buildings

MDPI

Article

Comparison of Deterministic, Stochastic, and Energy-Data-Driven Occupancy Models for Building Stock Energy Simulation

Salam Al-Saegh, Farhang Tahmasebi, Rui Tang * and Dejan Mumovic

Institute for Environmental Design and Engineering, University College London, London WC1E 6BT, UK; s.saegh@ucl.ac.uk (S.A.-S.); f.tahmasebi@ucl.ac.uk (F.T.); d.mumovic@ucl.ac.uk (D.M.)
* Correspondence: rui.tang@ucl.ac.uk

Abstract: Accurate modelling of occupancy patterns is critical for reliable estimation of building stock energy demand, which is a key input for the design of district energy systems. Aiming to investigate the suitability of different occupancy-modelling approaches for the design of district energy systems, the present study examines a set of standard-based schedules (from the UK National Calculation Methodology), a widely used stochastic occupancy model, and a novel energy-data-driven occupancy model. To this end, a dynamic energy model of a higher education office building developed within a stock model of London's Bloomsbury district serves as a testbed to implement the occupancy models, explore their implications for the estimation of annual and peak heating and cooling demand, and extrapolate the findings to the computationally demanding building stock stimulations. Furthermore, the simulations were conducted in two years before and after the COVID-19 pandemic to examine the implications of hybrid working patterns after the pandemic. From the results, the energy-data-driven model demonstrated superior performance in annual heating demand estimations, with errors of ±2.5% compared to 14% and 7% for the standard-based and stochastic models. For peak heating demand, the models performed rather similarly, with the data-driven model showing 28% error compared to 29.5% for both the standard-based and stochastic models in 2019. In cooling demand estimations, the data-driven model yielded noticeably higher annual cooling demand and lower peak cooling demand estimations as compared with the standard-based and stochastic occupancy models. Given the adopted building-modelling approach, these findings can be extended to district-level investigations and inform the decision on the choice of occupancy models for building stock energy simulation.

Keywords: occupancy modeling; building energy simulation; district energy systems; energy data-driven methods; building stock modeling

check for updates

Citation: Al-Saegh, S.; Tahmasebi, F.; Tang, R.; Mumovic, D. Comparison of Deterministic, Stochastic, and Energy-Data-Driven Occupancy Models for Building Stock Energy Simulation. *Buildings* **2024**, *14*, 2933. https://doi.org/10.3390/buildings14092933

Academic Editors: Laura Gabrielli and Ruggeri Aurora Greta

Received: 28 July 2024
Revised: 5 September 2024
Accepted: 14 September 2024
Published: 17 September 2024

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 energy flexibility at the building level to allow for a better interaction between buildings' and neighbourhoods' energy systems and reach the best design configurations for a cluster of buildings and neighbourhoods. Walker at 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.

In this setting, utilisation of District Energy Systems (DESs) is an efficient way of decarbonising the heating and cooling provided to buildings [7], which also secures 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 be used otherwise, and their effectiveness and carbon-saving potential grows as they expand and connect to each other.

2. Literature Review

Previous research suggests that full realisation of energy networks' potentials 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 the changing climatic conditions [8,9]. Specifically, accurate assessment of building stock base and peak thermal loads is critical for optimum sizing of district energy networks [10]. Oversizing increases initial and running costs and reduces efficiency, and undersizing 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 [11]. This latter point, in particular, means that the accuracy and quality of heat demand models of individual buildings is crucial for the overall accuracy of DES models [12].

In this context, occupant behaviour is one of the most critical variables in zero-carbon design, especially due to the improvement of buildings' physical characteristics and energy systems [13–16]. Many studies have demonstrated that occupants play a significant role in shaping energy demand in buildings [17–20], 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, [21–23]).

However, whereas building stock modelling has relatively matured in terms of using of dynamic thermal simulation tools and relying on relatively 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 [24]. In particular, the paucity of high-resolution data on occupants' energy-related behaviour at an urban scale leads to major difficulties in accurate and granular estimation of buildings' energy consumption [25], leading to a more challenging performance gap at the urban scale [26].

Specifically, previous studies demonstrate three approaches for occupancy modelling in district-level energy simulations. The first approach relies on schedules of occupants' presence, use of light and equipment, and hot water. This approach, which is adopted by building performance standards such as ASHRAE standard 90.1 and the UK National Calculation Method (NCM) [27], has been traditionally used in building energy simulation studies for years. The use of fixed occupancy schedules in urban-scale energy modelling presents both advantages and disadvantages. On the plus side, they offer a simple, straightforward approach to modelling, which can be beneficial for large-scale urban energy assessments where detailed data may not be available. Also, the consistency of profiles allows for easier comparison across different buildings and scenarios [28]. In addition, these schedules are derived from established guidelines and are easy to implement in simulation tools, making them accessible for practitioners who may not have the resources to develop complex, data-driven models [29]. Standard schedules also provide a baseline for policymakers to evaluate energy performance and develop energy conservation measures at an urban scale, especially when detailed occupancy data are unavailable [29]. On the other hand, as the standard-based occupancy schedules do not reflect the actual occupancy patterns, their use can lead to significant discrepancies between predicted and actual energy consumption. This is particularly problematic in non-residential buildings where occupancy patterns can vary widely from one building to another [30–32].

Furthermore, while many studies in this area have primarily relied on standard peak values and schedules for occupancy across their sample size [33], this approach tends to overlook the large diversity in the operation of buildings by occupants [34], which can result in an overestimation of occupancy loads and directly impact the estimation of energy demands for district energy networks. For example, in their study, Happle et al. [26] demonstrated that the use of uniform schedules and fixed set points leads to an overestimation of peak cooling loads.

The second approach to modelling occupants in building stock modelling is based on the integration of computationally advanced stochastic models of occupant presence and behaviour [35]. These models are meant to, and could potentially, enhance the representation of occupants in building performance models. But, as suggested by other studies [36,37], the reliability of these models and their cost benefit, especially for the purpose of stock modelling, are rather debatable (and this will be further investigated in this study). In particular, the integration of these computationally expensive models into building stock energy models (which by their nature are computationally demanding themselves) is very challenging. This has led many researchers (including the authors) to rely on single runs of stochastic models (albeit across sub-hourly timesteps of annual simulations), which could to a large extent undermine the benefits of stochastic models to yield probable ranges of building performance indicators (rather than single values of these indicators).

The third approach to modelling occupancy in buildings relies on leveraging data analytics to predict and optimise building occupancy patterns, which can significantly enhance energy efficiency and operational effectiveness [38–40]. This approach utilises various data sources, such as plug load profiles and sensor data, to infer occupancy states and adjust building systems accordingly. For example, Vassiljeva et al. [39] developed an algorithm using energy consumption data to create occupancy profiles for optimising ventilation schedules in school buildings. This method relied on energy consumption data from air handling units to either predict the presence of occupants or estimate their number [39]. In another effort, Vosoughkhosravi et al. [41] utilised machine learning algorithms to detect occupancy in office spaces based on individual plug load data. The models, validated with seat pressure sensors, achieved high accuracy in real-time occupancy detection, demonstrating a cost-effective alternative to traditional sensor-based systems. While data-driven occupancy modelling offers substantial benefits in energy efficiency and operational optimisation, challenges such as data privacy, model generalisability, and the complexity of human–building interactions remain. Addressing these issues requires further research and development to refine these models and ensure their applicability across diverse building types and environments [42]. Emerging data-driven approaches, such as those using machine learning and clustering algorithms, offer more nuanced insights by capturing the variability in occupant behaviour and energy use patterns [29,30]. These methods can enhance the accuracy of energy models, although they require more detailed data and computational resources which may not be always available.

In a wider perspective, recent advancements in data technologies have provided opportunities for a deeper understanding of building occupant patterns [43]. 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 [44]. However, despite these developments, there is still a need for more comprehensive and high-quality datasets [45], along with development and examination of fit-for-purpose occupancy-modelling approaches to improve the accuracy of building energy modelling at the urban scale, which is of critical importance for demand-side modelling in the design and assessment of DESs.

In conclusion, existing literature emphasises the importance of decarbonising the building sector to combat climate change and highlights the potential of community-based solutions, particularly district energy systems (DESs). However, accurate modelling of these systems, specifically with regard to representation of occupant behaviour, remains a challenge, leading to discrepancies between predicted and actual energy consumption.

3. Research Aims and Scope

Within the context described in the introduction and literature review, this research aims to evaluate the suitability of different occupancy-modelling approaches for building stock energy modelling and assess their impact on building energy demand estimations as needed for the optimal design of district energy systems.

The current study makes use of a building stock model developed for UCL Bloomsbury campus in London, England [46], but focuses on one of the buildings from the campus to implement and test different computationally demanding occupant behaviour modelling approaches. Specifically, we populate the building model with three types of occupancy models (namely, deterministic standard-based schedules, a stochastic model of occupants' presence, and a novel energy-data-driven occupancy model) for estimation of annual and peak heating and cooling demands as key performance indicators considered in the design of DESs. The study examines these indicators for two years, namely 2019 and 2023, to capture pre- and post-pandemic occupancy patterns.

It should be noted that, while the scope of this study is limited to one building, as the building model is developed as a part of a district model (with typical characteristics of these models as detailed in Section 4.3), the findings of the study can directly inform modelling of occupants for the purpose of building stock energy modelling. As such, the studied building serves as a small-scale testbed to identify the most cost-effective occupancy models to be integrated into a much more computationally expensive building stock energy modelling campaign.

4. Method

4.1. Overview

To describe the research method systematically, in the following, we first discuss the studied building (Section 4.2), and then describe how the energy model of this building is developed (Section 4.3). Subsequently, Section 4.4 details the three occupancy-modelling approaches selected for this study. These are, namely, a standard-based approach using fixed schedules from UK NCM, a stochastic model selected from the literature generating random daily occupancy profiles, and a novel energy-data-driven model (EDDOS). Lastly, Section 4.5 describes the building performance indicators used to evaluate the studied occupant behaviour models in assessment of district energy systems. Figure 1 illustrates how these different components of the research make it possible to assess the suitability of different occupancy models for building stock energy simulations.

4.2. Case Study Building

As clarified in Section 3, this study focuses on one higher education office building from the UCL Bloomsbury campus in central London. The building, called UCL Central House, is a six-storey building, which mainly accommodates offices and meeting rooms (for academic staff and PhD students) along with the Bartlett Library on its ground floor. The building was selected as a case study since it was representative of one type of (office-dominated) Bloomsbury campus building and its metered gas and electricity use was collected for a number of years. Figure 2 illustrates the building's typical floor plan.

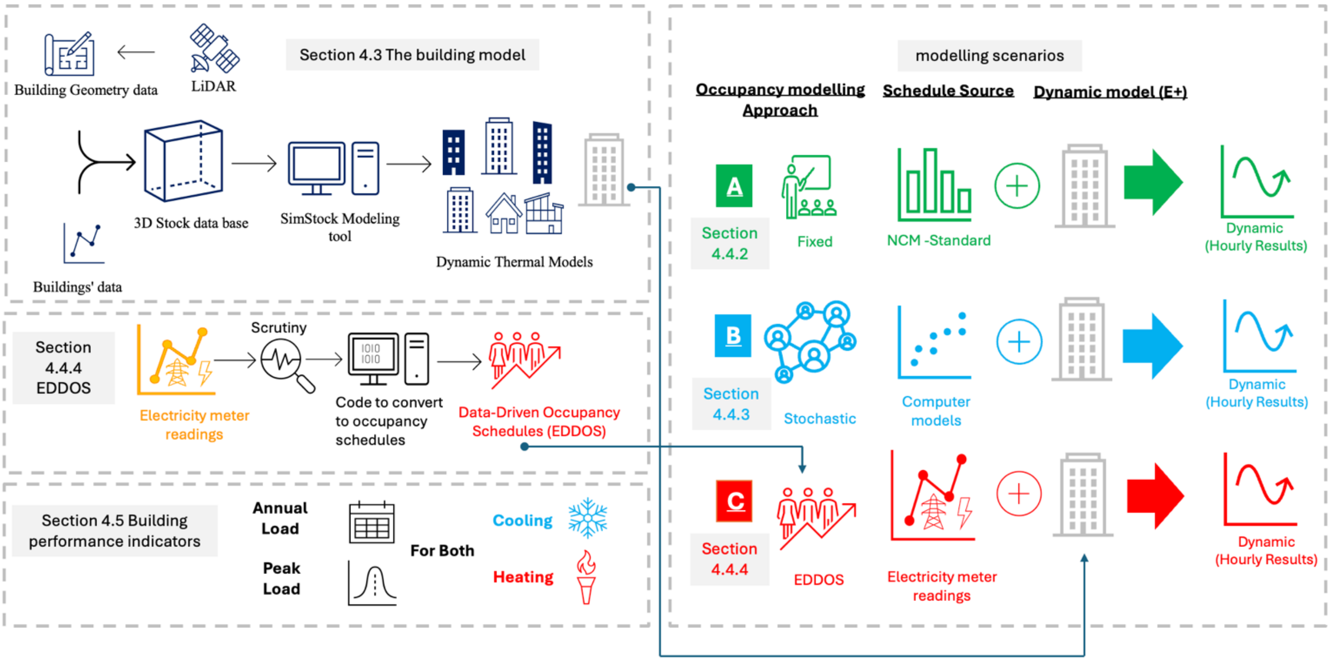

Figure 1. Graphical overview of research method (green: standard-based fixed schedules, blue: stochastic occupancy modelling, red: energy data-driven approach (EDDOS)).

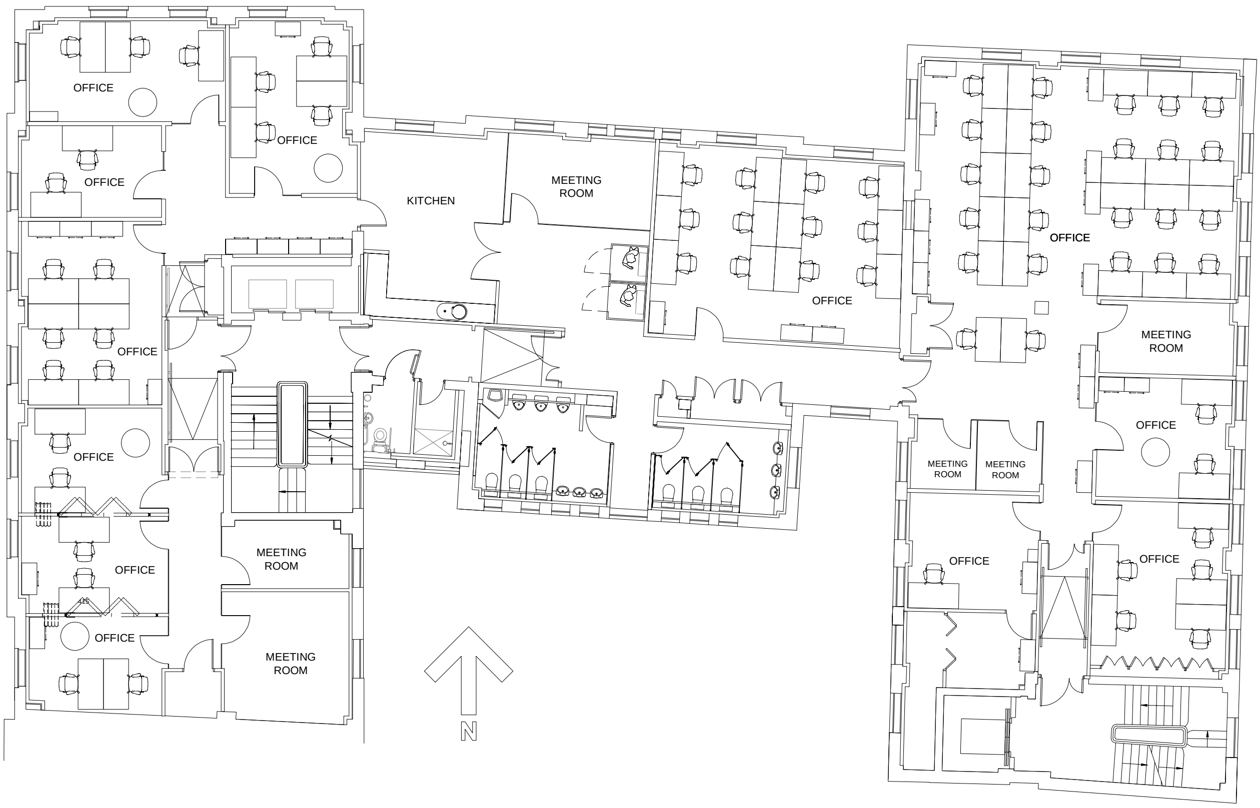

Figure 2. Case study building typical floor plan.

4.3. The Building Model

The building model used in this study is one of 120 models developed for the UCL Bloomsbury campus stock model, the detailed process of which is explained by Kourgiozou et al. [46]. These are building energy models generated using the extensively validated dynamic building energy modelling tool EnergyPlus in two main steps:

1. Developing a 3D building stock database integrating key building-level data including building name, geometry (including footprint and height), construction, and activity type, based on data sources such as UCL Estates, Display Energy Certificates, Ordnance Survey, and Valuation Office Agency (as detailed in [46]).
2. Using a Python-coded tool called SimStock [47] to automatically generate EnergyPlus thermal models for each building by extracting data from the 3D building stock database.

The generated building energy models maintain the actual building forms (instead of using geometric archetypes) and are comprised of one thermal zone per floor. They also represent glazing ratios across different facades as derived from a database developed in a previous study [48]. Four sets of operation schedules (involving occupancy, light, equipment, and hot water use) are assigned to the ground floor zone and the thermal zones representing upper floors as per data obtained from the UK National Calculation Methodology [27]. For the purpose of the current study on UCL Central House building, an EnergyPlus ideal loads air system is defined in the building model, which makes it possible to estimate building heating and cooling energy demands without a detailed definition of the HVAC system.

This bottom-up building stock modelling approach offers a balance between efficiency and accuracy and captures the dynamic relationship between the demand and supply sides as needed for design and assessment of DESs. Specifically, the dynamic nature of the underlying energy simulation engine allows for estimation of sub-hourly energy demands per building and provides a testbed for a feasibility study of integrating new technologies and renewables under future climate change scenarios [49]. Furthermore, this stock model can be populated with sub-hourly patterns of occupants' presence and behaviour, which makes it particularly useful for the purpose of this study. Conversely, the low granularity of thermal zones (one zone per floor) adds some inaccuracy to the estimation of building energy demands and limits the possibilities of integrating high-resolution occupant models (such as agent-based occupancy models, or models of occupant behaviour which capture interactions with the environmental control systems of dedicated offices).

4.4. Occupancy Modelling

4.4.1. Studied Approaches

Three occupancy-modelling scenarios are considered in this study, namely standard-based occupancy schedules, non-repeating occupancy schedules based on a stochastic occupancy model, and energy-data-driven occupancy schedules (EDDOSs). These are explained in more the detail in the following.

4.4.2. Fixed Standard-Based Occupancy Schedules

In this deterministic approach to occupancy representation, the UK National Calculation Method (NCM) [27] 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 the underlying dynamic building simulation methods and standard databases for building construction and activities. Table 1 gives the occupancy-related assumptions considered in this scenario and Figure 3 illustrates the UK NCM weekday schedules of occupancy, use of lights, and equipment used in this study.

Table 1. Occupancy-related assumptions under the standard-based scenario.

Model Input Parameter	Value
People maximum density	0.103 [persons/m^2]
Occupancy schedule	UK NCM D1_Edu_CellOff_Occ
Metabolic rate	123 [W/persons]
Latent heat fraction	0.4
Lighting maximum density	4.6 [W/m^2]
Lighting schedule	UK NCM D1_Edu_CellOff_Light
Equipment maximum density	11.99 [W/m^2]
Equipment schedule	UK NCM D1_Edu_CellOff_Equip
Hot water Supply	0.2369 (l/day/m^2)

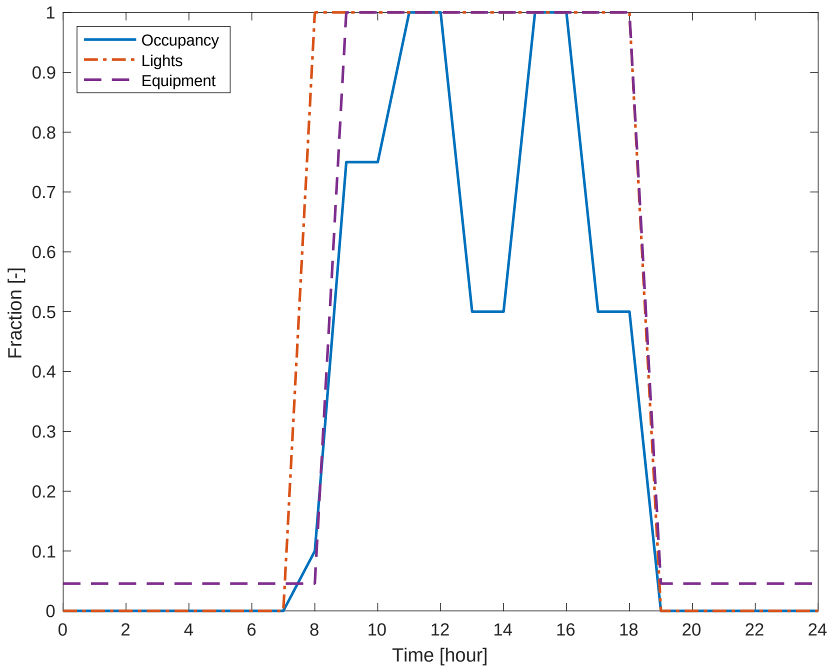

Figure 3. Weekday schedules of occupancy, lights, and equipment used in the standard-based scenario.

4.4.3. Non-Repeating Occupancy Profiles from a Stochastic Model

In this occupancy-modelling scenario, a stochastic occupancy model developed by Page et al. [50] was used to generate random daily occupancy profiles for the annual building performance simulation. The same occupancy schedule from the standard-based scenario was fed as input to this model, which along with a parameter of mobility (defined as the ratio of state change probability to state persistence probability) allowed the model to return non-repeating daily profiles of occupancy states (present or not present). The model was set to start from a state of vacancy for the first timestep of the day for all occupants to then determine the state of occupancy of each occupant at each time step based on the previous occupancy state and the probability of transition from this state to either the same state or its opposite state. To this end, for each time step, a random number between 0 and 1 was generated and compared with the transition probabilities to determine if a change of occupancy state occurred. Further details on this occupancy model can be found in [50].

The stochastic occupancy model was executed for 240 weekdays and 125 weekends and public holidays for each occupant to obtain random daily presence profiles originating from the UK NCM occupancy schedule. Figure 4 shows the zone-level weekday occupancy schedules resulting from averaging the randomly generated daily occupancy profiles of 84 occupants in one of the building floors. The resulting zone-level occupancy schedules were incorporated into the building model and were referenced in the definition of occupants in each thermal zone. The weekends and public holidays were considered in this process such that the days of the week were consistent in models with fixed standard-based and random occupancy profiles.

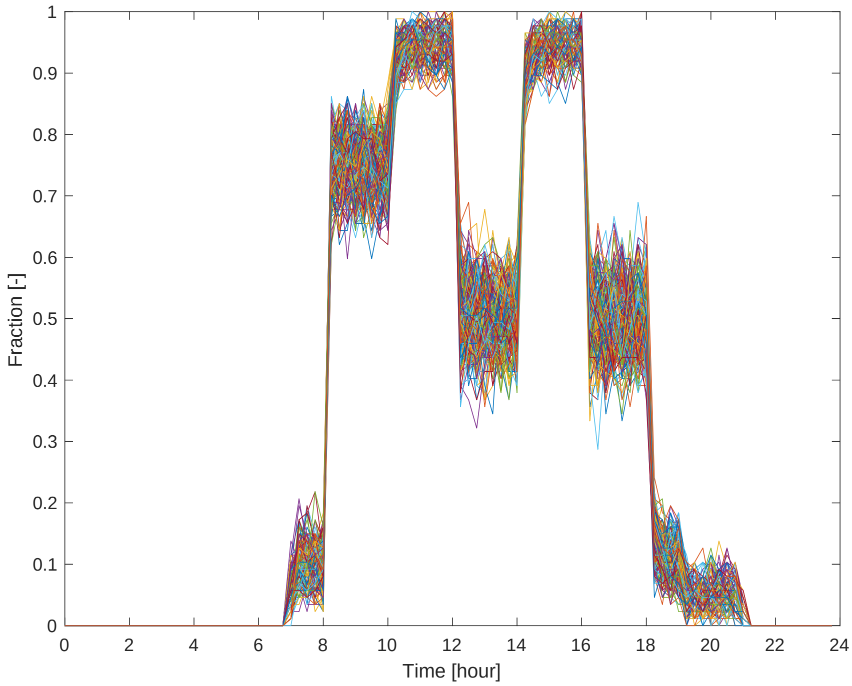

Figure 4. Non-repeating weekday occupancy schedules resulting from averaging the randomly generated presence profiles of 84 occupants in one of the building floors.

4.4.4. Energy-Data-Driven Occupancy Schedules (EDDOSs)

Aiming to model occupancy patterns more realistically, without relying on a stochastic modelling process, a novel procedure was developed and tested to derive occupancy schedules based on energy-use data. First, it was examined to what extent energy-use data correlated with the hours of the day. While metered gas data did not suggest a correlation with the hours of the day, electricity use and the hours of the day showed a good correlation (with an R-squared of 0.84) in the studied building (Figure 5). On this basis and following the periodic pattern of weekly electricity use in the building (see Figure 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 (Figure 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 were processed separately. Thus, the procedure returned non-repeating daily profiles of occupancy corresponding to the variations in electricity use. For the purpose of this study, year 2019 electricity-use data were used to identify the occupancy profiles, as this represented a building usage pattern before the COVID-19 pandemic.

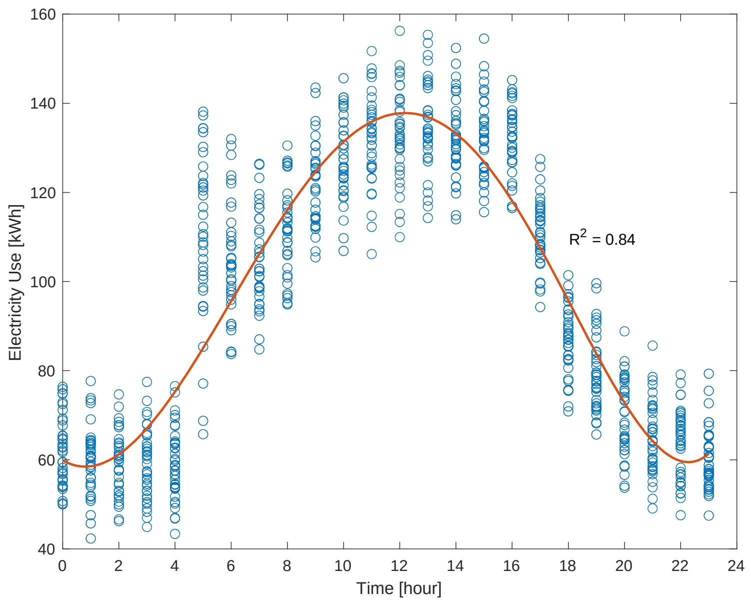

Figure 5. Electrical consumption versus hours of the day for the year 2019. Each circle represents metered electricity used at a certain hour.

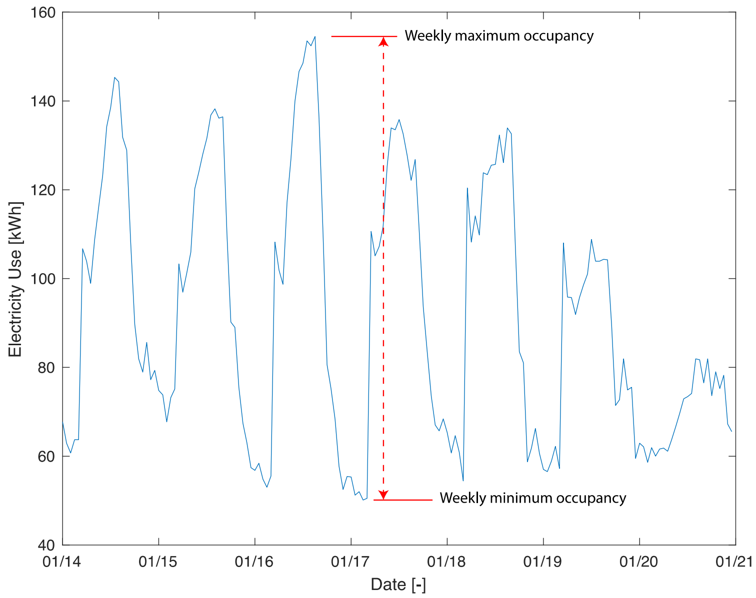

Figure 6. An example weekly profile of electricity consumption and assignment of minimum and maximum occupancy based on minimum and maximum electricity use.

4.5. Building Performance Indicators and Evaluation of Occupancy Models

The study focused on annual and peak heating and cooling demands as key building performance indicators for the design of district energy networks [51–55]. These indicators represent the annual total and hourly peak amount of heating and cooling energy that must be delivered to the building to maintain the heating and cooling set-points. In detail, these indicators can be defined as follows:

- Annual Heating Demand (*AHD*): This represents the total amount of heating energy required over a year to maintain the heating set-points:

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

where $Q_{heating}(t)$ is the heating energy demand at hour t, and 8760 is the total number of hours in a year.

- Peak Heating Demand: This represents the maximum heating energy required to maintain the heating set-points at a certain hour over a year.
- Annual Cooling Demand (*ACD*): This represents the total amount of cooling energy required over a year to maintain the cooling set-points:

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

where $Q_{cooling}(t)$ is the cooling energy demand at hour t, and 8760 is the total number of hours in a year.

- Peak Cooling Demand: This represents the maximum cooling energy required to maintain the cooling set-points at a certain hour over a year.

The estimated heating demand values were then divided by an efficiency factor of 0.7 assumed for the building heating system and compared with the building metered natural gas to assess the accuracy of building model estimations under different occupancy-modelling scenarios. In terms of cooling, however, as the building is not served by 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).

Furthermore, to test the occupancy models more extensively, the building model was simulated for two years, namely 2019 and 2023, as the former represents an occupancy pattern before the COVID-19 pandemic, and the latter represents an occupancy pattern once the post-pandemic hybrid way of working was well established. This also allowed the study to test the performance of the energy-data-driven model based on the year used for its training (2019) and a validation year (2023) with a different occupancy pattern. To this end, for both 2019 and 2023, the EDDOS maximum occupancy fraction (as specified in Figure 6) was set to 1.0, based on a systematic test of the model with different values of 0.4, 0.6, 0.8, and 1.0 for this parameter of the model.

5. Results and Discussions

5.1. Occupancy Profiles

Figure 7 illustrates the daily occupancy profiles resulting from the different occupancy-modelling methods investigated in this study: a standard-based fixed schedule used for all simulation weekdays (UK NCM) and an aggregated occupancy profile resulting from averaging all weekday profiles generated by the stochastic occupancy model (Stochastic Annual Average), along with a randomly selected profile from EDDOS (EDDOS Random Day) and its annual average profile (EDDOS Annual Average).

As can be seen in Figure 7, the UK NCM is the only occupancy profile that maintains 100% occupancy for a rather long period (11 am to 12 pm and 3 pm to 4 pm). The aggregated profile from the stochastic model has slightly reduced occupancy levels in these periods. However, it follows the UK NCM pattern in showing a clearly reduced occupancy level during the two-hour 'lunch break'. Moving on to the EDDOS profiles, which are generated based on electricity-use data, the EDDOS Random Day profile does not reach 100% occupancy and does not show a clear reduction during the lunch break. Most notably, however, the EDDOS Annual Average has clearly flattened midday occupancy peaks. This, firstly, suggests that EDDOS has generated daily profiles with more substantial variations in occupancy as compared with the stochastic model profiles illustrated in Figure 4. Secondly, the energy-data-driven profiles point to the rather negligible impact of the lunch break in reducing occupancy density in this building. This corresponds with the authors' observation in the building that most of the occupants have their lunch inside the building (and in the same level of the building) or leave the building for a very short period for lunch.

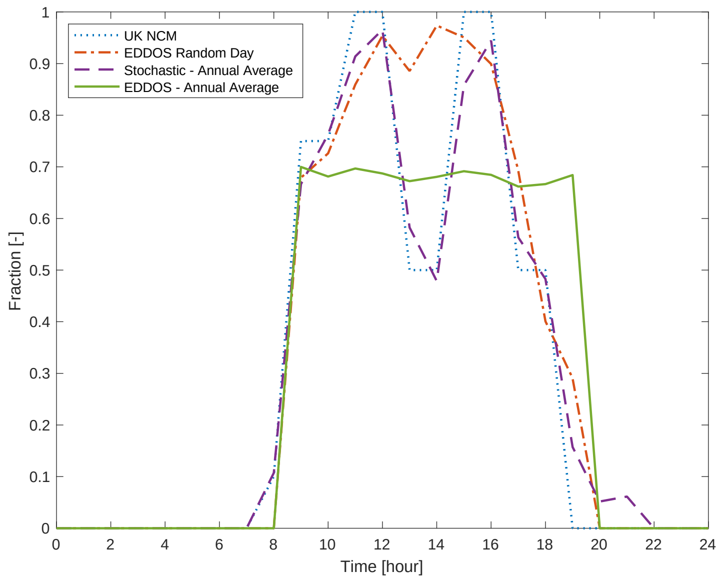

Figure 7. Weekday occupancy profile from different occupancy-modelling scenarios.

It should be noted that, to allow for a valid comparison of energy-demand estimations, all occupancy models are implemented in such a way that they return the same number of occupancy hours in total (namely 1824 h per occupant representing an average of around 35 working hours per week) for an annual simulation as per the UK NCM.

5.2. Annual Heating Demand

Table 2 and Figure 8 give the annual heating demand values as estimated by the building model populated with the standard-based occupancy schedules (NCM), the stochastic occupancy model, and the energy-data-driven model (EDDOS) in comparison with metered heating energy data in the years 2019 (pre-COVID occupancy pattern) and 2023 (post-COVID occupancy pattern). The metered data represent the actual use of natural gas in the building, which provides a basis for comparing the accuracy of energy demand estimations (after applying a building heating system efficiency factor of 0.7 to the estimated heating energy demands). The negative error values in Table 2 show an underestimation of the heating energy demand, while the positive error values depict an overestimation.

Table 2. Obtained annual heating demand values as estimated by the building model populated with standard-based (UK NCM) occupancy schedules, stochastic occupancy model, and energy-data-driven occupancy schedules (EDDOSs), along with the estimation errors as compared with metered heating energy in years 2019 and 2023.

Year	Indicators	Standard-Based Occupancy Schedule (UK NCM)	Stochastic Occupancy Model	Energy-Data-Driven Occupancy Schedules (EDDOSs)	Metered
2019	Annual Heating Demand [kWh/m^2]	32.34	32.41	36.73	37.66
	Estimation Error [%]	−14.1	−13.9	−2.5	-
2023	Annual Heating Demand [kWh/m^2]	25.50	25.57	28.32	27.62
	Estimation Error [%]	−7.7	−7.4	2.5	-

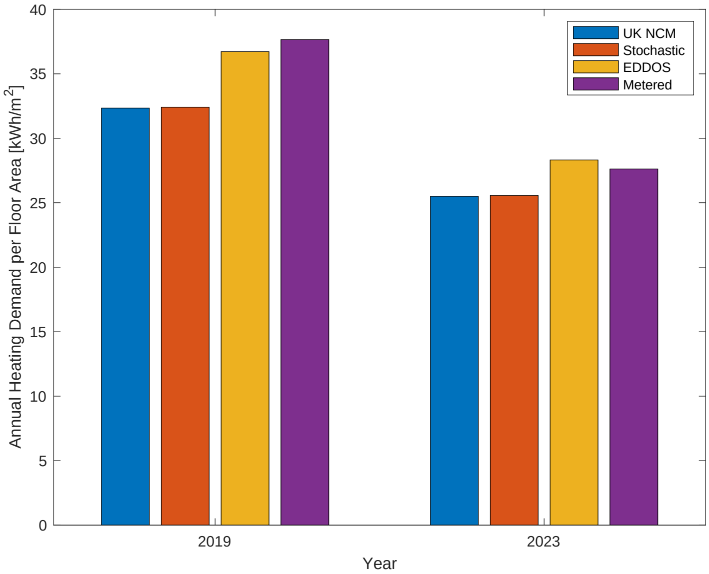

Figure 8. Annual heating demand as estimated by the building model populated with standard-based (UK NCM) occupancy schedules, stochastic occupancy model, and energy-data-driven occupancy schedules (EDDOSs), compared with metered heating energy in years 2019 and 2023.

The results, firstly, suggest that the stochastic occupancy model, despite its much higher computational cost, has not resulted in any significant improvement in the estimation of annual heating demand in both studied years. This finding aligns with what a number of previous studies have suggested (see, for example, [10]). Secondly, the building model equipped with energy-data-driven occupancy schedules (EDDOSs) demonstrates noticeably smaller errors in the estimation of annual heating demand compared to NCM and the stochastic model ($\pm$2.5% error compared with around 14% and 7% from NCM and the stochastic model in the years 2019 and 2023, respectively).

Notably, there is also a difference between 2019 and 2023 in terms of metered heating energy use (36.73 kWh/m^2 for 2019 and 28.32 kWh/m^2 for 2023), which, apart from potentially different occupancy and operation patterns, could also be attributed to weather variations. Based on the hourly weather data files used for this study obtained from Shinny weather data [56], Heating Degree Days (HDDs) for the years 2019 and 2023 are, respectively, 2491 and 2308.

5.3. Peak Heating Demand

Table 3 and Figure 9 show the hourly peak heating demand values estimated by the building models using the standard-based occupancy schedules (NCM), the stochastic occupancy model, and the energy-data-driven occupancy schedules (EDDOSs) in comparison with metered heating energy data in the years 2019 and 2023.

In contrast to annual heating demand results, the three models have performed rather similarly in estimating peak heating demand, although the energy-data-driven model (EDDOS) still shows a better performance. For 2019, EDDOS shows an error of 28% compared with that of 29.5% given by the stochastic model and NCM, and for 2023, EDDOS gives a very small error of 1.1% compared to that of 4.4% yielded by the stochastic model and NCM. It should be noted that the results illustrated here are absolute peak values (and not high percentiles of heating demands commonly used for system sizing). Prediction of absolute peak values in a year, with all the transient boundary conditions influencing the actual building energy use, poses a great challenge to any building energy model from a probability standpoint, and thus the different levels of errors observed in the studied years are not surprising. However, these discrepancies highlight the importance of refining predictive models to account for such variations in peak demand more accurately.

Table 3. Obtained hourly peak heating demand values as estimated by the building model populated with standard-based (UK NCM) occupancy schedules, stochastic occupancy model, and energy-data-driven occupancy schedules (EDDOSs), along with the estimation errors as compared with metered heating energy in years 2019 and 2023.

Year	Indicators	Standard-Based Occupancy Schedule (UK NCM)	Stochastic Occupancy Model	Energy-Data-Driven Occupancy Schedules (EDDOSs)	Metered
2019	Hourly Peak Heating Demand [kW]	117.76	117.76	116.38	90.93
	Estimation Error [%]	+29.5	+29.5	+28.0	-
2023	Hourly Peak Heating Demand [kW]	106.04	106.03	102.51	101.37
	Estimation Error [%]	+4.6	+4.6	+1.1	-

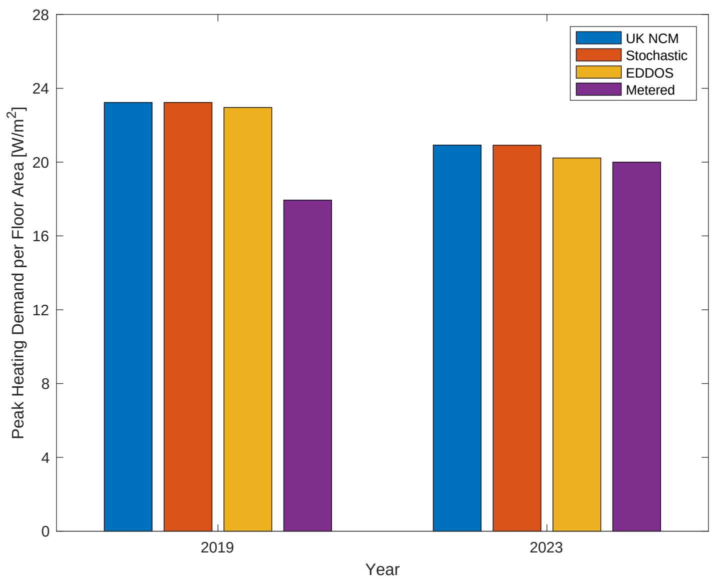

Figure 9. Hourly peak heating demand as estimated by the building model populated with standard-based (UK NCM) occupancy schedules, stochastic occupancy model, and energy-data-driven occupancy schedules (EDDOSs), compared with metered heating energy in years 2019 and 2023.

5.4. Annual and Peak Cooling Demands

Figures 10 and 11 present the obtained annual and peak cooling energy demands, respectively, as estimated using the three occupancy-modelling scenarios in the years 2019 and 2023. While the studied building does not have a dedicated electricity meter for the central cooling system to allow for assessing the accuracy of these estimations, comparison of the cooling demand values provides insights into the implications of using the different occupancy models. It should be noted that the building thermal model used in this study is only verified and validated against actual heating-energy-use data. However, as previous studies have also suggested [57–60], this process ensures that the model correctly captures the building's thermal dynamics, which are relevant to both heating and cooling. While this method cannot completely replace direct validation based on cooling data, it provides a solid framework for the comparative assessment of the building model's cooling demand estimations, as populated with different occupancy models.

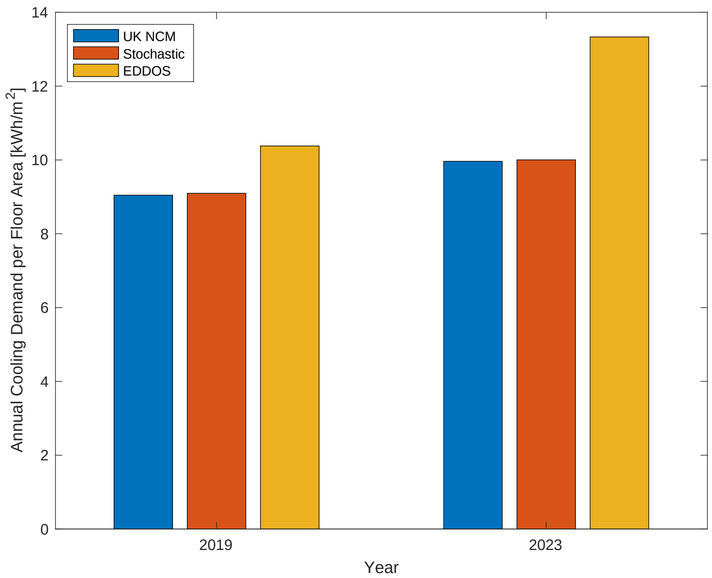

Figure 10. Annual cooling demand as estimated by the building model populated with standard-based (UK NCM) occupancy schedules, stochastic occupancy model, and energy-data-driven occupancy schedules (EDDOSs) in years 2019 and 2023.

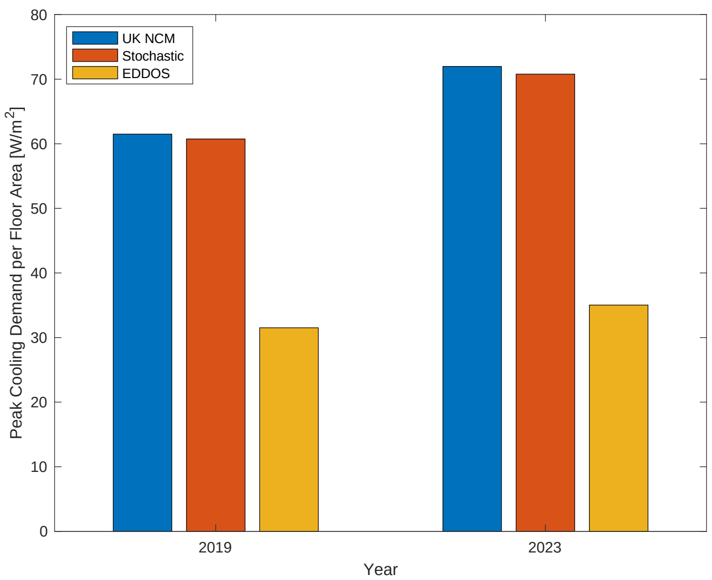

Figure 11. Peak cooling demand as estimated by the building model populated with standard-based (UK NCM) occupancy schedules, stochastic occupancy model, and energy-data-driven occupancy schedules (EDDOSs) in years 2019 and 2023.

For annual and peak cooling demand estimations in both 2019 and 2023, the NCM and stochastic models produced very similar results. In 2019, the annual cooling load was 9.05 kWh/m^2 for NCM and 9.10 kWh/m^2 for the stochastic model. The same trend continued in 2023, with annual cooling loads of 9.96 kWh/m^2 for NCM and 10 kWh/m^2 for the stochastic model. This suggests that the computationally expensive randomisation of the standard-based occupancy schedules via a stochastic occupancy model, at least through the method applied and documented in this study, does not change the building energy model performance in a meaningful manner. On the other hand, the energy-data-driven model (EDDOS) has reinterpreted the occupancy patterns in such a manner (while maintaining the total occupancy hours in a year) that it yields clearly different estimations of both annual and peak cooling energy demands. In both studied years, EDDOS has estimated noticeably higher annual cooling demands, 10.38 kWh/m^2 for 2019 and 13.34 kWh/m^2 for 2023, which can be mostly attributed to its higher occupancy densities during the lunch-break period (see Figure 7). In general, the year 2023 has shown higher annual cooling demands. This corresponds with Cooling Degree Days of 268 for 2023 compared to 246 for 2019, based on the hourly weather data file used for building performance simulation [56].

In the case of peak cooling demand, as illustrated in Figure 11, EDDOS has estimated noticeably lower values compared with the other two occupancy-modelling scenarios. Again, the current study does not have the possibility of comparing the accuracy of these predictions. However, the major difference in peak demand estimations of EDDOS and the standard-based schedule, as compared with the almost non-existent difference between the predictions of the stochastic and standard-based schedules, is worth highlighting. This seems to suggest that, contrary to the common argument for the use of stochastic occupancy models, in the current study, it is the rather simple EDDOS model (and not the stochastic model) that has provided a meaningfully diversified pattern of occupancy across the year. This finding raises important questions about the assumptions underlying occupancy modelling for building stock level energy modelling and emphasises the need for further investigation into the conditions under which different models may yield more accurate or useful predictions.

6. Conclusions

The study presented here conducted a comparison of three different occupancy-modelling approaches, including deterministic standard-based schedules (NCM), a stochastic model of occupants' presence, and a novel energy-data-driven occupancy model (EDDOS), for the purpose of building stock energy modelling. This involved implementation of these models in a building energy model built within a stock modelling campaign and testing the performance of the resulting occupied building models in estimating annual and peak heating and cooling demands. Key findings of the study were as follows:

1. The energy-data-driven model (EDDOS) demonstrated superior performance in estimating annual heating demands, with errors of ±2.5% compared to 14% and 7% for the standard-based and stochastic occupancy models in 2019 and 2023, respectively. This highlights the potential of energy-data-driven approaches in capturing realistic occupancy patterns.
2. The stochastic occupancy model, despite its higher computational cost, did not significantly improve heating demand estimations compared to standard-based schedules. For instance, in 2019, NCM estimated 32.34 kWh/m^2, while the stochastic model estimated 32.41 kWh/m^2, a negligible difference of 0.07 kWh/m^2.
3. For peak heating demand estimation, all models performed similarly, with EDDOS showing a slight edge (28% error compared to 29.5% for both stochastic and NCM in 2019).
4. In cooling demand estimations, the energy-data-driven model (EDDOS) yielded notably different results compared to standard-based and stochastic models. For example, in 2023, EDDOS estimated an annual cooling demand of 13.34 kWh/m^2, which was significantly higher than the estimations of the standard-based and stochastic models (9.96 kWh/m^2 and 10 kWh/m^2, respectively).
5. The energy-data-driven model estimated lower peak cooling demands compared to other occupancy models, suggesting a more diverse pattern of occupancy throughout the year.

Furthermore, the study revealed that even though all occupancy models were set up to maintain a consistent total number of occupancy hours (1824 h per occupant annually), the distribution of these hours throughout the day can significantly influence energy consumption patterns. In addition, the study's consideration of both pre-COVID (2019) and post-COVID (2023) scenarios highlights the adaptability of energy-data-driven models to changing occupancy patterns, which can be particularly useful in non-residential settings where occupancy patterns can vary widely over time and across different building types and evolve over time. In this context, the study suggested that the computationally expensive implementation of stochastic occupancy models in building stock models does not necessarily enhance their predictive performance. However, from the finding of this specific case study, the deployment of electricity use data to infer occupancy patterns seems to be a promising approach to capture the diverse and dynamic nature of occupants' presence for the purpose of building stock energy modelling.

In conclusion, while this study focused on a single building, its findings point to potentially transformative implications for urban-scale energy modelling and district energy system design. The energy-data-driven approach adopted in the development of the EDDOS model, if validated and refined further, could lead to more accurate assessments of urban energy systems and a more responsive and efficient design and operation of such systems.

Author Contributions: Conceptualization, S.A.-S. and F.T.; methodology, S.A.-S.; software, S.A.-S., F.T.; validation, S.A.-S.; formal analysis, S.A.-S.; investigation, S.A.-S.; data curation, S.A.-S.; writing—original draft preparation, S.A.-S.; writing—review and editing, S.A.-S., F.T.; visualization, S.A.-S.; supervision, F.T.; D.M.; project administration, R.T.; funding acquisition, F.T.; D.M. All authors have read and agreed to the published version of the manuscript.

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

Data Availability Statement: Data Availability Statement: The UK National Calculation Methodology (NCM) data used in this study are publicly available at https://www.uk-ncm.org.uk/ accessed on 10 September 2024. Weather data files were obtained from Shiny Weather Data (https://shinyweatherdata.com/ accessed on 10 September 2024). The building energy consumption data and specific building information used in this study are not publicly available. Interested researchers may contact the corresponding author for more information on the data used in this study.

Conflicts of Interest: The authors declare no conflict of interest.

Nomenclature

ACD	Annual Cooling Demand
AHD	Annual Heating Demand
ASHRAE	American Society of Heating, Refrigerating and Air-Conditioning Engineers
CDD	Cooling Degree Day
CO_2	Carbon Dioxide
DES	District Energy System
EDDOS	Energy-Data-Driven Occupancy Schedule
HDD	Heating Degree Day
HVAC	Heating, Ventilation, and Air Conditioning
NCM	National Calculation Method
UCL	University College London

References

1. Nejat, P.; Jomehzadeh, F.; Taheri, M.M.; Gohari, M.; Majid, M.Z.A. A global review of energy consumption, CO_2 emissions and policy in the residential sector (with an overview of the top ten CO_2 emitting countries). *Renew. Sustain. Energy Rev.* **2015**, *43*, 843–862.
2. Charani, S.C.; Rismanchi, B.; Foliente, G. Energy master planning for net-zero emission communities: State of the art and research challenges. *Renew. Sustain. Energy Rev.* **2021**, *137*, 110600.
3. Koch, A.; Girard, S.; McKoen, K. Towards a neighbourhood scale for low- or zero-carbon building projects. *Build. Res. Inf. Int. J. Res. Dev. Demonstr.* **2012**, *40*, 527–537.
4. Chow, Y. Utilizing district energy system as a cost-effective measure in meeting UK domestic 'zero carbon' targets. *Int. J. Low-Carbon Technol.* **2009**, *4*, 169–174. [CrossRef]
5. O'brien, W. *Conclusion, Research Needs, and Future Directions. Modeling, Design, and Optimisation of Net-Zero Energy Buildings*; Ernst & Sohn: Berlin, Germany, 2015.

6. Walker, S.; Labeodan, T.; Boxem, G.; Maassen, W.; Zeiler, W. An assessment methodology of sustainable energy transition scenarios for realizing energy neutral neighborhoods. *Appl. Energy* **2018**, *228*, 2346–2360. [CrossRef]
7. Nord, N.; Schmidt, D.; Kallert, A.M.D. Necessary Measures to Include more Distributed Renewable Energy Sources into District Heating System. *Energy Procedia* **2017**, *116*, 48–57. [CrossRef]
8. Lyu, Y.; Pan, Y.; Huang, Z. Development and Test of a New Fast Estimate Tool for Cooling and Heating Load Prediction of District Energy Systems at Planning Stage. *Buildings* **2022**, *12*, 1671. [CrossRef]
9. An, J.; Yan, D.; Hong, T.; Sun, K. A novel stochastic modeling method to simulate cooling loads in residential districts. *Appl. Energy* **2017**, *206*, 134–149. [CrossRef]
10. Zhang, J.; Hu, Y.; Yaun, Y.; Yuan, H. Accuracy improvement of the load forecasting in the district heating system by the informer-based framework with the optimal step size selection. *Energy* **2024**, *291*, 130347.
11. Menapace, A.; Santopietro, S.; Gargano, R.; Righetti, M. Stochastic Generation of District Heat Load. *Energies* **2021**, *14*, 5344. [CrossRef]
12. Aoun, N.; Baviere, R.; Vallee, M.; Sandou, G. Load shifting of space-heating demand in district heating systems based on a reduced-order building model identifiable at substation level. In Proceedings of the 4th Smart Energy Systems and 4th Generation District Heating Conference, Aalborg, Denmark, 13–14 November 2018.
13. Yan, D.; O'brien, W.; Hong, T.; Feng, X.; Gunay, H.B.; Tahmasebi, F.; Mahdavi, A. Occupant behavior modeling for building performance simulation: Current state and future challenges. *Energy Build.* **2015**, *107*, 264–278. [CrossRef]
14. Tahmasebi, F.; Mahdavi, A. The sensitivity of building performance simulation results to the choice of occupants' presence models: A case study. *J. Build. Perform. Simul.* **2015**, *10*, 625–635. [CrossRef]
15. Gaetani, I.; Hoes, P.-J.; Hensen, J.L.M. Occupant behavior in building energy simulation: Towards a fit-for-purpose modeling strategy. *Energy Build.* **2016**, *121*, 188–204.
16. Schweiker, M. Understanding occupants' behaviour for energy efficiency in buildings. *Curr. Sustain. Energy Rep.* **2017**, *4*, 8–14. [CrossRef]
17. Azar, E.; Menassa, C.C. A comprehensive analysis of the impact of occupancy parameters in energy simulation of office buildings. *Energy Build.* **2012**, *55*, 841–853. [CrossRef]
18. He, Z.; Hong, T.; Chou, S.K. A Framework for Estimating the Energy-saving Potential of Occupant Behaviour Improvement. *Appl. Energy* **2021**, *287*, 116591.
19. Hong, T.; Taylor-Lange, S.C.; D'oca, S.; Yan, D.; Corgnati, S.P. Advances in research and applications of energy-related occupant behavior in buildings. *Energy Build.* **2016**, *116*, 694–702. [CrossRef]
20. Maracchini, G.; Latini, A.; Giuseppe, E.G.; Gianangeli, A.; D'orazio, M. What Matters the Most? The Role of Actual Occupancy Patterns and Automatic Model Calibration in Reducing the Building Energy Performance Gap in an Italian Case Study. In *Sustainability in Energy and Buildings 2023*; Littlewood, J.R., Jain, L., Howlett, R.J., Eds.; Springer Singapore: Singapore, 2024; pp. 237–247.
21. Carpino, C.; Loukou, E.; Heiselberg, P.; Arcuri, N. Energy performance gap of a nearly Zero Energy Building (nZEB) in Denmark: The influence of occupancy modelling. *Build. Res. Inf.* **2020**, *48*, 899–921. [CrossRef]
22. Barthelmes, V.M.; Becchio, C.; Corgnati, S.P. Occupant behaviour lifestyles in a residential nearly zero energy building: Effect on energy use and thermal comfort. *Sci. Technol. Built Environ.* **2016**, *22*, 960–975.
23. Ebuy, H.T.; Bril El Haouzi, H.; Benelmir, R.; Pannequin, R. Occupant Behavior Impact on Building Sustainability Performance: A Literature Review. *Sustainability* **2023**, *15*, 2440. [CrossRef]
24. Pristerà, G.; Allacker, K.; Rock, M.; Sala, S. Archetype selection process for the development of a building stock model. *IOP Conf. Ser. Earth Environ. Sci.* **2023**, *1196*, 12013.
25. Yang, Y.; Gu, Q.; Wei, H.; Liu, H.; Wang, W.; Wei, S. Transforming and validating urban microclimate data with multi-sourced microclimate datasets for building energy modelling at urban scale. *Energy Build.* **2023**, *295*, 113318. [CrossRef]
26. Happle, G.; Fonseca, J.A.; Schlueter, A. A review on occupant behaviour in urban building energy models. *Energy Build.* **2018**, *174*, 276–292.
27. NCM. UK's National Calculation Method for Non Domestic Buildings. 2021. Available online: https://www.uk-ncm.org.uk/index.jsp (accessed on 10 September 2024).
28. Zhang, D.; Javanroodi, K.; Nik, V.M.; Chen, Y. Using urban building energy modeling to quantify the energy performance of residential buildings under climate change. *Build. Simul.* **2023**, *16*, 1629–1643.
29. Spyridon, S. Generating Occupancy Profiles for Building Simulations Using a Hybrid GNN and LSTM Framework. *Energies* **2023**, *16*, 4638. [CrossRef]
30. Cui, Y.; Zhu, Z.; Zhao, X.; Li, Z. Energy Schedule Setting Based on Clustering Algorithm and Pattern Recognition for Non-Residential Buildings Electricity Energy Consumption. *Sustainability* **2023**, *15*, 8750. [CrossRef]
31. Malatesta, T.; Li, Q.; Breadsell, J.K.; Eon, C. Distinguishing Household Groupings within a Precinct Based on Energy Usage Patterns Using Machine Learning Analysis. *Energies* **2023**, *16*, 4119. [CrossRef]
32. Hong, T.; Malik, J.; Krelling, A.; O'Brien, W.; Sun, K.; Lamberts, R.; Wei, M. Ten questions concerning thermal resilience of buildings and occupants for climate adaptation. *Build. Environ.* **2023**, *244*, 110806. [CrossRef]
33. Nejadshamsi, S.; Eicker, U.; Wang, C.; Bentahar, J. Data sources and approaches for building occupancy profiles at the urban scale—A review. *Build. Environ.* **2023**, *238*, 110375. [CrossRef]

34. Krishnan, D.; Kelly, S.; Kim, Y. A Meta-Analysis Review of Occupant Behaviour Models for Assessing Demand-Side Energy Consumption. *Energies* **2022**, *15*, 1219. [CrossRef]
35. Haldi, F.; Robinson, D. Interactions with window openings by office occupants. *Build. Environ.* **2009**, *44*, 2378–2395. [CrossRef]
36. Mahdavi, A.; Tahmasebi, F. The deployment-dependence of occupancy-related models in building performance simulation. *Energy Build.* **2016**, *117*, 313–320.
37. Tahmasebi, F.; Mahdavi, A. An inquiry into the reliability of window operation models in building performance simulation. *Build. Environ.* **2016**, *105*, 343–357.
38. Vassiljeva, K.; Matson, M.; Ferrantelli, A.; Petlenkov, E.; Thalfeldt, M.; Belikov, J. Data-Driven Occupancy Profile Identification and Application to the Ventilation Schedule in a School Building. *Energies* **2024**, *17*, 3080. [CrossRef]
39. Li, T.; Liu, X.; Li, G.; Wang, X.; Ma, J.; Xu, C.; Mao, Q. A systematic review and comprehensive analysis of building occupancy prediction. *Renew. Sustain. Energy Rev.* **2024**, *193*, 114284. [CrossRef]
40. Rugani, R.; Picco, M.; Salvadori, G.; Fantozzi, F.; Marengo, M. A numerical analysis of occupancy profile databases impact on dynamic energy simulation of buildings. *Energy Build.* **2024**, *310*, 114114. [CrossRef]
41. Vosoughkhosravi, S.; Thomas, M.A.; Jafari, A. Developing Data-Driven Occupancy Detection Models Based on Individual Plug Load Profiles in Office Spaces. *Comput. Civ. Eng.* **2023**, *2023*, 979–986.
42. Zhou, H.; Yu, J.; Zhao, Y.; Chang, C.; Li, J.; Lin, B. Recognizing occupant presence status in residential buildings from environment sensing data by data mining approach. *Energy Build.* **2021**, *252*, 111432. [CrossRef]
43. Johari, F.; Shadram, F.; Widén, J. Urban building energy modeling from geo-referenced energy performance certificate data: Development, calibration, and validation. *Sustain. Cities Soc.* **2023**, *96*, 104664. [CrossRef]
44. Hou, H.; Pawlak, J.; Sivakumar, A.; Howard, B. Building occupancy modelling at the district level: A combined copula-nested hazard-based approach. *Build. Environ.* **2022**, *225*, 109661.
45. Samareh Abolhassani, S.; Zandifar, A.; Ghourchian, N.; Amayri, M.; Bouguila, N.; Eicker, U. Improving residential building energy simulations through occupancy data derived from commercial off-the-shelf Wi-Fi sensing technology. *Energy Build.* **2022**, *272*, 112354.
46. Kourgiozou, V.; Salam, A.-S.; Ivan, K.; Mark, D.; Andrew, C.; Rui, T.; Dimitrios, R.; Dejan, M. Development of a dynamic building stock model for smart energy transition decision support: University campus stock case study. In Proceedings of the Building Simulation 2023: 18th Conference of IBPSA. IBPSA, Shanghai, China, 4–6 September 2023.
47. SimStock. Available online: http://cee.ac.uk/simstock/ (accessed on 10 October 2023).
48. Dong, J.; Schwartz, Y.; Korolija, I.; Mavrogianni, A.; Mumovic, D. London School Building Stock Model for Cognitive Performance Assessment. In Proceedings of the 5th Building Simulation and Optimization Conference, Loughborough, UK, 21–22 September 2020.
49. Ghiassi, N.; Tahmasebi, F.; Mahdavi, A. Harnessing buildings' operational diversity in a computational framework for high-resolution urban energy modeling. *Build. Simul.* **2017**, *10*, 1005–1021. [CrossRef]
50. Page, J.; Robinson, D.; Morel, N.; Scartezzini, J.-L. A generalised stochastic model for the simulation of occupant presence. *Energy Build.* **2007**, *40*, 83–98. [CrossRef]
51. Divkovic, D.; Knorr, L.; Meschede, H. Design approach to extend and decarbonise existing district heating systems—Case study for German cities. *Int. J. Sustain. Energy Plan. Manag.* **2023**, *38*, 141–156. [CrossRef]
52. Gonzalo, F.D.A.; Santamaría, B.M.; Burgos, M.J.M. Assessment of Building Energy Simulation Tools to Predict Heating and Cooling Energy Consumption at Early Design Stages. *Sustainability* **2023**, *15*, 1920. [CrossRef]
53. Edtmayer, H.; Fochler, L.-M.; Mach, T.; Fauster, J.; Schwab, E.; Hochenauer, C. High-resolution, spatial thermal energy demand analysis and workflow for a city district. *Int. J. Sustain. Energy Plan. Manag.* **2023**, *38*, 47–64. [CrossRef]
54. Rodríguez, L.R.; Ramos, J.S.; Domínguez, S. Building geometry data from online maps for accurate thermal simulations of districts. *Build. Simul.* **2023**, *16*, 1659–1674. [CrossRef]
55. Martinazzoli, G.; Pasinelli, D.; Lezzi, A.M.; Pilotelli, M. Design of a 5th Generation District Heating Substation Prototype for a Real Case Study. *Sustainability* **2023**, *15*, 2972. [CrossRef]
56. Shiny Weather Data. 2021. Available online: https://shinyweatherdata.com/ (accessed on 10 October 2023).
57. Raftery, P.; Keane, M.; Costa, A. Calibrating whole building energy models: Detailed case study using hourly measured data. *Energy Build.* **2011**, *43*, 3666–3679. [CrossRef]
58. Coakley, D.; Raftery, P.; Keane, M. A review of methods to match building energy simulation models to measured data. *Renew. Sustain. Energy Rev.* **2014**, *37*, 123–141. [CrossRef]
59. Chaudhary, G.; New, J.; Sanyal, J.; Im, P.; O'Neill, Z.; Garg, V. Evaluation of "Autotune" calibration against manual calibration of building energy models. *Appl. Energy* **2016**, *182*, 115–134.
60. Im, P.; Joe, J.; Bae, Y.; New, J.R. Empirical validation of building energy modeling for multi-zones commercial buildings in cooling season. *Appl. Energy* **2020**, *261*, 114374. [CrossRef]