Energy & Buildings 329 (2025) 115269

Contents lists available at ScienceDirect

Energy & Buildings

journal homepage: www.elsevier.com/locate/enb

Investigating building stock energy and occupancy modelling approaches for district-level heating and cooling energy demands estimation in a university campus

Salam Al-Saegh *, Vasiliki Kourgiozou, Ivan Korolija, Rui Tang, Farhang Tahmasebi, Dejan Mumovic

UCL University College London, London, United Kingdom

ARTICLE INFO

Keywords:
Building Stock Energy Simulation
Occupancy Modelling
District-Level Energy Demand
Energy Data-Driven Occupancy Schedule (EDDOS)
Heating and Cooling Energy Demand
Dynamic Thermal Simulation
Energy Performance
Urban Scale Energy Modelling
District Energy Systems (DES)
Energy Demand Assessment

ABSTRACT

The urgency of decarbonizing the built environment requires precise modeling of building stock energy performance for effective large-scale planning and retrofitting. Despite advancements in data and modeling techniques, uncertainties persist in balancing model complexity and accuracy, especially in representing occupancy patterns and their impact on energy demand at district and urban scales. This study examines various approaches to building stock energy simulation and occupancy modeling for district-level heating and cooling energy demand, using 19 buildings at a Central London campus as a case study. Five scenarios were evaluated: Scenario A employs THERMOS, a data-driven approach; Scenario B uses a single dynamic thermal simulation model for the entire inventory; Scenario C applies a thermal model with a uniform occupancy schedule across all buildings; Scenario D uses a thermal model with five distinct occupancy profiles; and Scenario E assigns unique occupancy profiles based on energy use data. Results showed that Scenario E's annual heating demand estimation closely matched metered data (12 % difference), while Scenario A underestimated by 44 %. Complex occupancy models improved peak heating load predictions, with Scenario E showing only a 4 % difference from metered data, though it may not always be feasible due to data and computational constraints. Scenario D emerged as a promising balance between accuracy and efficiency. For cooling demand, significant differences among scenarios (56.43 to 6.1 kWh/m^2/Y) underscored the importance of accurate occupancy modeling. This research identifies the optimal balance between model complexity and prediction accuracy, introduces the Energy Data-Driven Occupancy Schedule (EDDOS) method, and highlights the potential of data-driven approaches to enhance energy demand assessments.

1. Introduction

As buildings account for approximately 40 % of global energy consumption [1], decarbonisation of buildings at scale is a necessity to tackle climate change. Previous studies have demonstrated that the building sector's zero carbon targets are far more within reach at the community level due to the synergies and efficiencies gained through mixed energy use, the economy of scale, and better integration of renewables [2]. Specifically, thermal energy supply and load balancing are suggested to be more effective on scales larger than individual buildings [3]. Chow [4] demonstrated that achieving zero carbon targets at the district level needs up to 70 % less capital cost as compared with retrofitting individual buildings. O'Brien [5] suggested that to achieve zero energy buildings, it is critical to consider design flexibility at the building level to allow for better interaction between buildings and neighbourhoods' energy systems and reach the best design configurations for a cluster of buildings and neighbourhoods. Walker et al. [6] also underlined the importance of planning for built environment energy systems at the neighbourhood scale as this enhances the incorporation of distributed energy systems to achieve energy neutrality.

Utilization of District Energy Systems (DESs) is an efficient way of decarbonizing the heating and cooling provided to buildings [7], which also secures the energy supply by diversifying energy sources and the demand for heating and cooling. DESs offer a unique opportunity to use large scale renewable energies and recovered heat sources that cannot

* Corresponding author.
E-mail addresses: s.saegh@ucl.ac.uk (S. Al-Saegh), vasiliki.kourgiozou.14@ucl.ac.uk (V. Kourgiozou), i.korolija@ucl.ac.uk (I. Korolija), rui.tang@ucl.ac.uk (R. Tang), f.tahmasebi@ucl.ac.uk (F. Tahmasebi), d.mumovic@ucl.ac.uk (D. Mumovic).

https://doi.org/10.1016/j.enbuild.2024.115269
Received 21 October 2024; Received in revised form 30 December 2024; Accepted 30 December 2024
Available online 7 January 2025

Nomenclature

EDDOS	Energy Data-Driven Occupancy Schedule
DES	District Energy System
NCM	National Calculation Method
AHD	Annual Heating Demand
ACD	Annual Cooling Demand
HDD	Heating Degree Days
HVAC	Heating, Ventilation, and Air Conditioning
UK	United Kingdom
UCL	University College London
3D	Three-Dimensional
ASHRAE	American Society of Heating, Refrigerating and Air-Conditioning Engineers
CO2	Carbon Dioxide
GNN	Graph Neural Network
LSTM	Long Short-Term Memory
Wi-Fi	Wireless Fidelity
nZEB	nearly Zero Energy Building

be used otherwise, and their effectiveness and carbon-saving potential grow as they expand and connect to each other.

However, arguably, the full realisation of DES potential depends on reliable modelling and assessment of their thermal efficiency, CO_2 emissions, capital and running costs, along with their resilience in providing thermal comfort for building occupants under changing climatic conditions. Specifically, accurate assessment of building stock base and peak thermal loads is critical for optimum sizing of district energy networks. Oversizing increases initial and running costs and reduces efficiency, and under-sizing could affect occupants' comfort significantly.

Modelling DESs presents several challenges. First, the complexity of these systems, with multiple heat sources, a complex distribution network, and many end users, demands powerful computing resources and laborious efforts for reliable modelling and optimisation. Second, the fluctuating and complex nature of renewables on the supply side adds further challenges to modelling and simulation. Thirdly, capturing the fluctuating nature of space heating and cooling demands in buildings, which vary based on indoor and outdoor environmental conditions along with occupants' energy-related behaviour, poses another challenge [8]. This latter point means that the accuracy and quality of heat demand models of individual buildings are crucial for the overall accuracy of DES models [9].

In this context, occupant behavior is one of the most critical variables in zero carbon design especially due to the improvement of buildings' physical characteristics and energy systems. [10–14]. Many studies have demonstrated that occupants play a significant role in shaping energy demand in buildings [15–17], and are identified as the main source of the discrepancy between estimated and actual energy consumption and the resulting uncertainty in evaluating building energy consumption (see, for example, [18–20].

However, whereas building stock modelling has matured relatively in terms of using dynamic thermal simulation tools and relying on detailed and accurate geometric definitions of buildings, representation of occupants and their interactions with building environmental control systems remains a challenge. This aligns with the challenge underlined by previous studies that using single archetypes to represent all buildings in a segment of the building stock can result in a loss of detail and accuracy in the model [21]. In particular, the paucity of high-resolution data on occupants' energy-related behaviour at the urban scale leads to major difficulties in accurate and granular estimation of buildings' energy consumption [22], leading to a more challenging performance gap at the urban scale [23].

The main concern with occupant behaviour on the urban scale is the diversity among buildings of the same occupancy activity type, which should be accounted for to obtain realistic energy demand patterns [23]. Many studies in this area have primarily relied on standard peak values and schedules for occupancy and use of lights, equipment, and systems consistently across their sample size [24]. This approach tends to overlook the large diversity in the operation of buildings by occupants [25], which can result in an overestimation of occupancy loads with a direct impact on the estimation of energy demands for district energy networks. For example, in their study, Happle et al. [23] demonstrated that the use of uniform schedules and fixed set points lead to an overestimation of peak cooling loads.

Other efforts to model occupants in building stock modelling have integrated computationally advanced stochastic models of occupant presence and behaviour. These models are meant to, and could potentially, enhance the representation of occupants in building performance models. But, as suggested by other studies [26,27], the reliability of these models and their cost-benefit especially for the purpose of stock modelling are rather debatable. Particularly, integration of these computationally expensive models into building stock energy models (which by their nature are computationally demanding themselves) is incredibly challenging. This has led many researchers to rely on single runs of stochastic models (albeit across sub-hourly timesteps of annual simulations), which could largely undermine the benefits of stochastic models to yield probable ranges of building performance indicators (rather than single values of these indicators).

In a wider perspective, recent advances in data technologies have provided opportunities for a deeper understanding of building occupant patterns [28]. For example, the use of passive Wi-Fi sensing methods has shown potential for estimating occupant behaviour on an urban scale, providing a low-cost and privacy-friendly solution [29]. However, despite these developments, there is still a need for more comprehensive and high-quality datasets [30] along with the development and examination of fit-for-purpose occupancy modelling approaches to improve the accuracy of building energy modelling at urban scale, which is of critical importance for demand-side modelling in design and assessment of DESs.

Occupancy modelling plays a critical role in forecasting energy consumption within buildings. Two difficulties delay the credibility and efficacy of occupant behavior models in evaluating energy consumption and efficiency at the aggregate level:

1. Data Inadequacy: Obtaining comprehensive occupancy data is limited, which impairs the ability to gather essential occupancy information [24,31,32]
2. *Inadequate Assumptions*: simplifying assumptions to depict the complexities of real-world occupant behavior, potentially resulting in mistaken predictions of energy usage and comfort levels [33].

Selecting the right level of representing Occupancy Profiles is a delicate matter, as fixed occupancy profiles often do not account for the variability and diversity in occupancy within the same buildings' occupancy profile due to factors such as seasonality and other factors. This can result in substantial errors in estimating building energy consumption [24]. Modelling this diversity is challenging, as it needs detailed information about individual occupants' activities and thermal comfort preferences [31]. Most existing models model aggregate occupant behavior schedules, rather than capturing the diversity among different social demographics and building types [33].

The complexity of these models can range from simple static assumptions to detailed dynamic simulations. The choice of modelling approach affects the accuracy of energy demand estimations, which in turn influences the effectiveness of decarbonization strategies in district energy systems.

To conclude, the main research gap identified in literature was the miss-representation of occupancy diversity at stock level and lack of reliable alternatives for stochastic models at stock level. In this context, the research aims to answer the following research questions: 1. What is the ideal balance between model complexity and prediction accuracy in the context of district level energy demand estimations? 2. How does the use of different occupancy modeling techniques affect energy demand assessments? 3. To which extent energy data-driven occupancy schedules could enhance the reliability of energy demand assessments at district level as an alternative to stochastic models?

2. Method

2.1. Overview

Five different scenarios of modeling occupant behavior and building stock energy were proposed to conduct comparative analysis. The five scenarios (A-E) are visualized in Fig. 1. These scenarios range from simple and fast (Scenario A) to complex and time-consuming (Scenario E).

These models aim to estimate annual and peak heating and cooling demands, which are key performance indicators (KPIs) for designing District Energy Systems (DES) and built environment decarbonization. Although the study focuses on a sample of 19 buildings, the employed building models represent typical characteristics of building stock models. This allows the findings to be applied to large-scale building stock energy simulations, identifying the most cost-effective occupancy models for integration into computationally intensive building stock energy modeling endeavors.

Scenario A represents a data-driven method with the tool producing its own deterministic occupancy schedules. While Scenarios (B-E) use a bottom-up method to estimate energy demand in buildings through physics-based modelling, specifically using EnergyPlus software to generate dynamic thermal models. The scenarios of the bottom-up method are independent of historical data but require numerous inputs and high computing data and effort. The advantages of a bottom-up approach include the ability to simulate various retrofitting and future weather scenarios. In this research, it was assumed that, due to the nature of educational buildings, the occupants are unable to change the fixed thermostat settings. Occupancy modeling methods for building stock energy demand can be categorized into deterministic, stochastic, and data-driven approaches. Deterministic methods use fixed input parameters and predefined schedules, providing a straightforward but less flexible approach. Stochastic methods introduce randomness to account for variability in occupant behavior and environmental conditions, enhancing the realism of simulations. Data-driven approaches leverage actual energy consumption data to generate occupancy patterns, offering high accuracy and adaptability. In this research two occupancy modelling approaches were implemented within scenarios (B-E) building models: deterministic standard-based schedules and a novel energy data-driven occupancy model. Stochastic occupancy behavior modeling at the stock level presents significant computational challenges due to the inherent complexity and variability of occupant behaviors [34–36]. So, it has been excluded from this research.

2.2. Case study buildings

Nineteen buildings of UCL's Bloomsbury campus, were selected based on the data availability and stock representation of various types of occupancy types and Heating Ventilation and Air Conditioning (HVAC) systems. HVAC systems were defined by Display Energy Certificate (DEC) as in Table 1.

The five distinct occupancy types used in scenarios (B-E) were based on the National Calculation Method (NCM) include: Administration office (A), Teaching classroom (T), Library (Lib), Chemistry/Physics Lab

Table 1
HVAC Type.

DEC building environment categories	HVAC type
Heating and mechanical ventilation (Air conditioning)	HVAC 1: Heating, cooling, and mechanical ventilation system
Heating and natural ventilation (Air conditioning)	HVAC 2: Heating, cooling, and natural ventilation system
Heating and mechanical ventilation	HVAC 3: Heating and mechanical ventilation system
Heating and natural ventilation	HVAC 4: Heating system and natural ventilation

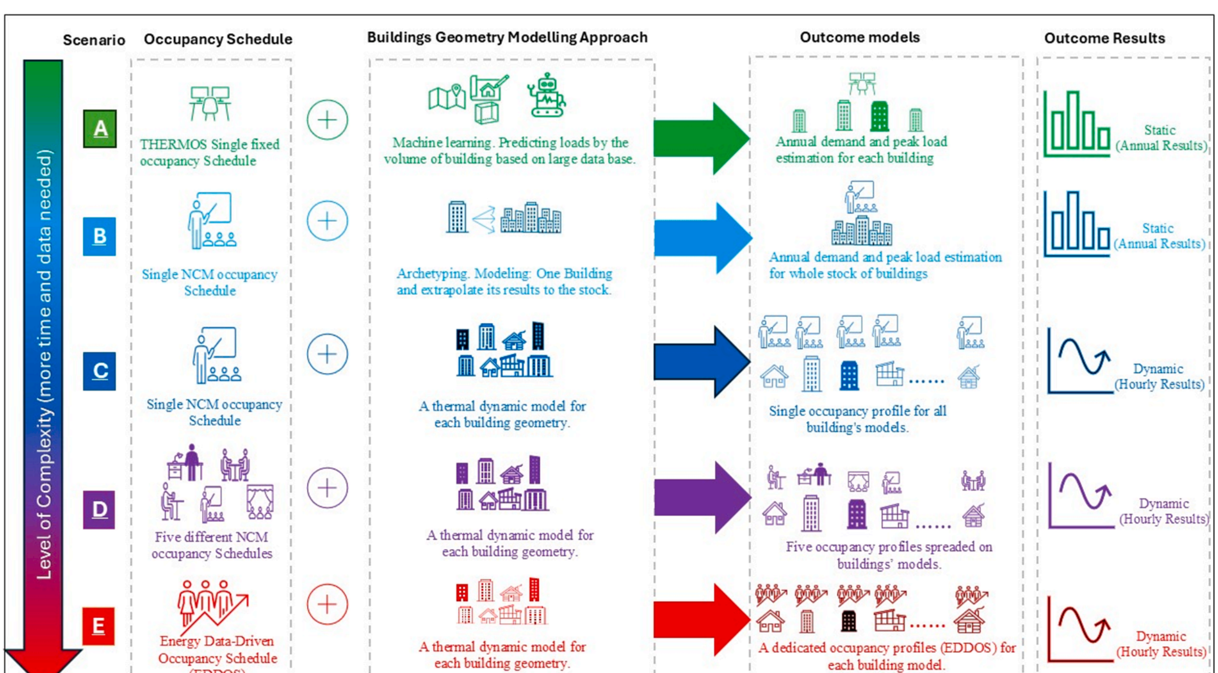

Fig. 1. Simulation Scenarios (A-E) Comparative Analysis of building stock energy and Occupancy Modeling Approaches.

(LC), Engineering Lab (LE), and Biology/Medicine Lab (LM). Academic Chemistry/Physics Lab (LC) profile is not included, but it has a similar profile to LE & LM profiles. See Table 2.

2.3. Building performance indicators and evaluation of occupancy models

These indicators represent the annual total and hourly peak amount of heating and cooling energy that must be delivered to the buildings to maintain the heating and cooling set-points.

$$AnnualHeatingDemand = \sum_{t=1}^{8760} Q_{heating}(t)$$

$$AnnualCoolingDemand = \sum_{t=1}^{8760} Q_{Cooling}(t)$$

The estimated heating demand values were then multiplied by an efficiency factor and compared with the metered heating energy to assess the accuracy of building model estimations under different occupancy modelling approaches. In terms of cooling, however, as the UCL building stock is not served with a central cooling system, the estimation of cooling demand only served to discuss the implications of different occupancy modelling approaches on a theoretical level (and to inform future development of the DES serving the building). Previous research demonstrated that models calibrated using heating season data could provide acceptable cooling load predictions but emphasized the need for careful consideration of model parameters that affect both heating and cooling performance [37]. Raftery et al. (2011) investigated the transferability of calibrated energy models. They found that models calibrated for one season (e.g., heating) could provide reasonable predictions for another season (e.g., cooling)[38]. Coakley et al. (2014) reviewed calibration techniques for building energy simulation and noted that heating-calibrated models can provide insights into cooling performance [39]. Other researchers considered a similar approach [40].

Table 2
List of the 19 buildings were chosen for stock buildings test.

Building Name	Area in m^2	Occupancy Type	HVAC
25 Gordon Street	6,497	Administration office (A)	UCL – HVAC 4
Central House	5,068	Administration office (A)	UCL – HVAC 3
IOE – 24–28 Woburn Square	781	Administration office (A)	UCL – HVAC 4
22 Gordon Street	7,252	Engineering Lab (LE)	UCL – HVAC 1
Malet Place Engineering Building	1,212	Engineering Lab (LE)	UCL – HVAC 1
Physics Building	5,071	Engineering Lab (LE)	UCL – HVAC 3
26 Bedford Way	710	Library (Lib)	UCL – HVAC 2
Bloomsbury Theatre	3,653	Library (Lib)	UCL – HVAC 4
DMS Watson Building	5,686	Library (Lib)	UCL – HVAC 3
Alexandra House	2,025	Biology/Medicine Lab (LM)	UCL – HVAC 3
Andrew Huxley Building	1,963	Biology/Medicine Lab (LM)	UCL – HVAC 1
Cruciform Building	14,509	Biology/Medicine Lab (LM)	UCL – HVAC 1
Rayne Institute	5,391	Biology/Medicine Lab (LM)	UCL – HVAC 3
1–4 Malet Place	8,573	Teaching classroom (T)	UCL – HVAC 4
16–18 Gordon Square	1,160	Teaching classroom (T)	UCL – HVAC 4
33–35 Torrington Place	1,039	Teaching classroom (T)	UCL – HVAC 3
Egyptology	1,090	Teaching classroom (T)	UCL – HVAC 3
Foster Court	2,471	Teaching classroom (T)	UCL – HVAC 3
Gordon House	1,751	Teaching classroom (T)	UCL – HVAC 3

2.4. Building stock energy and occupancy modelling scenarios

2.4.1. Scenario A- Data driven approach

This Scenario is based on the data driven approach using a tool called THERMOS. So, this tool was selected to represent the data-driven approach for estimating thermal demand at district level. This web-based, free-access tool quickly analyzes different thermal network possibilities [41]. THERMOS uses advanced machine learning modeling methods and algorithms, as well as real data from thermal networks, to estimate annual heat demand in kWh/yr and peak heat demand in kW. The building's peak heat demand is calculated from its annual heat demand using a linear model. The equation was derived from a large sample of published UK half-hourly domestic gas consumption data. The main value estimated is the annual heat demand, which consists of space heat demand and hot water demand, calculated using different models and combined [42].

THERMOS has a deterministic occupancy schedule that can be edited for the entire building stock collectively. It does not explicitly account for occupant behavior at the individual building level, which limits its ability to provide advanced insights into the role of occupant behavior modeling.

2.4.2. Scenario B – Simplified archetyping

This approach utilizes a streamlined archetyping strategy, where a single exemplary building thermal model is applied across the entire building inventory. The process involves constructing representative building archetypes, which serve as templates to forecast energy usage and evaluate the impact of the building sector on energy efficiency. These archetypes are simplified representations of actual buildings, chosen based on their dominant characteristics and energy consumption patterns. The objective is to create a manageable and easily comprehensible model that provides a broad overview of the energy efficiency performance of various building types within the sector. This approach allows for a simplified analysis of the building stock, enabling the identification of areas where energy savings can be achieved and the prioritization of interventions to improve energy efficiency on a large scale.

In this research, five archetypes of existing buildings were selected to represent the entire building stock: Alexandria House, Andrew Huxley, Physics Building, Cruciform Building, and Rayne Institute. These selected archetypes were analyzed using the thermal simulation tool EnergyPlus to predict the annual and peak heating and cooling demands for each building. The results of the five thermal simulations were averaged to obtain a single value (mean archetype) for each load, which was then extrapolated to the stock of 19 buildings based on the area ratio of the mean archetype building to the whole stock. While the five models provided dynamic hourly results, only the annual values were used for extrapolation. This approach utilized a fixed standard-based occupancy schedule, namely NCM-Academic, which is further explained in the next section.

2.4.2.1. Fixed standard-based occupancy schedules. In this deterministic approach to occupancy representation, the UK National Calculation Method (NCM) [43] is used as the reference for occupancy-related assumptions. NCM is a modelling guide provided in the UK in support of building performance simulation and assessments mandated in Approved Document Part L (Conservation of fuel and power). NCM comprises of the underlying dynamic building simulation methods and standard databases for building construction and activities.

Table 3 gives the occupancy-related assumptions considered in this study. Scenario B&C will use a single schedule: "Academic." Where scenario D will use five schedules of this table 3. Fig. 2 illustrates UK NCM weekday schedules of occupancy, use of lights and equipment used in this study.

2.4.3. Scenario C – Thermal Dynamic Models with single occupancy schedules

In Scenario (C) a Bottom-up physical approach was considered, creating a building-by-building thermal dynamic model for the whole stock. The same determinist single occupancy schedule of scenario (B): NCM-Academic, was used again and assigned for each building. A Dynamic Building Stock Model (DBSM) was developed to generate building by building thermal dynamic models for UCL University Campus, see Fig. 3.

DBSM is flexible in collecting data; it relies on available data with minimal requirements. Even so, it provides the opportunity for further data, in case of availability, to fine-tune the models. DBSM synergizes the combination of established stock modelling approaches with the pioneered automation of the SimStock modelling tool [44]. This is by building a DBSM 3DStock (database source) to feed Sim stock and generate building-by-building thermal models. DBSM development consists of two parts:

2.4.3.1. DBSM 3DStock (database). The initial step involves integrating multiple datasets to establish sets of inputs that are commonly utilized by individual building energy models. These data came from a public source, and it links modelling information to gather: Building name, Energy Performance Certificate (EPC), Display Energy Certificate (DEC), Building usage, footprint, external geometrical definition, Unique Property Reference Number (UPRN) (1). Databases can be classified into: **(1) General building information**. Building names and addresses can be found from various sources such as: UCL website, Ordnance Survey (OS), Display Energy Certificate, Google Maps, campus monitoring platform (Fabric) and Valuation Office Authority (VOA). The buildings name list provided by UCL estates (UCL Website) was considered in this study, as some building names could slightly vary.

(2) Geometrical building data. Geometrical data is what makes this approach unique when compared to archetyping approaches which do not take into consideration the shape of each building. GIS-based datasets can be taken from Ordnance Survey (OS) which provides geometrical data for the building stock of Great Britain. OS provides 2D polygons for buildings in OS Master Map, topography datasets also include the average height of each entity [45].

Lidar data is a reliable source of GIS-based polygons of buildings areas. It can be used to check against OS data and fill in the missed data. Building areas can also be found from VOA, Fabric, UCL estate and, for checking purposes, Google Maps. For Window to Wall Ratio (WWR) values were proposed in guidance with a survey for London schools [46] as in Table 4. Unfortunately, specific data of WWR for university buildings has not been found in the literature.

(3) Non-geometric building data. Building construction thermal properties need to be considered carefully, as they are especially important in the modelling process for their significant impact on the results. rdSAP method [47] was used as it provides U-value assumptions for various building age groups. As in the following Table 5 buildings were categorized based on their age, and for each category there is an assumption of U-Value for each building component (Wall, Roof, Floor and Glassing).

Building activities were provided by both DEC and UCL Estate, UCL Estate information was more correct in respect of building activity taking place as in Table 6.

2.4.3.2. SimStock (Thermal models generator). SimStock, a Python-based tool, automates the creation of thermal building models. It achieves this by supplying data synthesized from the DBSM 3DStock to the EnergyPlus engine, as described in part 1. SimStock was developed at UCL, it uses EnergyPlus and directly extracts information on building geometry, materials, and activities from 3DStock [48]. The resulting thermal models are in idf format, which aligns with the requirements of the EnergyPlus simulation program. Unlike archetyping, this method does not oversimplify or disregard the unique shapes of buildings, allowing for modelling unique specific parameters of each building.

2.4.3.3. DBSM validation. In the process of validating our models, we have adopted the guidelines provided in CIBSE TM63 Operational Performance: Building Performance Modelling [49]. These guidelines pertain to single-building measurement and verification practices. While the guide was initially designed for individual buildings, we have determined that its methodology can be adapted to suit the requirements of our stock model [50], as illustrated in Fig. 4.

2.4.4. Scenario D- Thermal dynamic models with five occupancy schedules

This is a bottom-up physical approach as well, was built on scenario (C) to consider building by building thermal dynamic model for the whole stock. Five NCM occupancy schedules (A, T, LE, LM and LIB) as in Table 3 were spread on the buildings of the stock. The aim of this scenario is to evaluate the diversification effect on the energy consumption of the whole stock.

2.4.5. Scenario E- Thermal dynamic models with EDDOS occupancy schedules

This scenario uses the same approach as scenarios (C&D), which involves building thermal models for individual buildings. The occupancy schedules in this scenario are derived from each building's energy use data, resulting in a unique occupancy schedule for each building. This captures the unique occupancy patterns of each building more realistically compared to standard profiles used in previous studies.

Table 3
NCM Occupancy-related assumptions under the standard-based Approach.

Model input parameter	Administration (A)	Teaching (T)	Library (Lib)	Academic (Chemistry/Physics (LC), Engineering (LE), biology/medicine (LM))
People maximum density [persons/m^2]	0.103	0.2034	0.0986	0.10625
Occupancy schedule	C2_Uni_Office	Uni_ClassRm_Occ	LibMusGall_CellOff_Occ	Uni_Lab_Occ
Metabolic rate [W/persons]	123	140	123	160
Latent heat fraction	40 %	50 %	40 %	39 %
Lighting maximum density (Lux)	300	300	300	500
Lighting schedule	Uni_CellOff_Light	Uni_ClassRm_Light	LibMusGall_CellOff_Light	Dwell_DomToilet_Light
Equipment maximum density [W/m2]	11.99	4.74	13.91	8.73
Equipment schedule	Uni_CellOff_Equip	Uni_ClassRm_Equip	LibMusGall_CellOff_Equip	Uni_Lab_Equip
Hot Water Supply (l/day/m^2)	0.2369	0.4678	0.1965	0.244375

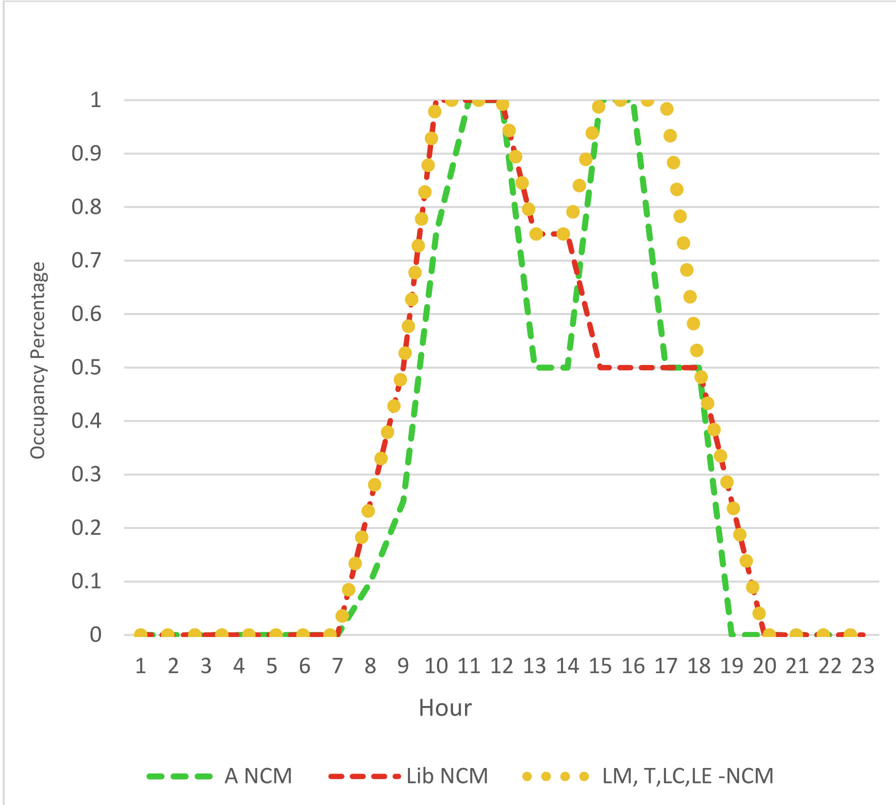

Fig. 2. Weekday schedules of NCM occupancy schedules.

The new procedure was pioneered to generate Energy data driven occupancy schedules (EDDOS) [35], it aims to model occupancy patterns and represent diversity among buildings more realistically, without relying on a stochastic modelling process, First, the extent to which energy use data correlated with the hours of the day. While mete red gas data did not suggest a correlation with the hours of the day, electricity use showed a good correlation (with an R-squared of 0.84) in the studied building Fig. 5. On this basis and following the periodic pattern of weekly electricity use in the building see Fig. 6, an algorithm was set up to translate the fluctuating electricity use into occupancy schedules. To this end, the minimum and maximum values of hourly electricity consumption were identified in each week, and these were assigned to the lowest and highest occupancy levels respectively Fig. 6. Between these two ends, the occupancy variations were assumed to follow the changes in electricity use linearly. To account for weather-dependant and seasonal changes in electricity use, each week data was processed separately. Thus, the procedure returned non-repeating daily profiles of occupancy corresponding to the variations of electricity use. The Procedure of generating EDDOS schedule is explained in Fig. 7. For this study, year 2019 electricity use data was used to identify the occupancy profiles, as this represented a building usage pattern before Covid-19 pandemic. 2019 was the only year before COVID-19 that had all the necessary data for our research, including hourly gas and electricity meter readings from nineteen buildings.

Furthermore, the building models were simulated for the year 2019, as it represents an occupancy pattern before the Covid-19 pandemic. This also allows the study to assess the performance of the energy data-driven model based on the year used for its training (2019).

3. Results

3.1. Annual heating demand

Fig. 8 presents a comparison of annual heating loads for 19 buildings at University College London's Bloomsbury campus, using different modelling scenarios (A-E) against metered data from the Fabric system and CIBSE good Practice. This represents the actual metered data and serves as the baseline for comparison. (A) Thermos: A data-driven scenario using machine learning, showed a 16 % underestimation compared to the metered data, which is similar to CIBSE good practice. (B) Archetyped: A simplified scenario using a single archetype for all buildings, resulting in a significant 44 % underestimation. (C) Single Occupancy: A more detailed scenario using individual building thermal models but with a single occupancy schedule, showing a 31 % underestimation. (D) 5 Patterns/Stock: This scenario incorporates five distinct occupancy profiles, resulting in a 17 % underestimation. (E) Single pattern/Building: The most complex scenario, assigning a unique occupancy profile to each building, showing the closest estimation with only a 12 % underestimation.

These results demonstrate several key findings: 1. Increasing complexity in occupancy modelling generally improves the accuracy of annual heating load predictions. This is evident in the progression from the highly simplified Archetyped scenario (B) to the more complex Single pattern/Building scenario (E). 2. The data-driven Thermos scenario (A) performs well, with only a 16 % underestimation. This suggests that machine learning techniques can provide relatively accurate estimates with less computational complexity. 3. The significant improvement in accuracy between the Archetyped scenario (B) and the Single Occupancy scenario (C) highlights the importance of using individual building thermal models rather than a single archetype for the entire stock. 4. The further improvements seen in scenarios D and E underscore the value of incorporating more detailed occupancy profiles. This aligns

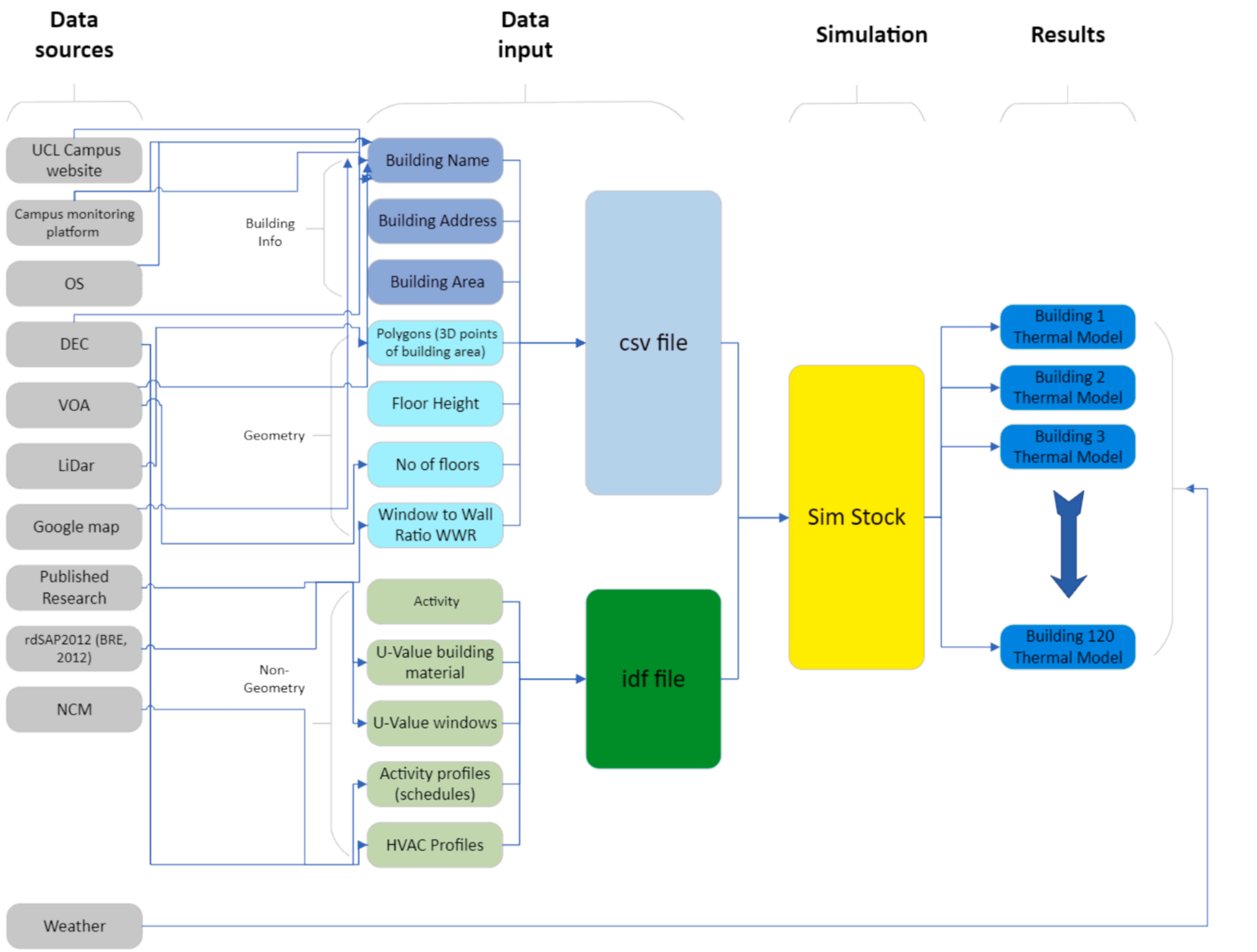

Fig. 3. DBSM development.

Table 4
Window to wall ratio Assumptions.

Campus stock model age bands	Window to Wall Ratio (WWR)
Pre 1914	33 %
1918–1939	35 %
1945–1980	38 %
Post 1980	30 %

with the study's emphasis on the importance of occupancy modelling in building energy simulations. 5. The most complex scenario (E) provides the closest estimate to the metered data. However, the marginal improvement over scenario D (5 Patterns/Stock) raises questions about the trade-off between increased computational complexity and improved accuracy.

3.2. Peak heating demand

Fig. 9 illustrates the comparison of peak heating demand estimates across different modelling scenarios for the set of 19 buildings. The "Fabric (metered)" value of 5,450 kW represents the actual measured peak heating demand, serving as the reference point for assessing the accuracy of different modelling approaches. scenario A (Thermos): This data-driven approach underestimates the peak demand by 43 %, suggesting limitations in capturing peak loads using machine learning techniques. Scenario B (Archetyped): This simplified archetype method significantly overestimates the peak demand by 53 %, highlighting the risks of oversimplification in building stock modelling. Scenarios C-E:

Table 5
U-Values assumptions.

DEC building age categories	rdSAP age bands	Campus stock model age bands	Wall U-Values (W/m^{2}K)	Roof U-Values (W/m^{2}K)	Floor U-Values (W/m^{2}K)	Glazing U-Values (W/m^{2}K)
Pre-world war I (Pre 1914)	A. Before 1900, B. 1900–1929	Pre 1914	1.7	2.3	1.5	4.8
Inter war (1918–1939)	B. 1900–1929, C. 1930–1949	1918–1939	1.7	2.3	1.5	4.8
Post-war regeneration and expansion (1945–1980)	C, D, E, F: 1930–1982	1945–1980	1.35	1.5	1.4	4.8
Modern (post-1980)	G, H. I, J, K, L: 1983–2012	Post-1980	0.4	0.4	0.94	3.1

Table 6
Occupancy (activity) schedules.

UCL DEC Tracker Building Use		NCM Profiles Associated				
A	Admin	Uni_CellOff_Occ	Uni_CellOff_Light	Uni_CellOff_Equip	Uni_CellOff_Heat	Uni_CellOff_Cool
T	Teaching	Uni_ClassRm_Occ	Uni_ClassRm_Light	Uni_ClassRm_Equip	Uni_ClassRm_Heat	Uni_ClassRm_Cool
LIB	Libraries	LibMusGall_CellOff_Occ	LibMusGall_CellOff_Light	LibMusGall_CellOff_Equip	LibMusGall_CellOff_Heat	LibMusGall_CellOff_Cool
LC	Lab-Chemistry	Uni_Lab_Occ	Uni_Lab_Light	Uni_Lab_Equip	Uni_Lab_Heat	Uni_Lab_Cool
LE	Lab-Engineering	Uni_Lab_Occ	Uni_Lab_Light	Uni_Lab_Equip	Uni_Lab_Heat	Uni_Lab_Cool
LM	Lab-Medicine	Uni_Lab_Occ	Uni_Lab_Light	Uni_Lab_Equip	Uni_Lab_Heat	Uni_Lab_Cool
R	Residential	Uni_Bed_Occ	Uni_Bed_Light	Uni_Bed_Equip	Uni_Bed_Heat	Uni_Bed_Cool

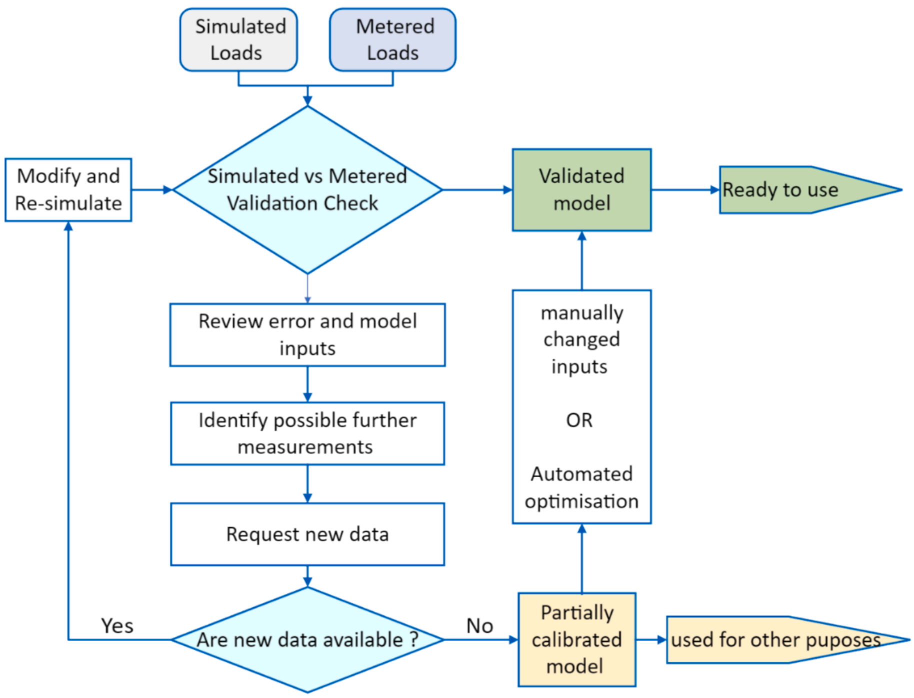

Fig. 4. DBSM Validation process.

These represent increasingly complex dynamic simulation methods with varying occupancy modelling strategies. The transition from a single occupancy type (C) to multiple occupancy patterns (D) improves the peak demand estimate, reducing overestimation from 17 % to 5 % for diversified (Div) calculations. Scenario (E) is the most detailed scenario, using building-specific occupancy profiles, shows mixed results with slight improvements in some cases but not consistently better than scenario D.

For Aggregated vs. Diversified Peaks: scenarios C-E, both aggregated (Agg) and diversified (Div) peak calculations are presented. Diversified peaks generally provide closer estimates to the measured value, highlighting the importance of considering load diversity in district energy system planning. Fig. 10, explain how peaks are calculated on hourly basis.

There is a general trend of improved accuracy from scenarios A to E, except for the archetype scenario (B). This suggests that increasing the complexity of occupancy modelling and using dynamic simulations can lead to more accurate peak demand estimations.

Scenarios E and D show the closest matches, within 4 % and 5 %, respectively. This indicates that using multiple occupancy patterns may offer a good balance between model complexity and accuracy for peak load estimation.

3.3. Annual and peak cooling demands

Fig. 11 illustrates the results of a comparative study on different scenarios to modeling cooling energy demand for the 19 buildings. Due to the absence of a dedicated cooling meter, the study uses the model (E) as the baseline for comparison as it shown best accuracy for heating loads. Comparisons are for reference purposes only and do not represent calibrated results for cooling loads.

Fig. 11 presents five different scenarios to modeling cooling demand: (A) THERMOS: A data-driven approach using machine learning, estimating 56.4 kWh/m2/year, 825 % higher than the baseline. The data for space cooling demand is limited in Europe [51], this makes the Machine learning tools like THERMOS not suitable to predict cooling, so THERMOS is showing unrealistic results. (B) Archetyped: A simplified scenario using a single archetype for all buildings, estimating 21.3 kWh/m2/year, 249 % higher than the baseline. (C) Single Occ: Using individual building thermal models but with a single occupancy schedule, estimating 15.35 kWh/ m2/year, 152 % lower than the baseline. (D) 5 Patterns/Stock: Incorporating five distinct occupancy profiles, estimating 7.17 kWh/year, 17 % higher than the baseline. (E) Single pattern/Building: The most complex scenario, assigning a unique occupancy profile to each building, estimating 6.1 kWh/m2/year, which is set as the baseline (0 % difference). CIBSE benchmark is 14 kWh/m2/Year. These results reveal several important insights: There is a substantial variation in cooling demand estimates across different modeling scenarios, ranging from 6.1 kWh/m2/year to 56.43 kWh/m2/year. This highlights the critical impact of modeling methodology on cooling demand predictions. The THERMOS scenario significantly overestimates cooling demand compared to other methods. This could be due to

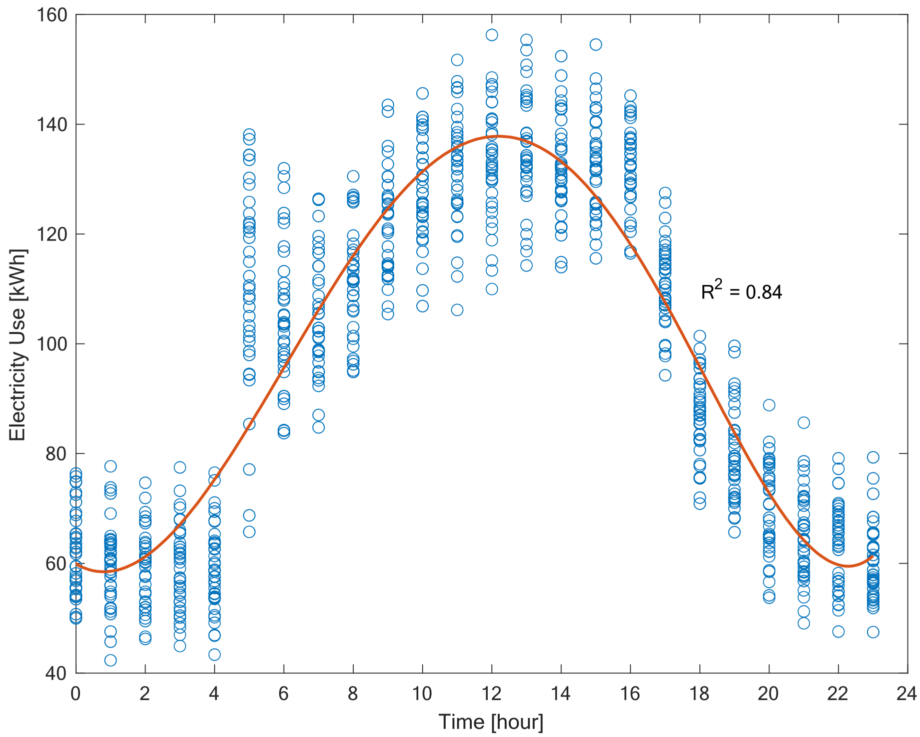

Fig. 5. Electrical consumption versus hours of the day for the year 2019.

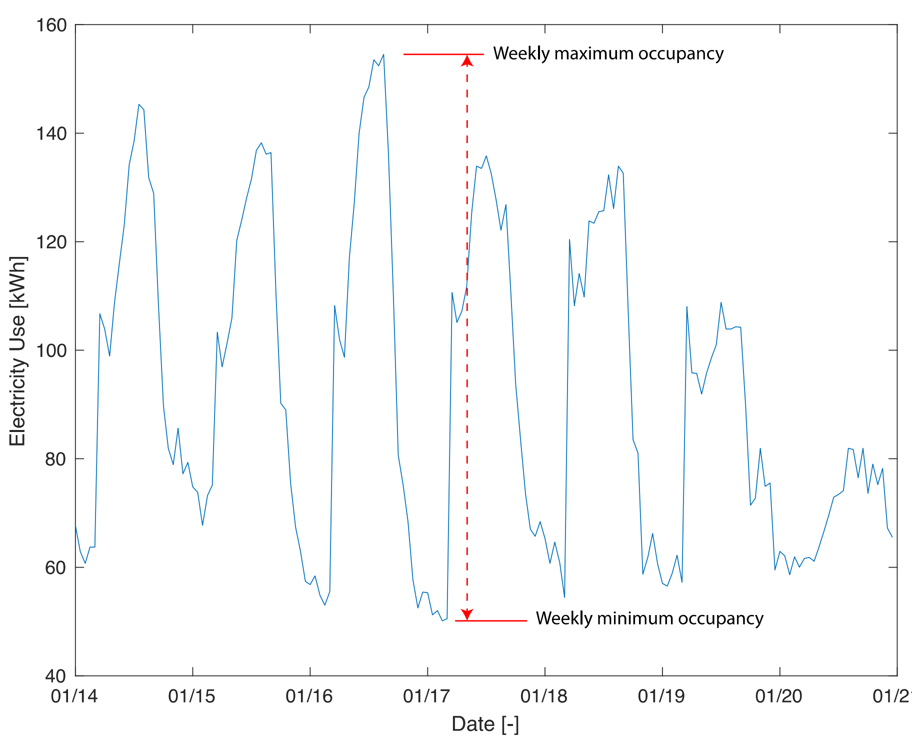

Fig. 6. An example weekly profile of electricity consumption and assignment of minimum and maximum occupancy based on minimum and maximum electricity use.

limitations in the machine learning model's training data or assumptions, particularly for cooling in the specific context of the UCL campus. The progression from the Archetyped scenario to more complex occupancy models (Single Occ, 5 Patterns/Stock, Single pattern/Building) shows a general trend of decreasing cooling demand estimates. This suggests that more detailed occupancy modeling tends to result in lower cooling demand predictions. The estimates from the two most complex scenarios (D and E) are close, indicating a potential convergence of results as model complexity increases. This suggests that the additional computational effort required for the most complex scenario (E) may not provide significantly different results from the slightly simpler scenario (D).

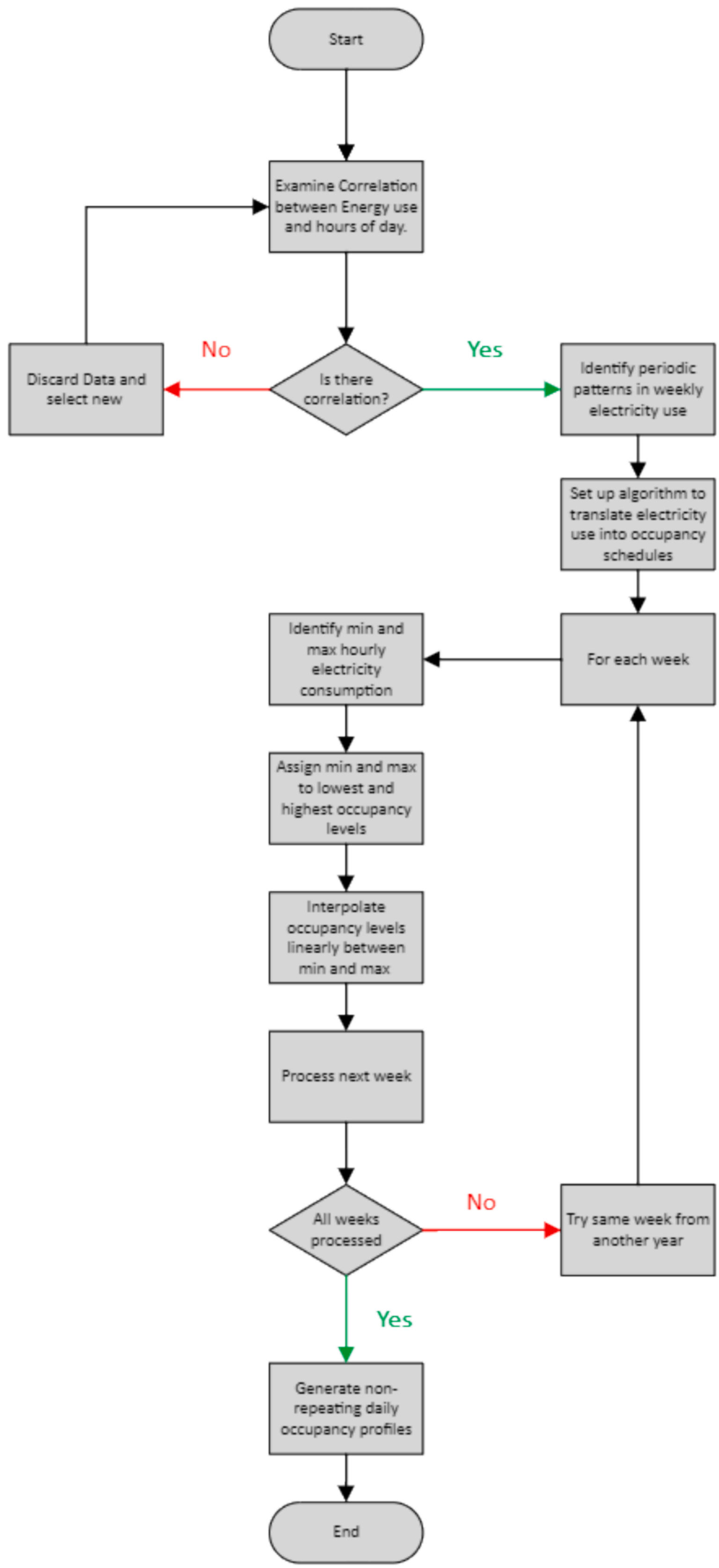

Fig. 7. EDDOS Development Procedure.

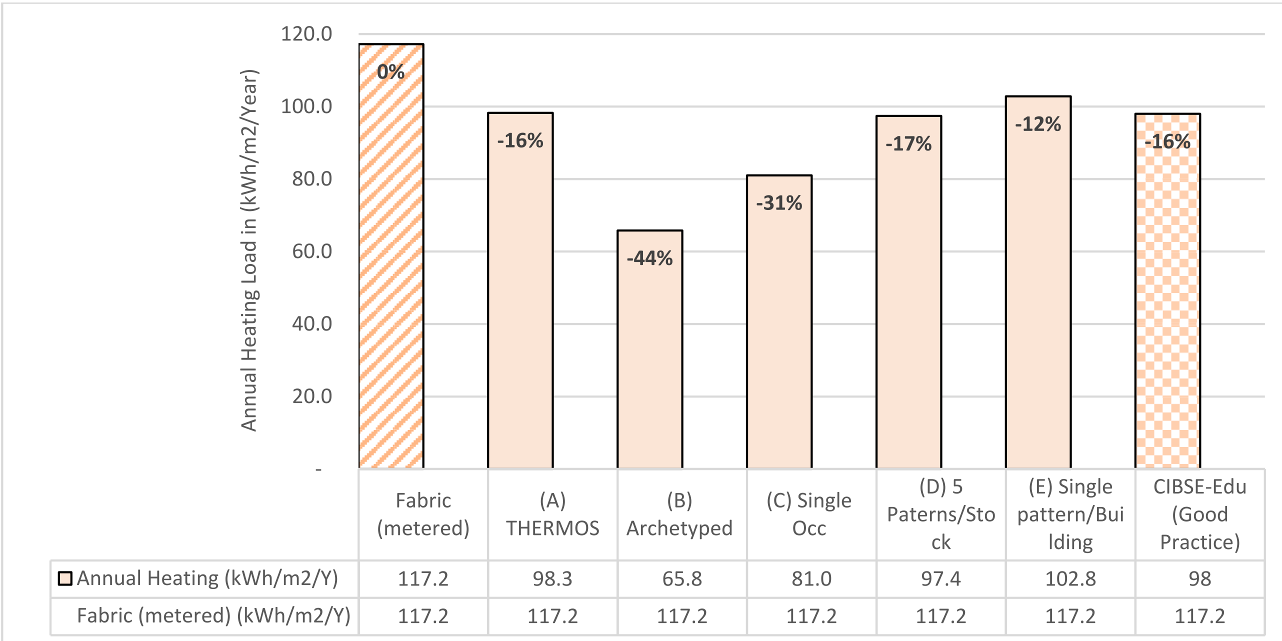

Fig. 8. Annual Heating Load ((kWh/m2/Year)) for nineteen buildings.

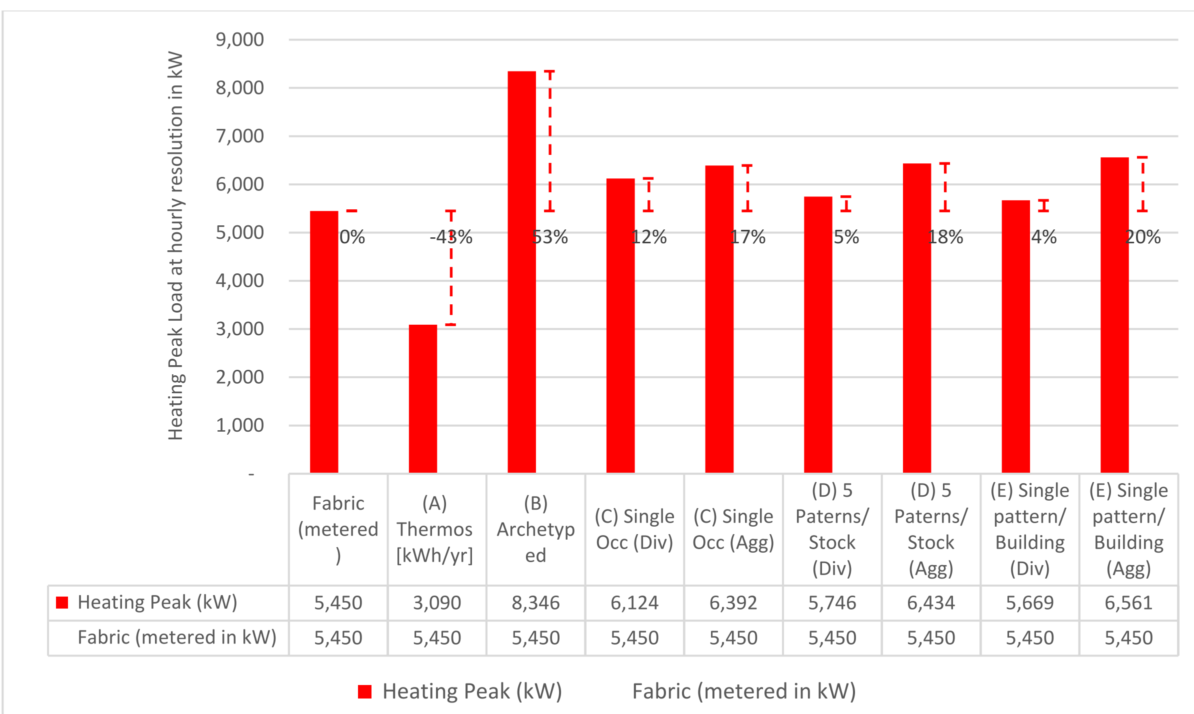

Fig. 9. Staggered vs aggregated Peak Heating Loads For stock of 19 Buildings (in kW).

Fig. 12 compares peak cooling load estimates for the set of buildings. The results reveal significant variability in peak cooling load estimates across different modelling scenarios. The THERMOS scenario (A), a data-driven method, shows the highest estimate at 15,069 kW, which is 530 % higher than the baseline (Scenario E Div.). This substantial overestimation suggests potential limitations in its application to cooling load estimation for this specific building stock. The Archetyped scenario estimates 7,053 kW, 195 % above the baseline, while the Single Occ scenario yields 6,073 kW (diversified) and 6,141 kW (aggregated), 154 % and 157 % above the baseline respectively.

The 5 Patterns/Stock scenario estimates 2,302 kW (diversified) and 2,383 kW (aggregated), only −4% and −0.3 % above the baseline. The most complex Single pattern/Building scenario predicts 2,390 kW for diversified (with zero difference and serving as the baseline) and 2,514 kW for aggregated, which is 5 % above the baseline. This pattern indicates that more detailed occupancy modeling tends to result in lower peak cooling load predictions.

The graph also presents both diversified (Div) and aggregated (Agg) peak calculations for the more complex scenario. Consistently lower diversified peak estimates compared to aggregated ones (e.g., 2,302 kW

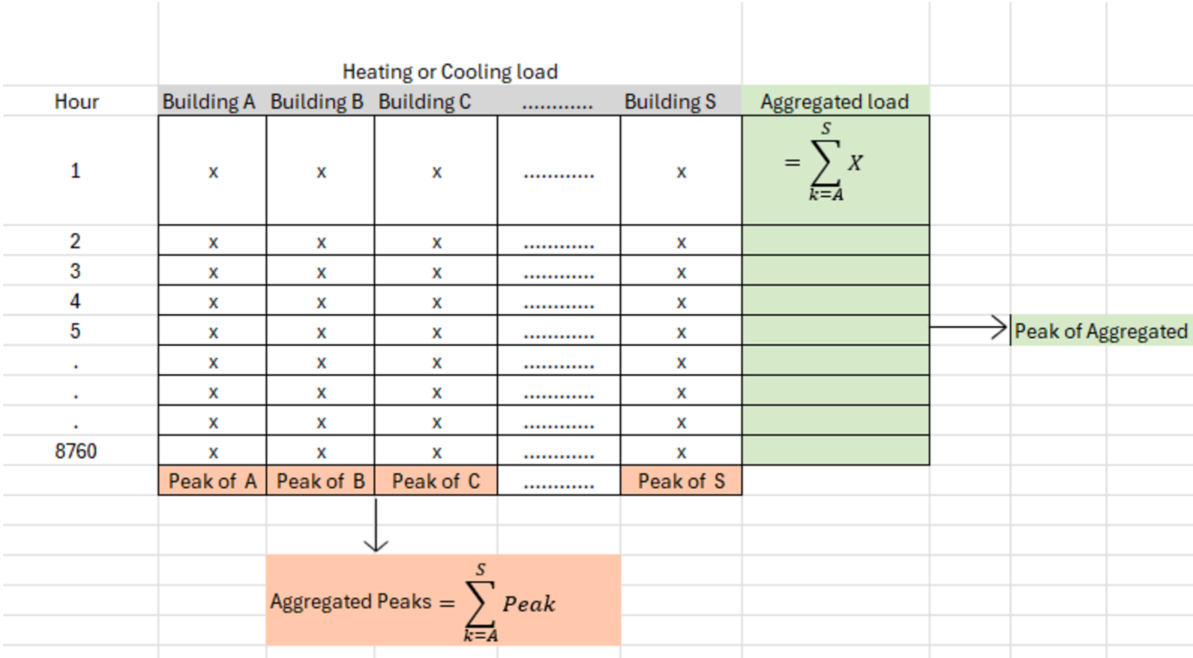

Fig. 10. Peak of aggregated (diversified) vs aggregated peak (non-diversified).

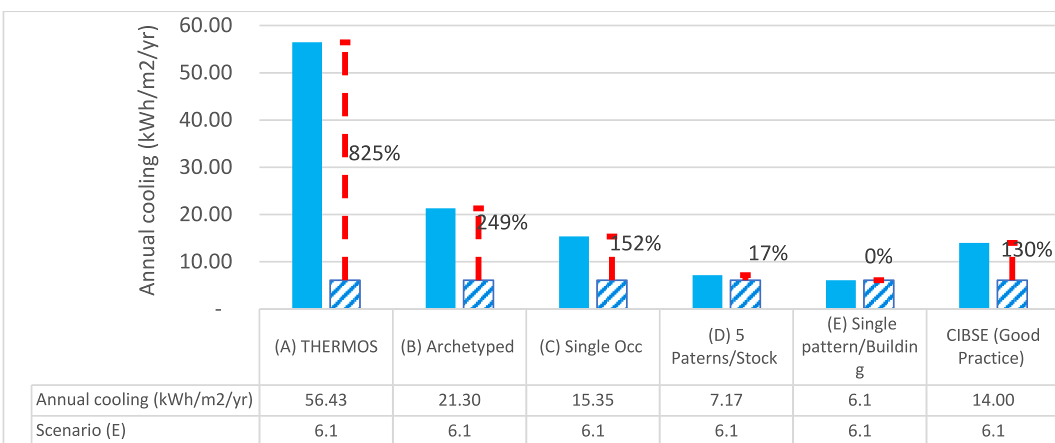

Fig. 11. Annual Cooling load for 19 Buildings (in kWh/m2/Year).

vs 2,383 kW for 5 Patterns/Stock) emphasize the importance of considering load diversity in district cooling system design. The convergence of results between the 5 Patterns/Stock and Single pattern/Building scenarios (2,302 kW vs 2,390 kW for diversified peaks) suggests a potential optimal point of complexity in modeling.

3.4. The role of occupancy profiles

From the previous results, it was observed that the fluctuation in the outcomes is higher among scenarios A, B, and C, where completely different approaches for modeling building energy stocks were considered in each scenario. On the other hand, there are much fewer differences between scenarios D and E, as the same building-by-building modeling approach was used in both. The main difference was in occupancy modeling approach, where a unique building-by-building occupancy model was used in scenario E, while scenario D considered the 5 NCM deterministic schedule. So, this led to focus on the differences between NCM schedules and EDDOS. As this has a significant importance in understanding diversification effect at stock level.

Figs. 13-18 illustrate the daily averaged occupancy profiles over one year resulted from generating different Energy Data Driven Occupancy Schedules (EDDOS) compared with UK-NCM occupancy schedules investigated in this study.

As it can be seen in the figures that EDDOS averaged profiles are lower that NCMs in all occupancy profiles. Fig. 13, Fig. 15, Fig. 16, are showing two different EDDOS occupancy profiles of each building for the same occupancy profiles (T, LE, LIB). This confirms the purpose of this study of having a unique profile of each building to assess the effect of diversification of occupants. The other profiles: T, LM and LC has only one suitable building electricity data to generate the required EDDOS file, so one profile was generated for each profile.

Fig. 19 explains the principle of occupancy Fraction-Hours (F-H) per year of each schedule. Real EDDOS generated values of (F-H) for buildings are shown as in Fig. 20. This research aims to examine EDDOS approach at the stock level by checking the effect on building energy performance and how close to real metered data. Fig. 20 also shows that the general trend of EDDOS Fraction-Hours per year is lower than NCM. EDDOS Average is 1823, while NCM average is 2187. This trend can be

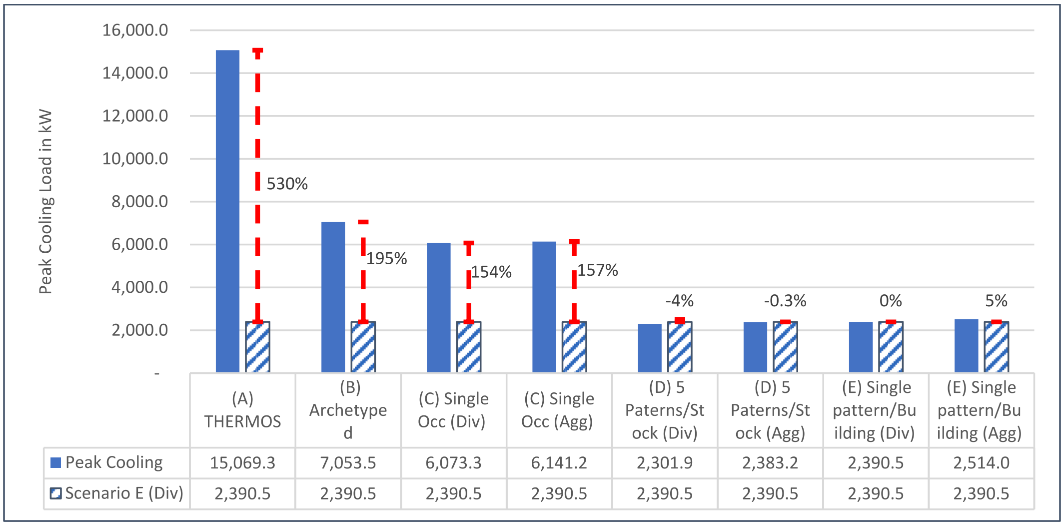

Fig. 12. Cooling Peak Loads in kW.

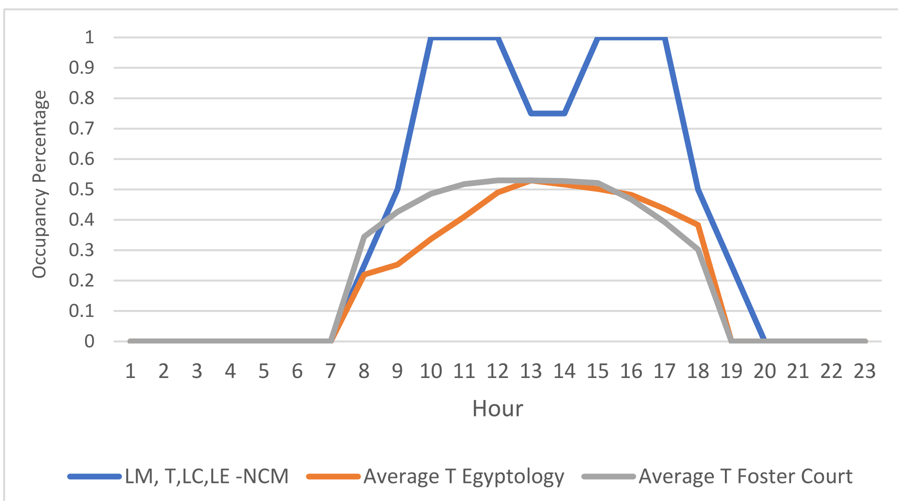

Fig. 13. Averaged EDDOS Teaching (T) Occupancy profile vs NCM.

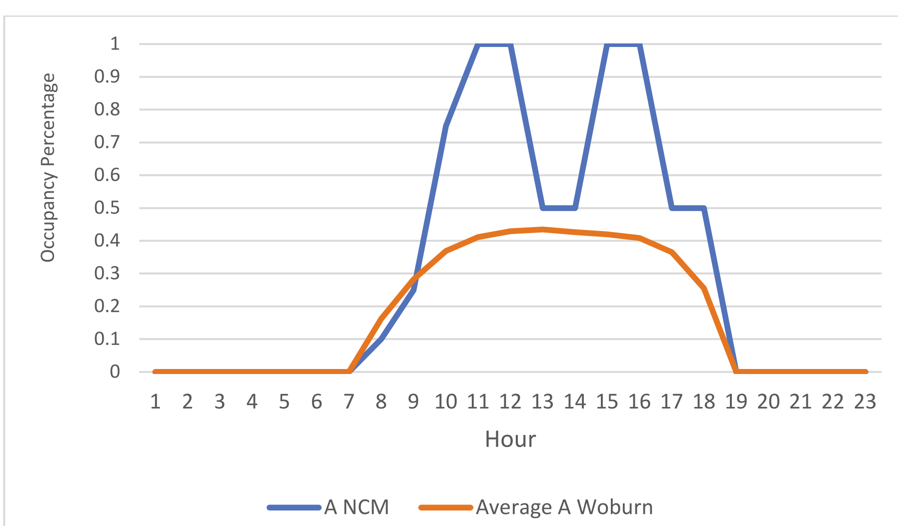

Fig. 14. Averaged EDDOS Administration (A) Occupancy profile vs NCM.

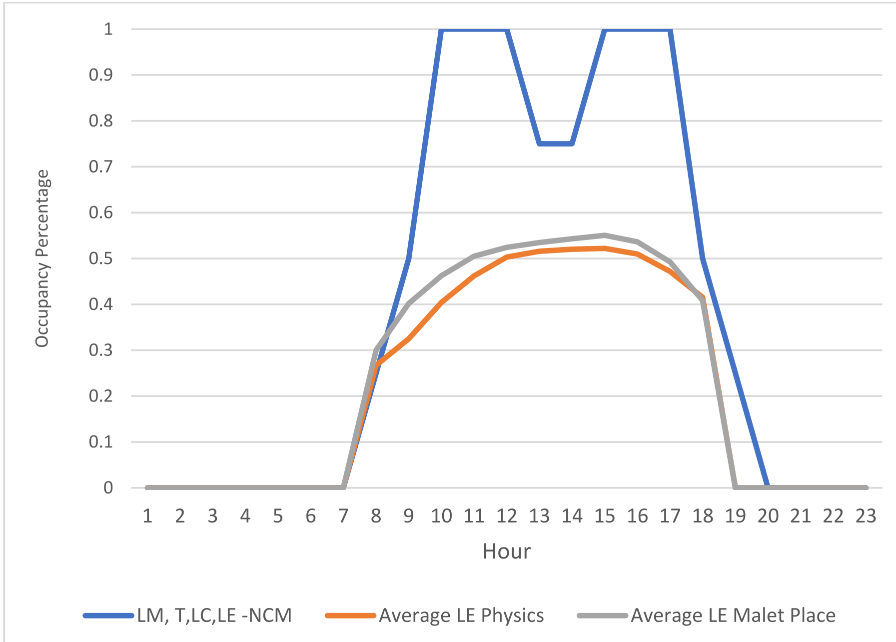

Fig. 15. Averaged EDDOS LE Occupancy profile vs NCM.

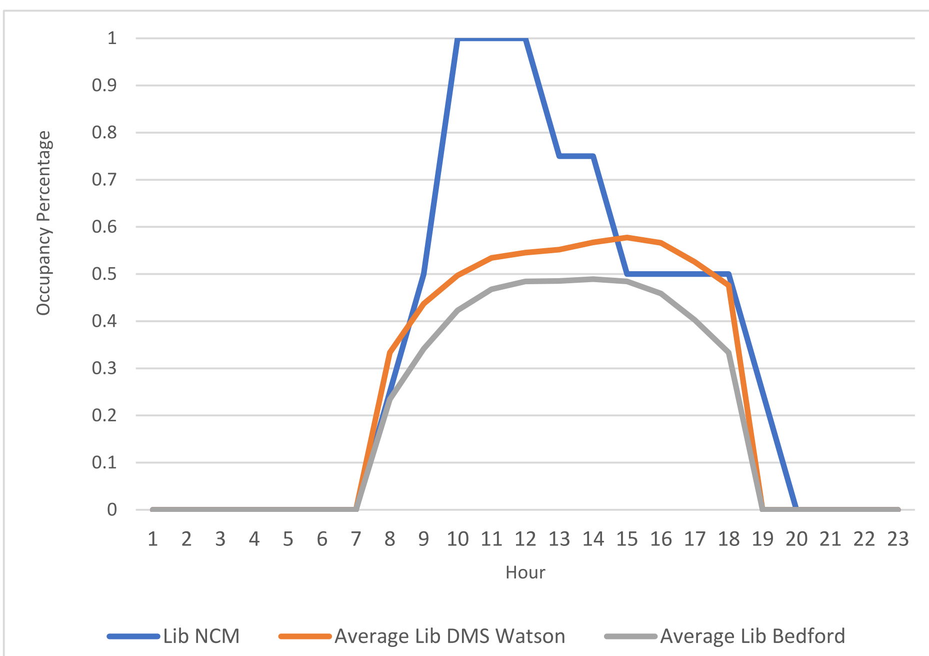

Fig. 16. Averaged EDDOS Library (LIB) Occupancy profile vs NCM.

seen for each occupancy profile except in administration profile (A) where NCM is equal or lower than EDDOS. It also shows many EDDOS schedules when data is available for same occupancy type: EDDOS1, EDDOS2, ……. etc.

4. Discussion

This research provides several deeper insights. One significant insight is the importance of accurately representing occupant behavior in energy models. The novel Energy Data-Driven Occupancy Schedule (EDDOS) method introduced in this study demonstrated the potential to capture building-specific occupancy patterns more realistically compared to standard profiles. This method leverages actual energy use data to generate occupancy schedules, which can significantly enhance the accuracy of energy demand predictions. The findings align with previous work by Tahmasebi and Mahdavi (2017), who emphasized the sensitivity of building performance simulation results to the choice of occupants' presence models [12,52]. Additionally, the study highlights the value of using multiple occupancy patterns to balance accuracy and computational efficiency, a concept supported by Gaetani et al. (2016), who advocated for fit-for-purpose modeling strategies [13,52].

The research also uncovered some challenging findings regarding the resolution of occupant behavior models. One of the primary challenges is the data inadequacy for high-resolution occupancy modeling at the

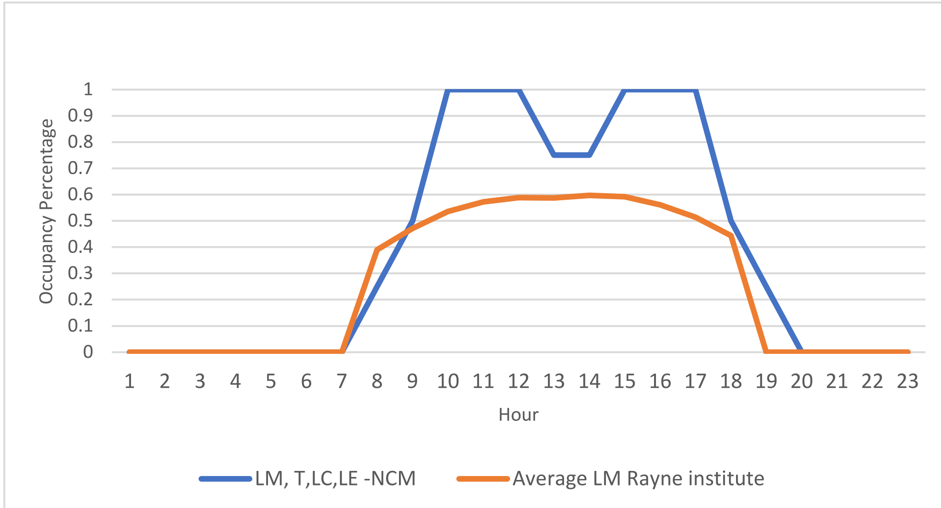

Fig. 17. Average EDDOS Lab Medicine (LM) occupancy profile vs NCM.

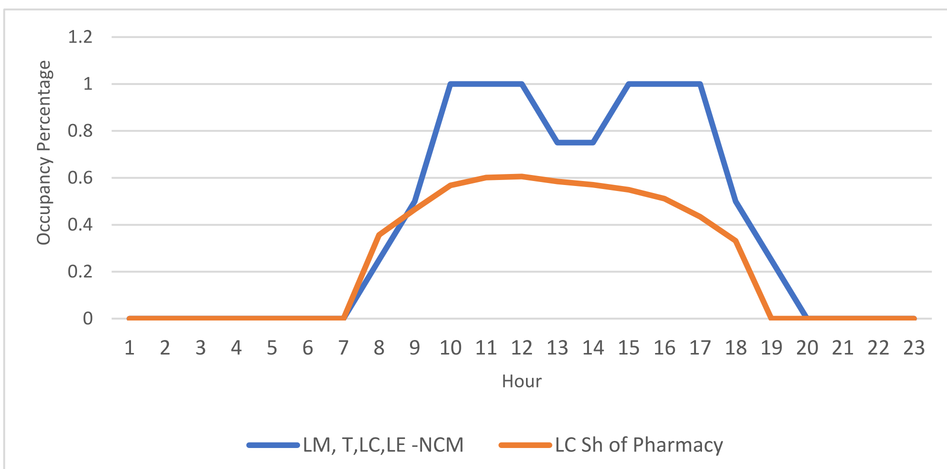

Fig. 18. Average EDDOS Lab Chemistry vs NCM.

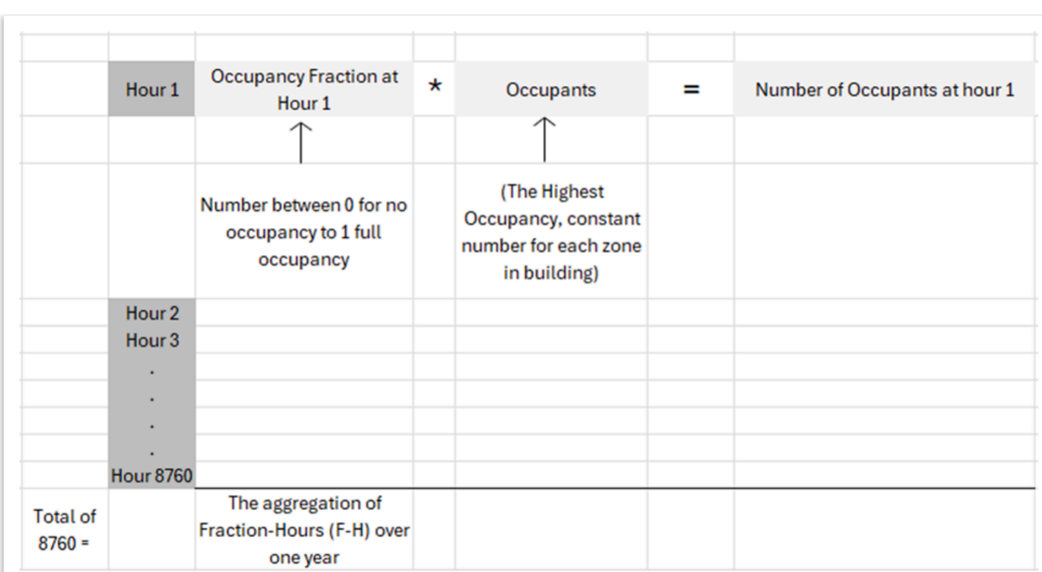

Fig. 19. Fraction-Hours for occupancy over one year.

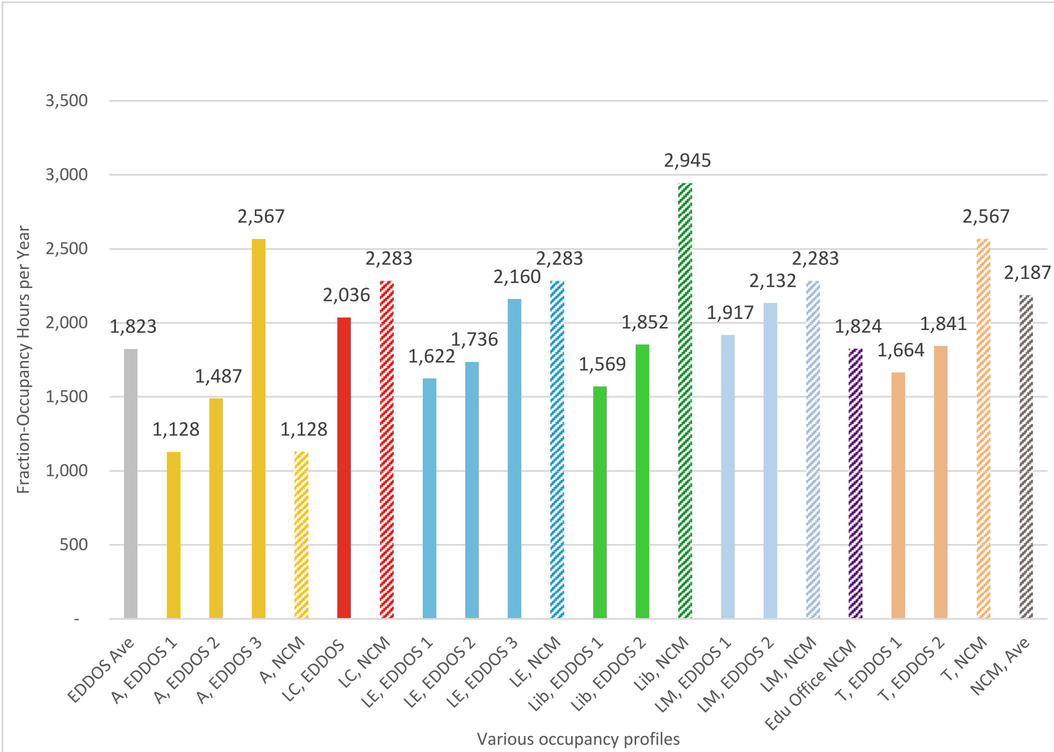

Fig. 20. Annual occupancy Fraction-Hours for various NCM and EDDOS profiles.

urban scale. The lack of comprehensive and high-quality datasets impairs the ability to gather essential occupancy information, leading to substantial errors in estimating building energy consumption. This issue is compounded by the simplifying assumptions often made to depict the complexities of real-world occupant behavior, which can result in mistaken predictions of energy usage and comfort levels. The study's comparison of different modeling scenarios revealed that more complex occupancy approaches generally improved the accuracy of energy demand estimations. However, the marginal improvement in accuracy between the most complex scenario (E) and the diversified scenario (D) raises questions about the trade-off between increased computational complexity and improved accuracy.

The findings also underscore the importance of considering load diversity in district energy system planning. The research demonstrated that diversified peak calculations generally provided closer estimates to measured values, highlighting the need to account for the variability and diversity in occupancy within the same building type. This aligns with the work of Happle et al. (2018), who showed that the use of uniform schedules and fixed set points could lead to an overestimation of peak cooling loads [52].

5. Conclusion

This study examines various approaches to building stock energy simulation and occupancy modelling for district-level heating and cooling energy demand assessment, using nineteen buildings at University College London's Bloomsbury campus. Five scenarios of increasing complexity were evaluated, from a machine learning tool (THERMOS) to building-specific occupancy profiles.

The results show that occupancy modelling significantly influences energy demand estimations, especially for peak loads. For annual heating demand, the most complex scenario with building-specific occupancy profiles provided the closest estimation to metered data (12 % difference), compared to a 44 % underestimation by the simplest scenario. More complex occupancy approaches improved peak heating load predictions, with the diversified scenario showing only a 4 % difference from metered data.

The novel Energy Data-Driven Occupancy Schedule (EDDOS) method showed promise in capturing building-specific occupancy patterns, resulting in lower occupancy fraction-hours per year compared to standard profiles.

Cooling demand estimations varied substantially between scenarios, highlighting the importance of appropriate occupancy modelling. Peak cooling load estimates ranged from 15,069 kW (THERMOS) to 2,514 kW (most complex scenario). The multiple occupancy patterns scenario emerged as a promising compromise, offering improved accuracy over simpler methods without full complexity.

These findings have important implications for district energy system planning and design. The results suggest that using dynamic building stock energy simulation with multiple occupancy patterns could provide a good balance between accuracy and computational efficiency for large-scale assessments.

Study limitations include the lack of measured cooling data for validation and the need for testing across a larger, more diverse building stock. The impact of climate change on future cooling demands underscores the need for robust modelling approaches.

In conclusion, this study emphasizes the importance of carefully considering occupancy modelling complexity in building stock energy simulations for district energy system planning. While the most complex modelling approaches can provide the highest accuracy, the additional computational demands may not always be justified. A balanced approach, such as using multiple occupancy patterns, could offer an optimal solution for many applications in district energy system planning and building stock energy modelling. The authors' analysis focuses solely on a university campus at University College London (UCL) in the UK, which experiences a climate dominated by heating needs. These occupancy patterns may not be typical of configurations found in other areas or under different weather conditions.

Funding

This research was funded by the UK Engineering and Physical Sciences Research Council grant number EP/N509577/1 and EP/T517793/1.

CRediT authorship contribution statement

Salam Al-Saegh: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Vasiliki Kourgiozou:** Software. **Ivan Korolija:** Writing – review & editing, Software. **Rui Tang:** Writing – review & editing. **Farhang Tahmasebi:** Writing – review & editing, Supervision, Software, Funding acquisition. **Dejan Mumovic:** Writing – review & editing, Supervision.

Declaration of competing interest

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

Data availability

The authors do not have permission to share data.

References

[1] P. Nejat, et al., A global review of energy consumption, CO2 emissions and policy in the residential sector (with an overview of the top ten CO2 emitting countries), Renew. Sustain. Energy Rev. 43 (2015) 843–862.

[2] S. Charani Shandiz, B. Rismanchi, G. Foliente, Energy master planning for net-zero emission communities: state of the art and research challenges, Renew. Sustain. Energy Rev. 137 (2021) 110600.

[3] A. Koch, S. Girard, K. McKoen, Towards a neighbourhood scale for low- or zero-carbon building projects, Build. Res. Inf. 40 (4) (2012) 527–537.

[4] Y. Chow, Utilizing district energy system as a cost-effective measure in meeting UK domestic 'zero carbon' targets, International Journal of Low-Carbon Technologies 4 (3) (2009) 169–174.

[5] A. Athienitis, W. O'Brien, Conclusion, research needs, and future directions, in Modeling, Design, and Optimization of Net-Zero Energy Buildings. 2015. p. 351-354.

[6] S. Walker, et al., An assessment methodology of sustainable energy transition scenarios for realizing energy neutral neighborhoods, Appl. Energy 228 (2018) 2346–2360.

[7] N. Nord, D. Schmidt, A.M.D. Kallert, Necessary measures to include more distributed renewable energy sources into district heating system, Energy Procedia 116 (2017) 48–57.

[8] A. Menapace, et al., Stochastic generation of district heat load, Energies (Basel) 14 (17) (2021) 5344.

[9] N. Aoun, et al., Load shifting of space-heating demand in district heating systems based on a reduced-order building model identifiable at substation level, in 4th Smart Energy Systems and 4th Generation District Heating Conference, Aalborg, Denmark, 2018.

[10] A.S. Fathi, W. O'Brien, Considering diverse occupant profiles in building design decisions, Build. Environ. (2024) 111857.

[11] D. Yan, et al., Occupant behavior modeling for building performance simulation: current state and future challenges. Energy and Buildings , 107 pp. 264-278. (2015), 2015.

[12] F. Tahmasebi, A. Mahdavi, The sensitivity of building performance simulation results to the choice of occupants' presence models: a case study, J. Build. Perform. Simul. 10 (5–6) (2017) 625–635.

[13] I. Gaetani, P.-J. Hoes, J.L.M. Hensen, Occupant behavior in building energy simulation: towards a fit-for-purpose modeling strategy, Energ. Buildings 121 (2016) 188–204.

[14] M. Schweiker, Understanding occupants' behaviour for energy efficiency in buildings, Current Sustainable/renewable Energy Reports. 4 (1) (2017) 8–14.

[15] E. Azar, C.C. Menassa, A comprehensive analysis of the impact of occupancy parameters in energy simulation of office buildings, Energ. Buildings 55 (2012) 841–853.

[16] Z. He, T. Hong, S.K. Chou, A framework for estimating the energy-saving potential of occupant behaviour improvement, Appl. Energy (2021).

[17] T. Hong, et al., Advances in research and applications of energy-related occupant behavior in buildings, Energ. Buildings 116 (C) (2016) 694–702.

[18] C. Carpino, et al., Energy performance gap of a nearly Zero Energy Building (nZEB) in Denmark: the influence of occupancy modelling, Build. Res. Inf. 48 (8) (2020) 899–921.

[19] V.M. Barthelmes, C. Becchio, S.P. Corgnati, Occupant behavior lifestyles in a residential nearly zero energy building: effect on energy use and thermal comfort, Sci. Technol. Built Environ. 22 (7) (2016) 960–975.

[20] H.T. Ebuy, et al., Occupant behavior impact on building sustainability performance: a literature review, Sustainability (Basel, Switzerland) 15 (3) (2023) 2440.

[21] G. Pristerà, et al., Archetype selection process for the development of a building stock model, IOP Conf. Ser.: Earth Environ. Sci 1196 (1) (2023) 12013.

[22] Y. Yang, et al., Transforming and validating urban microclimate data with multi-sourced microclimate datasets for building energy modelling at urban scale, Energ. Buildings 295 (2023) 113318.

[23] G. Happle, J.A. Fonseca, A. Schlueter, A review on occupant behavior in urban building energy models, Energ. Buildings 174 (2018) 276–292.

[24] S. Nejadshamsi, et al., Data sources and approaches for building occupancy profiles at the urban scale – a review, Build. Environ. 238 (2023) 110375.

[25] D. Krishnan, S. Kelly, Y. Kim, A Meta-analysis review of occupant behaviour models for assessing demand-side energy consumption, Energies (Basel) 15 (3) (2022) 1219.

[26] A. Mahdavi, F. Tahmasebi, The deployment-dependence of occupancy-related models in building performance simulation, Energ. Buildings 117 (2016) 313–320.

[27] F. Tahmasebi, A. Mahdavi, An inquiry into the reliability of window operation models in building performance simulation, Build. Environ. 105 (2016) 343–357.

[28] F. Johari, F. Shadram, J. Widén, Urban building energy modeling from geo-referenced energy performance certificate data: development, calibration, and validation, Sustain. Cities Soc. 96 (2023) 104664.

[29] H. Hou, et al., Building occupancy modelling at the district level: a combined copula-nested hazard-based approach, Build. Environ. 225 (2022) 109661.

[30] S. Samareh Abolhassani, et al., Improving residential building energy simulations through occupancy data derived from commercial off-the-shelf Wi-Fi sensing technology, Energ. Buildings 272 (2022) 112354.

[31] M. Mosteiro-Romero, et al., A data-driven agent-based model of occupants' thermal comfort behaviors for the planning of district-scale flexible work arrangements, Build. Environ. 257 (2024) 111479.

[32] A. Banfi, M. Ferrando, F. Causone, Occupant behaviour (OB) modelling to support urban building energy simulation, An Overview. (2023).

[33] F.D. Salim, et al., Modelling urban-scale occupant behaviour, mobility, and energy in buildings: a survey, Build. Environ. 183 (2020) 106964.

[34] Y. Chen, T. Hong, X. Luo, An agent-based stochastic occupancy simulator, Build. Simul. 11 (1) (2018) 37–49.

[35] S. Al-Saegh, et al., Comparison of deterministic, stochastic, and energy-data-driven occupancy models for building stock energy simulation, Buildings 14 (9) (2024) 2933.

[36] Z. Liu, et al., Exploring the impacts of heterogeneity and stochasticity in air-conditioning behavior on urban building energy models, Sustain. Cities Soc. 103 (2024) 105285.

[37] G. Chaudhary, et al., Evaluation of "Autotune" calibration against manual calibration of building energy models, Appl. Energy 182 (C) (2016) 115–134.

[38] P. Raftery, M. Keane, A. Costa, Calibrating whole building energy models: detailed case study using hourly measured data, Energ. Buildings 43 (12) (2011) 3666–3679.

[39] D. Coakley, P. Raftery, M. Keane, A review of methods to match building energy simulation models to measured data, Renew. Sustain. Energy Rev. 37 (2014) 123–141.

[40] P. Im, et al., Empirical validation of building energy modeling for multi-zones commercial buildings in cooling season, Appl. Energy 261 (C) (2020) 114374.

[41] Thermos. what is thermos. 2020; Available from: https://www.thermos-project.eu/thermos-tool/what-is-thermos/.

[42] Thermos. Heat Demand Model. Demand 2020; Available from: https://tool.thermos-project.eu/help/demand/demand-model.html.

[43] NCM, UK's National Calculation Method for Non Domestic Buildings. 2021.

[44] Sim-Stock, U. Sim-Stock. Available from: https://github.com/UCL/simstock.

[45] Survey, O. OS data download and API products. 2023; Available from: https://www.ordnancesurvey.co.uk/products.

[46] J. Dong, et al., London School Building Stock Model for Cognitive Performance Assessment, in IBPSA. 21-22 Sep 2020, IBPSA: Loughborough, UK.

[47] BRE, Appendix S: Reduced Data SAP for existing dwellings. 2012.

[48] Ruyssevelt, P., Modelling London's Building Stock and Its Associated Energy Use, in ASHRAE winter Conference. 2019.

[49] CIBSE_TM63, CIBSE TM 63. 2020.

[50] K. Vasiliki, et al., Development of a dynamic building stock model for smart energy transition decision support: university campus stock case study, in Proceedings of Building Simulation 2023: 18th Conference of IBPSA. 2023, IBPSA. p. 3764-3772.

[51] S. Werner, European space cooling demands, Energy 110 (2016) 148–156.

[52] J. Li, et al., Development and improvement of occupant behavior models towards realistic building performance simulation: a review, Sustain. Cities Soc. 50 (2019) 101685.