–49.
- DOE (2015). M&V Guidelines: Measurement and verification for performance-based contracts (Version 4.0). Technical report. US Department of Energy.
- EU (2012). Directive 2012/27 /EU of the European Parliament and of the Council of 25 October 2012 on energy efficiency. Official Journal of the European Union, L 315, 14 November 2012, pp. 1–56. Available at <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32012L0027>
- EU (2023). Directive (EU) 2023/1791 of the European Parliament and of the Council of 13 September 2023 on energy efficiency and amending Regulation (EU) 2023/955 (recast). Official Journal of the European Union, L 231/1, 20 September 2023, pp. 1–60. Available at <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32023L1791>
- Fernández Bandera C, Ramos Ruiz G (2017). Towards a new generation of building envelope calibration. *Energies*, 10: 2102.
- Google Earth (2024). Google Earth Imagery of 39°28’44” N, 6°20’35” W. Available at <https://earth.google.com/>. Accessed 26 Jul 2024.

- Gutiérrez González V, Ramos Ruiz G, Fernández Bandera C (2020). Empirical and comparative validation for a building energy model calibration methodology. *Sensors*, 20: 5003.
- Han J, Kamber M, Pei J (2011). *Data Mining: Concepts and Techniques*, 3rd edn. Waltham, MA, USA: Morgan Kaufmann
- Hensen JLM, Lamberts R (2019). *Building Performance Simulation for Design and Operation*, 2nd edn. London: Routledge.
- Herbinger F, Vanden Hof C, Kummert M (2023). Building energy model calibration using a surrogate neural network. *Energy and Buildings*, 289: 113057.
- Ikotun AM, Ezugwu AE, Abualigah L, et al. (2023). K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences*, 622: 178–210.
- Jeong D, Park C, Ko YM (2021). Missing data imputation using mixture factor analysis for building electric load data. *Applied Energy*, 304: 117655.
- Jeong C, Byon E (2024). Calibration of building energy computer models via bias-corrected iteratively reweighted least squares method. *Applied Energy*, 360: 122753.
- Kim J, Yoon S (2023). Virtual PMV sensor towards smart thermostats: comparison of modeling approaches using intrusive data. *Energy and Buildings*, 301: 113695.
- Lee Y, Kim W (2024). Fault detection and diagnosis for variable refrigerant flow systems by using virtual sensors and deep learning. *Energy Reports*, 11: 471–482.
- Lillstrang M, Harju M, del Campo G, et al. (2022). Implications of properties and quality of indoor sensor data for building machine learning applications: two case studies in smart campuses. *Building and Environment*, 207: 108529.
- Martin D, Kühl N, Satzger G (2021). Virtual sensors. *Business & Information Systems Engineering*, 63: 315–323.
- Menezes AC, Cripps A, Bouchlaghem D, et al. (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.
- Miller C, Nagy Z, Schlueter A (2015). Automated daily pattern filtering of measured building performance data. *Automation in Construction*, 49: 1–17.
- Ministerio de Economía y Competitividad (2013). Smartpolitech: Sistema Inteligente de Eficiencia Energética para la Escuela Politécnica. Convocatoria de Infraestructura científico-tecnológica Ref: UNEX13-1E-1716 (2013–2016). <https://www.unex.es> (in Spanish)
- Montalbán Pozas B, Muriel Holgado B, Lucas Bonilla M, et al. (2022). Iterative optimization of a social inmotics-based method in order to make buildings smart and resilient. *Sustainable Cities and Society*, 82: 103876.
- Montalbán Pozas B, Lucas Bonilla M, Serrano Candela F, et al. (2023). A methodology for designing an automated system to improve the thermal performance of a large building in operation. *Buildings*, 13: 1938.
- Nepal B, Yamaha M, Yokoe A, et al. (2020). Electricity load forecasting using clustering and ARIMA model for energy management in buildings. *Japan Architectural Review*, 3: 62–76.
- Nikolaou TG, Kolokotsa DS, Stavrakakis GS, et al. (2012). On the application of clustering techniques for office buildings' energy and thermal comfort classification. *IEEE Transactions on Smart Grid*, 3: 2196–2210.
- Oyewole GJ, Thopil GA (2023). Data clustering: Application and trends. *Artificial Intelligence Review*, 56: 6439–6475.
- Panapakidis IP, Papadopoulos TA, Christoforidis GC, et al. (2014). Pattern recognition algorithms for electricity load curve analysis of buildings. *Energy and Buildings*, 73: 137–145.
- Papadimitriou A, Tsoukala V (2024). Evaluating and enhancing the performance of the k-means clustering algorithm for annual coastal bed evolution applications. *Oceanologia*, 66: 267–285.
- Pieri SP, IoannisTzouvadakis, Santamouris M (2015). Identifying energy consumption patterns in the Attica hotel sector using cluster analysis techniques with the aim of reducing hotels' CO₂ footprint. *Energy and Buildings*, 94: 252–262.
- Ramos Ruiz G, Fernández Bandera C, Gómez-Acebo Temes T, et al. (2016). Genetic algorithm for building envelope calibration. *Applied Energy*, 168: 691–705.
- Salvalai G, Zhu Y, Maria Sesana M (2024). From building energy modeling to urban building energy modeling: a review of recent research trend and simulation tools. *Energy and Buildings*, 319: 114500.
- Uti MN, Md Din AH, Yusof N, et al. (2023). A spatial-temporal clustering for low ocean renewable energy resources using K-means clustering. *Renewable Energy*, 219: 119549.
- Verbert K, Babuška R, De Schutter B (2017). Combining knowledge and historical data for system-level fault diagnosis of HVAC systems. *Engineering Applications of Artificial Intelligence*, 59: 260–273.
- Wu S, Clements-Croome D (2007). Understanding the indoor environment through mining sensory data—A case study. *Energy and Buildings*, 39: 1183–1191.
- Yoon S (2022). Virtual sensing in intelligent buildings and digitalization. *Automation in Construction*, 143: 104578.
- Yu W, Nakisa B, Ali E, et al. (2023). Sensor-based indoor air temperature prediction using deep ensemble machine learning: An Australian urban environment case study. *Urban Climate*, 51: 101599.
- Zhang Y, Korolija I (2010). Performing complex parametric simulations with jEPlus. In: *Proceedings of the 9th International Conference on Sustainable Energy Technologies (SET 2010)*, Shanghai, China.
- Zhang C, Zhao Y, Li T, et al. (2021a). Generic visual data mining-based framework for revealing abnormal operation patterns in building energy systems. *Automation in Construction*, 125: 103624.
- Zhang L, Leach M, Bae Y, et al. (2021b). Sensor impact evaluation and verification for fault detection and diagnostics in building energy systems: A review. *Advances in Applied Energy*, 3: 100055.
- Zhao Y, Zhang C, Zhang Y, et al. (2020). A review of data mining technologies in building energy systems: load prediction, pattern identification, fault detection and diagnosis. *Energy and Built Environment*, 1: 149–164.