energies

Review

Integrating Occupant Behaviour into Urban-Building Energy Modelling: A Review of Current Practices and Challenges

Alessia Banfi [1], Martina Ferrando [1], Peixian Li [2], Xing Shi [2] and Francesco Causone [1,*]

1 Energy Department, Politecnico di Milano, Via Lambruschini 4, 20156 Milano, Italy; alessia.banfi@polimi.it (A.B.); martina.ferrando@polimi.it (M.F.)
2 College of Architecture and Urban Planning, Tongji University, No. 1239 Si Ping Road, Shanghai 200092, China; lipx@tongji.edu.cn (P.L.); 20101@tongji.edu.cn (X.S.)
* Correspondence: francesco.causone@polimi.it

Abstract: Urban-Building Energy Modelling (UBEM) tools play a crucial role in analysing and optimizing energy use within cities. Among the available approaches, the bottom-up physics-based one is the most versatile for urban development and management applications. However, their accuracy is often limited by the inability to capture the dynamic impact of occupants' presence and actions (i.e., Occupant Behaviour) on building energy use patterns. While recent research has explored advanced Occupant Behaviour (OB) modelling techniques that incorporate stochasticity and contextual influences, current UBEM practices primarily rely on static occupant profiles, due to limitations in the software itself. This paper addresses this topic by conducting a thorough literature review to examine existing OB modelling techniques, data sources, key features and detailed information that could enhance UBEM simulations. Furthermore, the flexibility of available UBEM tools for integrating advanced OB models will be assessed, along with the identification of areas for improvement. The findings of this review are intended to guide researchers and tool developers towards creating more robust and occupant-centric urban energy simulations.

Keywords: Occupant Behaviour (OB); Urban-Building Energy Modelling (UBEM); building performance simulation; occupancy; human activities

check for updates

Citation: Banfi, A.; Ferrando, M.; Li, P.; Shi, X.; Causone, F. Integrating Occupant Behaviour into Urban-Building Energy Modelling: A Review of Current Practices and Challenges. *Energies* **2024**, *17*, 4400. https://doi.org/10.3390/en17174400

Academic Editor: Christian Inard

Received: 7 August 2024
Revised: 27 August 2024
Accepted: 30 August 2024
Published: 3 September 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The rapid increase in urban population demands innovative approaches to city design and energy management, aimed at enhancing living standards while reducing greenhouse gas emissions and mitigating local climate impacts [1]. The buildings sector alone contributes to over one-third of global final energy use and energy-related CO_2 emissions [2], highlighting the urgent need to reduce its environmental burden.

In this context, Urban-Building Energy Modelling (UBEM) represents a valuable approach for analysing and optimizing energy consumption within cities. By offering a comprehensive framework to simulate and evaluate energy use across a diverse building stock, UBEM tools can provide valuable insights for urban planners, policymakers, and energy managers [3]. Among the available approaches, the bottom-up physics-based one stands out for its versatility in urban development and management applications. This approach is based on the calculation, and subsequent aggregation, of each single building's energy consumption via energy balance equations [3,4]. Factors such as the building's geometry, materials, occupancy patterns, HVAC systems, weather conditions, solar radiation, and building orientation are considered to provide estimations with high spatiotemporal resolution [3,4]. For these reasons, this paper focuses only on this approach. and in the following sections the term "UBEM" will refer specifically to "bottom-up physics-based UBEM" unless otherwise specified.

Nevertheless, performing highly detailed large-scale simulations presents challenges, including significant computational time and effort. Consequently, multiple approximations are often adopted in UBEM modelling frameworks, which, combined with uncertainties in input data, can lead to discrepancies between simulated and measured energy data [5–7].

A similar performance gap exists in single Building Performance Simulation (BPS), primarily due to the description of building occupants' presence and energy-related actions, often referred to as Occupant Behaviour (OB) [8]. Occupants' presence (i.e., occupancy) directly impacts the internal loads of a building, as well as, for most end uses, its operational conditions [9,10]. Additionally, building operations are influenced by occupants' direct actions on thermostats, blinds, or light switches [10]. Over the years, efforts have been made to formalize OB modelling frameworks and requirements in BPS [11,12] and various models with different levels of detail have been proposed [10,13]. Nonetheless, the complex, dynamic, stochastic, and heterogeneous nature of OB, coupled with data gathering difficulties and simulation software limitations, led to the continued reliance on deterministic and static profiles, such as those proposed by ASHRAE 90.1 [14] or other national and international standards [15–18].

When scaling up the analysis from the single building to the urban level, the approach to model OB usually remains the same, as the complexities encountered in single BPS persist. In UBEM, to reduce computational time and effort, buildings are described with the use of "archetypes", which are fully characterized standard building models able to represent a cluster (i.e., group) of buildings sharing similar characteristics [3]. Archetypes comprise information about building use, fixed values of certain geometric variables (e.g., window-to-wall ratio or percentage of heated area), construction assemblies, systems and occupants [4]. Each building in the model is then associated with an archetype based on variables such as use type (e.g., residential, commercial, educational), construction period, or construction type (e.g., traditional or prefabricated) [4].

Occupant-related information in archetypes is often in the form of static schedules differentiated by building-use type, which can be either customized by modellers, with direct inputs or dedicated editing functions (depending on the tool flexibility) or predefined and extracted from the ones proposed for building-level applications by standards or codes [3,19]. Therefore, the adoption of archetypes in UBEM results in the same OB schedules assigned to all buildings of the same cluster. This approach, however, can lead to homogenous load-curve predictions within archetype and building-use categories, failing to capture the inherent randomness of human behaviour and urban dynamics [9]. While this level of detail can be suitable for studies focused on large spatial and temporal scales (e.g., annual energy use predictions for large building stock) [20,21] most urban applications aimed at decarbonization and sustainability (e.g., renewable energy community design, demand–response management, grid flexibility) require a more refined description of spatiotemporal fluctuations in buildings' energy demand [9]. To effectively support these initiatives, a realistic, diverse, and stochastic description of occupants' daily activities and occupancy patterns is essential [22]. This level of detail is crucial for achieving accurate simulations that can inform policies and strategies for urban energy management and sustainability [23].

Consequently, researchers have begun exploring solutions to achieve this goal by proposing new advanced mathematical models capable of representing OB with greater heterogeneity and even stochasticity [19]. These efforts are supported by the proliferation of new and opportunistic datasets able to inform about human habits, which have become available due to the expansion of the Internet of Things (IoT) in everyday life (e.g., call-detail records, location-based service applications data, etc.) [23–25]. However, despite these advancements, the implementation of these models is not always feasible, due to the inherent limitations of currently available UBEM simulation tools [3]. Therefore, it is essential to accurately analyse the OB inputs of these tools to understand their constraints and identify development opportunities. This paper addresses this need by starting with a literature review of the available studies proposing advanced OB models (i.e., OB models that differ from the standard ones proposed by codes) for urban applications and examining the occupant-related input of the main bottom-up physics-based UBEM tools. Through this analysis, the paper aims to highlight the limitations and opportunities for development, guiding improvements towards more accurate and reliable urban energy simulations.

1.1. OB in Building Performance Simulation: Scaling up from the Single Building to the Urban Level

Despite the absence of an established scientific definition, OB in building performance simulation is intended to reflect the presence of occupants and their energy-related actions [9,10]. Occupancy refers to the number of individuals in a space at a given time, usually expressed as a normalized daily, weekly or annual profile. Occupant's actions, however, include activities such as operating windows, adjusting solar shades, controlling lighting, modifying thermostat settings, using appliances, managing domestic hot water (DHW), and adjusting clothing. A further distinction is made between "adaptive" and "non-adaptive" actions: adaptive actions are driven by comfort needs and triggered by environmental factors, while non-adaptive actions are routine tasks unrelated to environmental adaptation [26]. This differentiation, however, is context-dependent rather than absolute, as many energy-related actions can be habitual or influenced by building design and control systems or sociocultural norms [9]. To simplify the terminology, in this study the following OB attributes are defined and analysed:

- Occupancy (i.e., presence and number of individuals in a space at any given time). It is often represented by occupancy schedules or profiles, which indicate the expected number of occupants during different times of the day, week, or year. It can also include details on occupancy density or sensible and latent occupant heat loads.
- Thermostat control (i.e., actions occupants take to adjust thermostat settings for temperature regulation). It can be modelled by specifying setpoint temperatures for heating and cooling, as well as schedules for when these setpoints change.
- Lighting control (i.e., management of artificial lighting.) It can include schedules that specify when lights are turned on or/off, dimming controls or adjustments based on occupancy or daylight levels, and the power density related to the lighting devices installed in the space. Additionally, this can encompass the control of blinds and internal/external shades, which not only regulate lighting but can also influence solar heat gains.
- Electric appliances control (i.e., operation of various electrical devices and appliances by occupants). This is often represented as usage pattern schedules and power consumption profiles.
- HVAC systems control (i.e., management of HVAC systems by occupants). It includes actions such as turning on/off devices, adjusting fan speeds or selecting heating or cooling modes.
- Windows operation (i.e., actions taken to open or close windows for ventilation and temperature control). It is often modelled using schedules or rules that determine when windows are opened or closed based on indoor and outdoor conditions.
- DHW usage (i.e., DHW usage for activities like showering, washing dishes, and laundry). It is generally modelled with schedules and flow rates that represent the demand for hot water throughout the day.

While extensive research has been conducted on OB in single buildings [13,27], scaling the modelling to the urban level presents unique challenges and complexities. Since UBEM tries to capture the interconnection between buildings, environment and local microclimate, considering people's movement patterns through urban areas, influenced by factors like public transport locations, public space design, and pedestrian connectivity, becomes important to truly understand the energy use across the city [7,22]. Additionally, changes in human behaviour over time must be considered for long-term urban analyses or scenario evaluations [28].

To realistically simulate energy use patterns for groups of buildings at various spatial and temporal scales adopting the archetype approach in UBEM, it is crucial to introduce diversity among buildings of the same archetype, trying to reflect the variability in human behaviour and the influence of the surrounding urban environment, socio-economic conditions, and cultural practices. Nonetheless, data sources and methodologies that have been successfully used in single-building OB studies often need adaptation for urban-scale models [21]. For example, detailed occupancy and behaviour data can be collected using sensors and surveys in single buildings, providing high-resolution data for model calibration. However, scaling this approach to an entire urban area can be logistically or financially challenging [9]. Therefore, the most common approach to modelling OB in UBEM still relies on the adoption of standard deterministic schedules proposed for single building analyses in standards or codes [9].

Deterministic and homogeneous OB schedules may provide a generalized estimation with sufficient accuracy, balancing out over- and underestimation across different buildings, when the primary goal of UBEM is to provide an overall prediction of energy use with a low spatiotemporal resolution (e.g., annual prediction at city level) [9,20,21]. However, they lack the flexibility to account for the dynamic nature of urban environments where people's presence and actions can vary significantly across different buildings and times. Incorporating realistic, heterogeneous and stochastic OB models into UBEM is especially important in applications where accurate prediction of the timing and distribution of energy use is vital for ensuring optimal matching between energy supply and demand, and for achieving energy efficiency goals.

Examples include, but are not limited to, demand-side management, renewable-energy communities design and optimization or grid decarbonization. In these applications, the precise timing and distribution of energy use are critical [29,30]. Deterministic and homogeneous schedules might miss peak- or low-demand periods, failing to capture the potential for flexibility in energy use or leading to inefficiencies in energy distribution and potential grid instability. Similarly, the effective integration of renewable energy resources, such as solar and wind, depends on accurately predicting when and where energy will be consumed. Advanced OB models, incorporating heterogeneity and stochasticity, can allow for a more precise matching of energy supply with demand, reducing reliance on non-renewable backup sources.

To address these challenges and enhance the accuracy and applicability of UBEM, there has been a growing focus on the development of advanced, context-specific OB models, particularly supported by the proliferation of IoT data in recent years [24]. These datasets have proven useful in other fields [31–33] and have shown promise also in pilot studies in building science [28,34,35]. IoT devices can collect real-time data on occupancy and energy use across multiple buildings, offering a more comprehensive dataset to inform advanced OB modelling techniques, encompassing randomness and probability. Nonetheless, the integration of advanced OB models into UBEM tools requires robust data-integration techniques and remains a significant challenge.

1.2. Existing Reviews and Contribution of the Present Study

In recent years, there has been increasing interest in modelling OB for large spatial-scale applications, with novel techniques and datasets being evaluated [19,24]. Researchers have critically reviewed available studies to consolidate past research findings, identify current gaps, and define future opportunities for urban-scale OB modelling. For instance, Happle et al. [19] categorized urban OB modelling techniques based on their ability to account for randomness, granularity, and inter-individual diversity. Salim et al. [23] and Dong et al. [36] reviewed urban-scale human behaviours influencing energy use, alongside the supporting datasets and techniques from urban science and beyond. Similarly, Dabirian et al. [24] described available OB modelling techniques focusing on occupant-related attributes influencing UBEM simulation, while Nejadshamsi et al. [25] and Ren et al. [37] specifically reviewed occupancy profiles at the urban scale. Doma and Ouf [38] investigated the inputs and workflows of existing UBEM tools, providing a breakdown of the occupant-related features. Banfi et al. [9] provided a user-oriented overview of the current State of the Art, trying to summarize the overlapping or combinable results of the previous analyses with unified terminology.

However, despite these extensive efforts, the practical integration of OB features into current simulation software, particularly in bottom-up physics-based UBEM tools, is underexplored. This paper addresses this topic, providing a critical review of the available data sources and OB modelling techniques, key features, and detailed information that could enhance UBEM simulations. Furthermore, the flexibility of existing UBEM tools in integrating advanced OB models is assessed, along with the identification of areas for improvement. The findings of this review aim to provide researchers and tool developers with a practical guide towards creating more robust and occupant-centric urban energy simulations.

1.3. Structure of the Paper

The paper is structured in the following sections:

- Section 1: Introduction. This section provides an overview of the topic of UBEM and the challenges associated with modelling OB. It sets the context for the study, outlines the research objectives, and highlights the contributions of the present work.
- Section 2: Literature review and study selection. This section details the literature review and study selection process. It describes the search strategy, screening, and selection criteria used to identify relevant studies, and summarizes the selected studies according to their OB attributes, data sources, modelling approaches, and validation methods. Furthermore, the selection process and key features of the studied bottom-up physics-based UBEM tools are reported
- Section 3: Advanced OB modelling to support UBEM simulation. This section focuses on the data sources and modelling techniques for OB in urban-scale applications. It contrasts traditional data sources (i.e., in situ measurements and surveys) with new datasets enabled by modern technologies (i.e., location-based service application data and network connectivity data) and discusses various OB modelling techniques, focusing on deterministic stochastic and agent-based modelling approaches.
- Section 4: Flexibility of UBEM tools in incorporating OB models. This section evaluates the flexibility of existing UBEM tools in integrating advanced OB models, detailing the allowed OB-related input with the current capabilities and limitations.
- Section 5: Current challenges and potentials of development in UBEM tools and OB modelling. This section identifies and analyses the key limitations of current UBEM tools and explores development opportunities. It suggests future research directions and potential advancements in UBEM and OB modelling.
- Section 6: Conclusions. This section summarizes the main findings of the paper.

2. Literature Review and Study Selection

2.1. OB Modelling Study Selection

This review followed a structured approach akin to the PRISMA methodology [39] to identify and select relevant studies. The search was conducted on Scopus using a comprehensive query including a multiple set of keywords, selected based on the key concepts of the research question, and Boolean operators, such as AND, OR, and NO, to refine the results.

The first set of keywords included terms related to OB, such as “occupant behaviour”, “human behaviour” and “occupancy” in combination with terms related to modelling, profiling, and scheduling. The terms related to buildings, such as “building”, “house”, and “office” were added and combined with keywords related to urban contexts (e.g., “urban”, “district”, “neighbourhood” and “city”) and energy modelling and simulation. Then, the resulting studies underwent a two-step screening, with the title and abstract first and the full text later, to exclude documents non-relevant to the research scope. To reduce the number of studies, only publications written in English and with available full text have been included.

Selection criteria for the final dataset included the following: (1) the goal of the paper should be the proposal of a novel methodology to describe OB, and (2) the scale of analysis should include more than a single building. Studies by Richardson et al. [40–42] and Page et al. [43], despite being of a household scale, have been introduced, since they have been adopted by other researchers in urban studies [44–46]. Figure 1 reports the workflow of the review methodology.

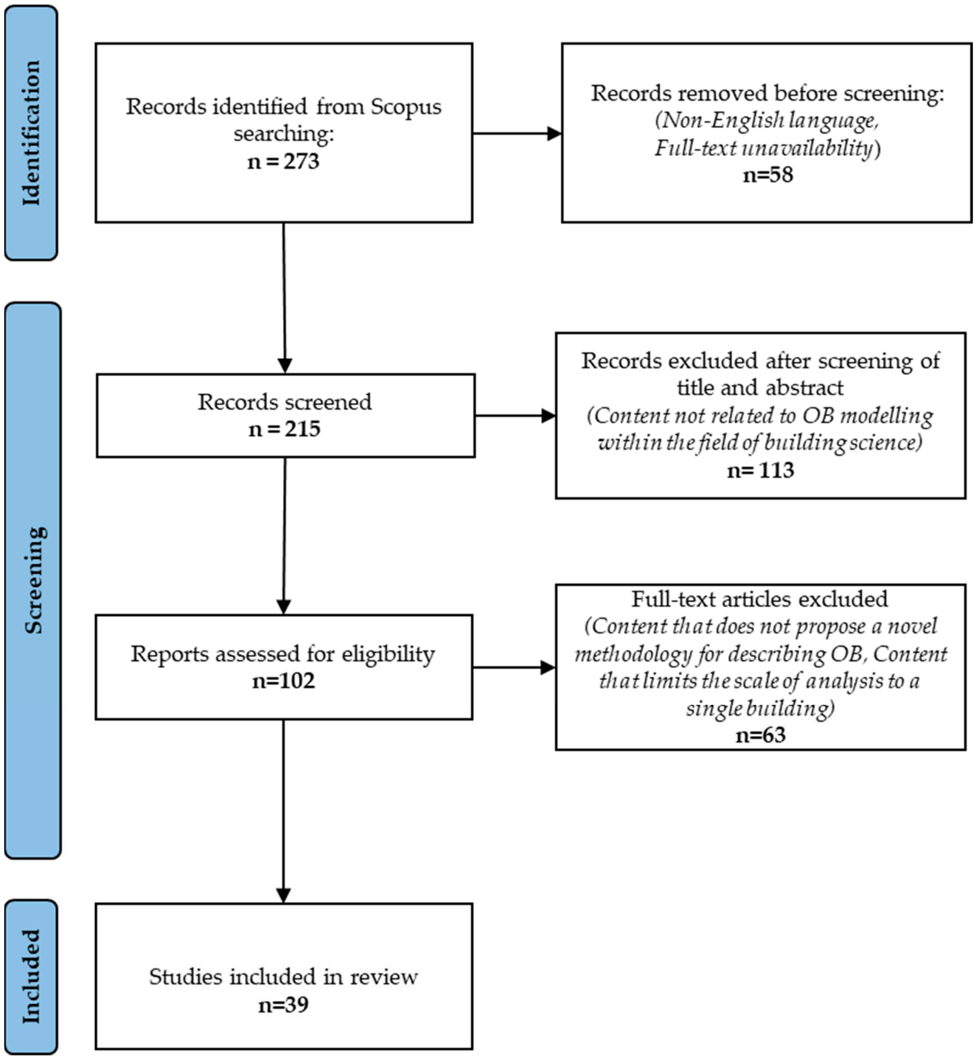

Figure 1. Workflow of the review methodology.

In accordance with to this procedure, 39 studies have been selected. These documents are summarized in Table 1, according to the modelled OB attributes (i.e., occupancy, thermostat control, lighting control, electric appliances control, HVAC systems control, window operations and DHW usage), the data sources and modelling approaches adopted for their definition, the validation procedure (if available) and, if included, the energy simulation tool employed to test their influence on energy output. Table 1 includes also the “Location” and “Building Use” columns, which refer, respectively, to the geographic and climatic context in which occupants are modelled and the building use to which OB attributes refer. Location is important because different climates can impact the way people move and occupy urban spaces, as well as their energy-related actions within buildings [47]. On the other hand, OB varies significantly across different building uses due to the distinct functions, activities, and occupancy patterns associated with each of them [9].

Table 1. Summary of the studies selected for review.

Authors (Year)	OB Modelled Attributes	OB Data Source	OB Modelling Approach	Building Type	Location (Climate *)	Models' Verification	Simulation in a Case Study	Energy Simulation Tool	Ref.
I. Richardson et al. (2008)	O	S	Sto	Residential	UK (Cfb)	Verification against UK TUS data	No		[40]
Page et al. (2008)	O	M	Sto	Residential, Office	Lausanne, Switzerland (Cfb)	Verification against measured O data	No		[43]
I. Richardson et al. (2009)	L	S	Sto	Residential	UK (Cfb)	Verification against Stokes et al. [48] model	No		[41]
I. Richardson et al. (2010)	EA	S	Sto	Residential	UK (Cfb)	Verification against measured electricity data	No		[42]
T. Rakha et al. (2014)	O	S	Det	n.d.	Massachussets, US (Dfa)	n.d.	No		[49]
R. Baetens et al. (2015)	O, EA, DHW, T	S	Sto	Residential	Belgium (Cfb)	Statistical verification	Yes	Modelica IDEA	[44]
E. Azar et al. (2016)	O	M	ABM	Residential, Office, Educational	Abu Dhabi, UEA (BWh)	Theoretical validation	No		[50]
J. Parker et al. (2017)	O	LBS	Det	Retail	UK (Cfb)	Comparison with reference schedules	No		[51]
J. An et al. (2017)	O, T, L, HVAC, W	S	Sto	Residential	Wuhan, China (Cfa)	Verification against measured energy data	Yes	DeST	[52]
T. Trondle et al. (2017)	O, HVAC	S	Sto	Residential	London, UK (Cfb)	n.d.	Yes		[53]
R. El Kontar et al. (2018)	O, L, HVAC, EA	M	Det	Residential	Austin, Texas, US (Cfa)	Verification against measured energy data	Yes	umi	[54]
A. Berres et al. (2019)	O	S	ABM	Office	Chicago, Illinois, US (Dfa)	n.d.	Yes	Energy Plus	[55]
D. M. Koupaei et al. (2019)	O	S	Sto	Residential	Des Moines, Iowa, US (Dfa)	n.d.	Yes	umi	[56]
G. Buttitta et al. (2019)	O, EA, L	S	Sto	Residential	UK (Cfb)	Verification against annual national energy use	Yes	EnergyPlus	[46]
C. Wang et al. (2019)	O	M	Det	Retail	Nanjing, China (Cfa)	Statistical verification	No		[57]
E. Barbour et al. (2019)	O	N	ABM	Residential, Retail, Industrial, Mixed-use	Boston, US (Dfa)	Comparison with reference schedules	Yes	umi	[34]
I. Mahmood et al. (2020)	O, EA	M	ABM	Residential	Islamabad, Pakistan (Cwa)	Verification against measured energy data	Yes	AnyLogic 8.2.4	[58]
W. Wu et al. (2020)	O	LBS	Det	Office, Retail, Restaurant	San Antonio, Texas, US (Cfa)	Comparison with reference schedules	Yes	CityBES	[59]
M. Mosteiro-Romero et al. (2020)	O, EA, DHW	N	ABM	Office, School	Zurich, Switzerland (Dfb)	Comparison with reference schedules	Yes	CEA	[60]
G. Happle et al. (2020)	O	LBS	Det	Retail, Restaurant	US	Comparison with reference schedules	No		[35]

Table 1. *Cont.*

Authors (Year)	OB Modelled Attributes	OB Data Source	OB Modelling Approach	Building Type	Location (Climate *)	Models' Verification	Simulation in a Case Study	Energy Simulation Tool	Ref.
Ueno et al. (2020)	T	M	Det	Residential	US	n.d.	No		[61]
M. Vellei et al. (2021)	O, T	S	Det for O ABM for T	Residential	Canada (Cfb)	Verification against measured energy data	Yes	DIMOSIM	[62]
K. Panchabikesan et al. (2021)	O	S	Det	Residential	Lyon, France (Cfb)	n.d.	No		[63]
X. Kang et al. (2021)	O	LBS	Det	Retail, Hospital, Transportation hubs	Beijing and Shanghai, China (Dwa, Cfa)	Comparison with reference schedules	Yes	Dest-C	[64]
Schumann et al. (2021)	O, T, EA	S	ABM	Residential	France	Verification against measured energy data	Yes	Modelica	[65]
H. Hou et al. (2022)	O	N	Sto	Office, Residential, Mixed-use	London, UK (Cfb)	Statistical verification	No		[28]
J. Chen et al. (2022)	DHW, L, EA O	S	Sto	Residential	US	Verification against measured energy data	Yes	ResStock	[66]
M. Ferrando et al. (2022)	O, EA	M	Det	Residential	Milan, Italy (Cfa)	Verification against measured energy data	Yes	umi	[21]
D. M. Koupaei et al. (2022)	O	S	Sto	Residential	US	Verification against American TUS data	No		[67]
Y. Wu et al. (2023)	T, HVAC	M	Det	Residential	China	Verification against measured energy data	Yes	DeST	[68]
X. Liu et al. (2023)	O	S	Det	Residential	Yinchuan and Chengdu, China (BWk, Cwa)	n.d.	No		[69]
M. Zhu et al. (2023)	O	LBS	Sto	Educational, Residential, Restaurant	Shanghai, China (Cfa)	Verification against measured O data	No		[70]
W. Jung et al. (2023)	O	M	Sto	Residential	US and Canada	n.d.	No		[71]
Z. Yu et al. (2023)	O, EA	S	ABM	Residential, Office, Commercial, Educational	Xi'an, China (Cwa)	Statistical verification	Yes	Dedicated model	[72]
D. Sood et al. (2023)	O	S	Det	Residential	UK	n.d.	Yes	EnergyPlus	[73]
W. Zhou et al. (2024)	O	S	Sto	Residential, Office, Educational	Lhasa, Tibet (BSk)	n.d.	Yes	umi	[74]
A. Doma et al. (2024)	O	M	Det	Residential	Canada	Verification against Canadian TUS data	No		[75]
Z. Liu et al. (2024)	O, HVAC	M	Sto	Residential	Hangzhou, China (Cfa)	n.d.	Yes	DeST	[76]
S.S. Abolhassani et al. (2024)	O	N	Det	Educational	Montreal, Canada (Dfb)	Statistical verification	Yes	Tool4Cities	[77]

* Climate classification according to Koppen [78]; O = occupancy, T = thermostat control, L = lighting control, EA = electric appliances control, HVAC = HVAC systems control, W = windows operation, DHW = domestic hot-water usage, M = in situ measurement, S = survey, LBS = location-based service application data, N = network connectivity data; Det = deterministic modelling, Sto = stochastic modelling, ABM = agent-based modelling.

Figure 2a provides a breakdown of the different OB attributes modelled in the reviewed studies. The predominant focus is on occupancy, with more than half of the studies concentrating on this attribute. This significant interest in occupancy can be linked to the lower complexity of gathering data about people's presence, rather than actions, in buildings, but also to the fundamental role played by this attribute. Among the aspects of OB, occupancy stands out as a fundamental and critical factor for its influence on various building performance aspects, including internal loads, ventilation requirements, and lighting needs [28]. Although occupants' actions also affect these factors, they are contingent on occupancy levels, display significant variability, and are not always controllable. Electric appliances control and thermostat control follow occupancy, with 15% and 11%, respectively, of studies modelling these factors. Fewer than 10% of the studies, however, investigate lighting control, windows operation, DHW usage and HVAC systems control.

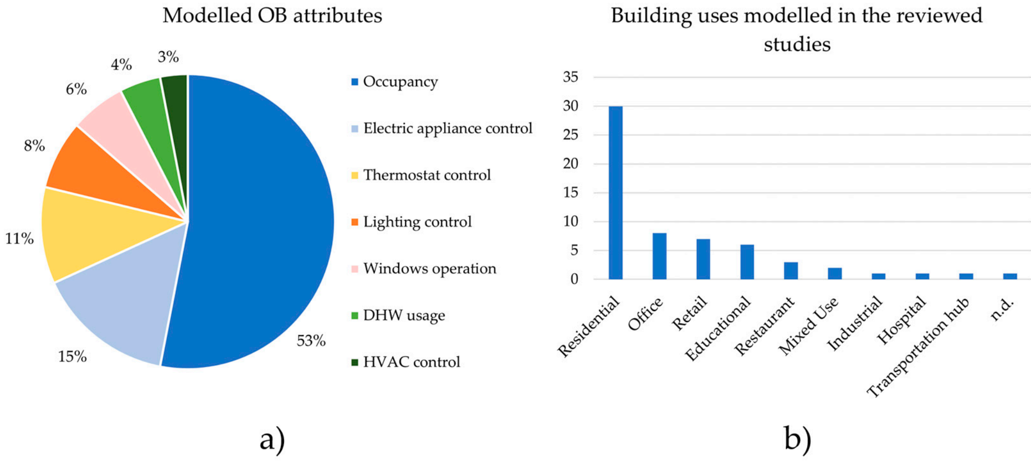

Figure 2. Breakdown of (**a**) the different OB attributes modelled in the reviewed studies and (**b**) the building uses that have been the focus of OB modelling.

Looking at the different building uses that have been the focus of OB modelling in the reviewed studies, Figure 2b reveals that residential buildings are the most frequently modelled, with nearly 30 studies focusing on this building use. Office buildings are the next most common, with around 10 studies. The interest in other building uses such as retail, educational or restaurants is less frequent. However, by modelling a diverse range of building uses, researchers can better understand and predict the complex interactions and energy flows within mixed-use districts, leading to more effective energy management and urban planning strategies. This holistic approach is essential for creating sustainable and efficient urban environments that cater to the diverse needs of their occupants.

Figure 3 shows where OB attributes have been modelled in various studies, using a colour gradient to indicate the number of studies conducted in each geographical region. The United States and China have the highest number of studies, indicating significant research interest. Europe, including the UK, and Canada show a moderate number of studies. South America, Africa, Oceania and parts of Asia are notably under-represented. This distribution highlights gaps in OB modelling research which might have different energy usage patterns and occupancy behaviours due to varying levels of infrastructure, economic development, and cultural practices. Also, tropical and arid climates are under-represented in the research, with a major focus on temperate climates such as in the United States and Europe, and humid subtropical climates, as in parts of China. Despite some exceptions, this result aligns with the trend observed by Carlucci et al. [10] in building-level applications. Nonetheless, understanding OB in different climates is crucial, as it influences how people use and occupy buildings.

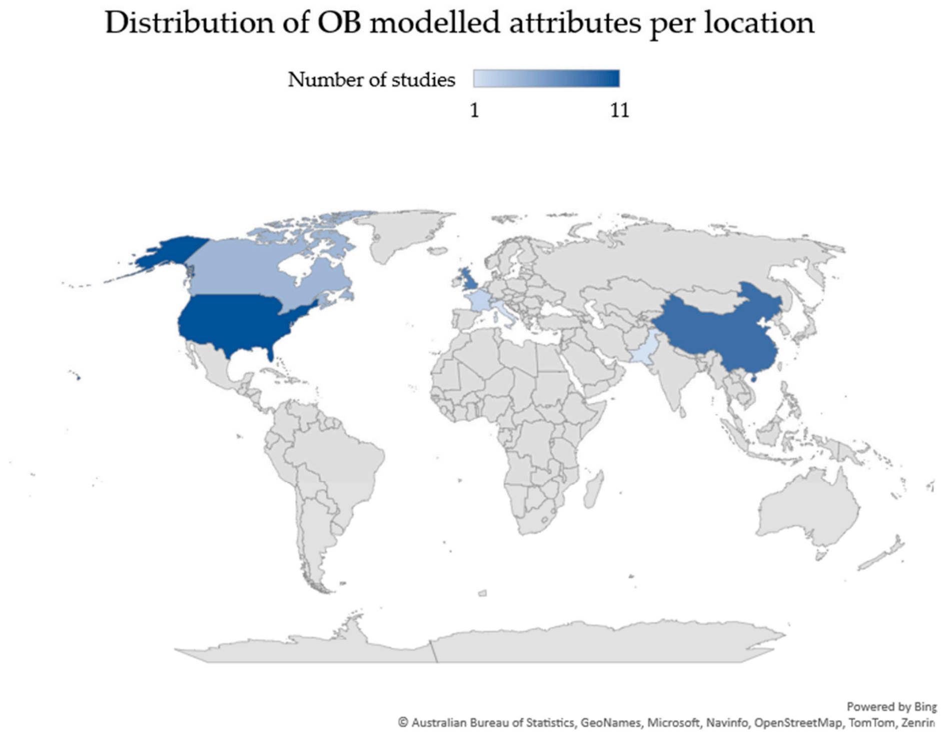

Figure 3. Distribution of OB modelled attributes by geographical location.

The main approach to verifying models consists of the comparison of newly defined advanced OB profiles against standard schedules, to check consistencies and differences. Verification of simulated energy results against metered data is another common verification procedure. However, around 30% of the reviewed studies lack a verification procedure, evidencing the need for a standardized method to validate advanced OB models. Of the 39 selected studies, 22 test the proposed OB models in case studies, and specifically, only 9 of them use bottom-up physics-based tools. This underscores the importance of improving the integration of OB in UBEM frameworks for accurately predicting energy consumption and optimizing building performance in urban environments.

2.2. Bottom-Up Physics-Based UBEM Tool Selection

To evaluate the capabilities of UBEM tools in handling advanced OB inputs the following tools were selected for detailed investigation: CitySim [79], Urban Modeling Interface (umi) [80], City Building Energy Saver (CityBES) [81], OpenIDEA [82], City Energy Analyst (CEA) [83], and IES-iCD [84]. The selection was based on the literature review of advanced OB studies and the tools employed therein, as performed in Section 2.1, their frequency and relevance in existing research, and their updated maintenance status.

CitySim [79], developed at the Solar Energy and Building Physics Laboratory of EPFL, supports the simulation and optimization of energy performance in urban environments. It is known for its advanced solar calculations and uses an equivalent reduced-order resistor–capacitor (RC) network for the building thermal model. It can simulate and optimise building-related resource flows based on four core models: the thermal model, radiation model, behavioural model and plant and equipment model.

Umi [80] is an urban modelling platform developed by the MIT Sustainable Design Lab as an extension for Rhinoceros 3D. It integrates analyses on building energy use, daylight, walkability, urban food production and district-level energy supply with a SketchUp-based interface and the EnergyPlus simulation engine.

CityBES [81] is a web-based data and computing platform developed by the Building Technology and Urban Systems Division of LBNL. It uses the CityGML schema and EnergyPlus to simulate the energy use of building stocks. Applications include energy benchmarking, energy retrofit analysis, district energy-system design, renewable energy analysis, heat resilience modelling, and urban mapping.

OpenIDEA [82] is an open framework for integrated building and district energy simulation, provided by KU Leuven University and based on the Modelica framework. This tool supports detailed and scalable energy analysis for both individual buildings and entire districts. It is inclusive of three modules: the Modelica IDEA library for system modelling, the stochastic residential occupant behaviour (StROBe) module and the Modelica FastBuildings library for building modelling.

CEA [83], released by ETH Zurich, supports the modelling and simulation of energy systems in cities, particularly in temperate and tropical climates, using an equivalent RC model. Equipped with a pre-stored database of building archetypes, it allows for energy-demand forecast, renewable-energy-potential assessment, and district energy-supply design and optimization.

Despite not being widely adopted in the literature, IES-iCD [83] is included in this study, as it represents the only commercially available UBEM tool. Commercialized by Integrated Environmental Solution as part of an integrated suite of four different tools, it allows energy use, solar potential and PV integration, water and accessibility assessments and simulations.

Tool4Cities [85], a suite of digital tools created by Concordia University to support urban sustainability and resilience, despite its promising integrated urban-energy modelling approach and its use in an advanced OB modelling case study [77], is still under development and, therefore, it has not been selected in the present study. Among the energy simulation tools employed in the reviewed studies, ResStock [86], although a comprehensive tool for residential building stock modelling, was not selected, due to its primary focus on national-scale residential building analysis rather than urban-scale modelling. Similarly, engines like Modelica [87], EnergyPlus [88], and DeST [89], although used to simulate advanced OB [46,52,55,64,65,68,73,76], as evidenced by the literature review performed (Table 1), were not included in the review because they are not considered bottom-up, physics-based UBEM tools. Instead, they are primarily used for detailed, component-level simulations of individual buildings and systems, rather than for integrated urban-scale energy modelling.

The decision not to conduct a comprehensive literature review on all UBEM tools is based on the existence of multiple detailed reviews that already cover various aspects of these tools. For example, Ferrando et al. [3] provide an extensive review of bottom-up physic-based UBEM tools, focusing on their development, capabilities, and applications. Hong et al. [7] pose and answer ten significant questions on UBEM, exploring advanced data models, datasets, modelling approaches, simulation tools, and model calibration. Reinhardt et al. [6] discuss input organization, thermal-model generation and execution, and result validation of UBEM tools. These reviews provide a solid foundation for understanding the general landscape of UBEM tools, allowing this study to focus more specifically on the tools' capabilities in advanced OB modelling. Readers interested in a more comprehensive understanding of UBEM tools, including their general capabilities and limitations, are encouraged to consult these multiple reviews.

3. Advanced OB Modelling to Support UBEM Simulation

3.1. Data Sources

Modelling OB at an urban scale presents significant challenges, due to the complexity and diversity of human activities and interactions within a city. The spatial and temporal stochastic nature of OB directly links to difficulties in collecting sufficient and reliable data about occupants [9]. Furthermore, OB encompasses both the presence of individuals and their energy-related actions, requiring data to be gathered and integrated from multiple and diverse sources [9].

This section aims to provide an overview of the various data sources emerging from the literature for OB characterization, highlighting whether they can offer information on occupants' presence, actions, or both. By categorizing these sources—ranging from traditional in situ measurements and surveys to modern location-based service applications and network connectivity— the inherent complexities in urban-scale OB data gathering are addressed, and potential pathways for improvement are proposed.

3.1.1. In Situ Measurements

Multiple technologies have been developed over the years to monitor OB, providing real-time data collection specific to the location being investigated [10]. These technologies can provide either direct or indirect measurements of OB attributes and are typically used to model people's presence and movement patterns, electric appliances and lighting control, thermostat control or HVAC control. However, they can be exploited also for windows operation and DHW usage [10]. Despite reliability, they come with significant drawbacks, including high sensor-installation and maintenance costs, limited coverage areas, and the necessity for physical access to measurement sites.

The most promising technologies for urban-scale applications emerge to be vision-based sensors, smart meters and connected thermostats [9]. Vision-based sensors, such as cameras or infrared detectors, represent a type of direct OB-sensing technology. These sensors can count the number of occupants, track their movements, and even detect specific activities in buildings and outdoor areas. Furthermore, they can be supported by machine learning algorithms to analyse visual data [90]. Their adoption in model urban OB is still limited, but they have been widely employed in building-level applications [10]. In particular, the advantage of vision-based sensors in urban studies is the possibility of relying on existing cameras (e.g., surveillance cameras), rather than installing new sensors. This approach avoids installation costs and potential resistance from citizens, but it can generate privacy concerns that should be accurately investigated.

Direct counting of people entering buildings, as performed by Wang et al. [57] to obtain the hourly occupancy profile of 12 commercial buildings in China, is another direct approach. This approach, however, is extremely time-consuming, hindering its adoption in large-scale applications.

On the other hand, smart meters and connected thermostats belong to the class of indirect OB attribute-sensing technology. Smart meters provide detailed electricity usage, typically at time resolutions of 15 min. Analysis of these data via machine learning techniques has proven to discover typical electric load curves and infer OB schedules [91]. El Kontar and Rakha [54] proposed a framework based on k-means clustering to model occupancy and related energy (i.e., lighting and electric appliances) loads based on metered data. Similarly, Ferrando et al. [21] clustered smart meter registration to create data-driven schedules for electric use and occupancy, adding diversity to UBEM simulations. Despite the potential, smart meters are not widely used, and their installation can encounter population resistance, due to privacy concerns.

Smart thermostats not only control heating and cooling systems but also collect data on temperature preferences and occupancy patterns. Equipped with sensors, they can detect when people are present in a room and adjust the temperature accordingly. They have been recently employed by Jung et al. [71] and Doma et al. [75] to generate residential occupancy profiles and by Ueno et al. [61] to model thermostat control in households. These three studies leveraged an open dataset of smart thermostat registration called Donate Your Data (DYD) [92]. The DYD dataset derives from an initiative launched in 2015 by Ecobee, one of the leading vendors of smart thermostats in the US and Canada, where users were invited to voluntarily share their smart thermostat data for research purposes. Data include 5 min-interval collections of indoor and outdoor temperature and humidity, heating and cooling temperature setpoints, HVAC mode (i.e., heating or cooling), motion data, and user-defined schedules. This initiative is still ongoing and currently collects data from more than 200,000 households, making the DYD dataset a valuable resource for studies on building science, and beyond.

Wu et al. [68] investigated HVAC control strategies and thermostat adjustments in the Chinese hot-summer–cold-winter climate zone, based on monitoring data of more than 1200 household variable refrigerant flow systems. The collected variables included on/off status, operation mode, fan-speed gear and setting temperatures, and helped to identify five typical HVAC behaviour patterns. However, different from DYD, the dataset employed was not open source.

In summary, various technologies offer promising solutions for monitoring OB, although their adoption is influenced by factors such as cost, privacy concerns, and the scale of application. Initiatives like open data sharing are crucial in overcoming these challenges and maximizing the potential of these technologies for urban-scale applications. Additionally, by addressing privacy concerns and leveraging existing infrastructure, these technologies can be more effectively integrated into urban planning and building-performance simulations, ultimately enhancing energy efficiency.

3.1.2. Surveys

Surveys are a widely used method for collecting data on OB in urban environments. They offer unique insights into human behaviour, preferences, and subjective experiences that are often difficult to capture through sensor-based methods. They can be conducted by national institutes of statistics and government agencies, researchers or private companies [9]. Furthermore, they can be structured to reflect population characteristics at the national or regional level or be tailored to be representative of a specific case study. Examples of surveys dedicated to specific case studies can be found in the works of An et al. [52], Liu et al. [69], Yu et al. [72] and Zhou et al. [74]. These researchers created and submitted surveys in their investigated case study area, to obtain context and location-specific information about its population. Surveys can include general sociodemographic variables, such as age, sex or occupation status, but also more specific information about housing and household members, installed HVAC systems or appliances, preferences for indoor conditions, activities performed during the day, working hours or travel patterns [9].

A common type of survey employed for OB modelling is the Time Use Survey (TUS), collecting data on activities—including start and end times, and location—performed during a typical day by individuals selected to statistically represent the national population [93]. TUSs are currently available in more than 100 countries around the world, and can be structured as a full diary, light diary, and questionnaire [94]. To unify the data collection procedure among the member states, the European Union proposed a Harmonized European Time Use Survey (HETUS) [95] with a standardized design, to make data comparable across countries and over time.

TUSs provide high temporal resolution (1-to-10 min intervals) about a large variety of possible indoor and outdoor activities and locations, representing a promising solution to OB data-gathering complexities. American, French, UK and Belgian TUSs have already been successfully employed to model occupancy, electric appliances and DHW usage [40,44,53,65,67]. Richardson et al. [40], for example, leveraged a UK TUS to inform a model to stochastically generate residential occupancy patterns, diversified per household size. Subsequently, they proceeded to extend the modelling to lighting [41] and electric appliances control [42]. Similarly, surveys proved useful for modelling thermostat control and DHW usage [44,66]. Their integration for windows operation and HVAC control still needs to be tested.

Referring to Table 1, it is evident that most studies using TUSs focus on residential case studies, showing a lack of versatility for application in mixed-use urban areas. However, TUS data also include information on non-residential buildings, outdoor spaces (e.g., parks, lakes, beaches), and transportation modes (e.g., cars, trains, buses). This broader scope opens opportunities to integrate different urban domains, such as buildings and mobility, within a unified modelling framework.

TUSs can be coupled with other surveys (e.g., travel surveys, work surveys) and census data to include economic, and social information, detail on working hours, travel modes, durations, and purposes. An example of this integration can be found in the work of Buttitta et al. [46], where the UK TUS was combined with census data and analysed via data-mining techniques to develop occupancy-integrated UK residential archetypes. Qualitative sociological and psychological studies can offer additional valuable insights into the cultural, social, cognitive and emotional factors affecting human behaviour and energy use patterns in urban settings [9].

While surveys provide detailed representative data that can be aggregated at different spatial levels (e.g., individual, household, building, population), their reliability can be affected by researcher bias, the representativeness of the sample population, and participants' willingness to take part [93]. Additionally, carrying out large-scale surveys like TUSs involves significant resources, effort and time, limiting updates to every 5 to 10 years and hindering the use for simulation behavioural changes [93]. Nonetheless, surveys offer the significant advantage of capturing rich, detailed data on both occupants' presence and actions, and they can often be made available as open-source resources, facilitating broader access and collaborative research.

3.1.3. Location-Based Service Applications

Location-based service (LBS) applications are designed to operate using geolocation data from users. Thus, even though the purpose of these applications (e.g., social media, mapping platforms, tracking services, etc.) might be far from OB, they provide high spatiotemporal resolution data on mobility patterns, presence, and time spent in specific locations. While LBS data are limited in modelling energy-related actions in buildings, they can support the creation of realistic and context-specific occupancy profiles for large-scale applications.

Successful examples of data from social media check-ins and mapping platforms to model occupancy in both residential and non-residential buildings are present in the literature (Table 1). Google Popular Times [96], a feature available in Google Maps that provides real-time and historical data on the busiest times at various businesses and places, has been employed by Parker et al. [51] to extract occupancy profiles of twenty supermarkets in the UK, and by Happle et al. [35] to create data-driven occupancy schedules of restaurants and retail outlets in 13 different US cities. The advantage of Google Popular Times [96] data is their open-source availability and their presentation as normalized daily profiles with hourly time-steps (Figure 4), which simplifies the modelling process. Nonetheless, these data are limited to busy public places, and they may need integration with other data sources in case of simulation of mixed-use urban areas. Furthermore, despite the useful normalized form, the normalization factor of Google Popular Times [96] is unclear, adding uncertainty to simulations.

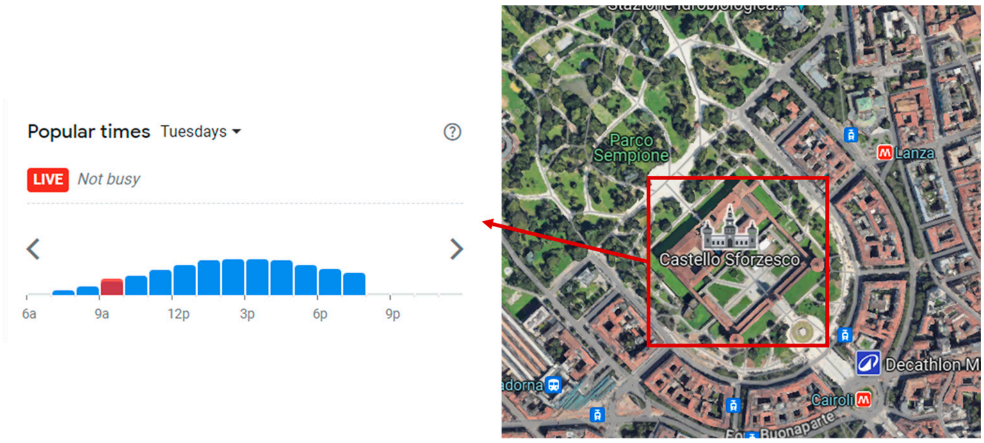

Figure 4. Example of real-time Google Popular Times [96] for a building of cultural interest in Milan, as extracted from Google Maps.

Kang et al. [64] leveraged social media data to extract, through clustering analysis, typical weekly occupancy profiles of 16 non-residential buildings (i.e., retail outlets, hospitals, railway stations and airports) in Beijing, China. The data employed in this study come from Tencent, one of the largest internet companies in China, and include positioning records from apps such as WeChat and QQ. However, privacy legislation and cost considerations may limit the accessibility and use of similar datasets.

Promising opportunities for occupancy modelling can also come from the use of GPS traces, as is often seen in the transportation field [33], and geo-tagged tweets from X Platform, whose frequency showed the potential to be a predictor of hourly, city-wide electricity use [97].

Overall, leveraging LBS data for urban OB modelling offers several advantages, including high spatiotemporal resolution and the ability to create detailed occupancy profiles for a wide variety of building uses. The popularity of using such datasets, especially open-source ones, is expected to increase.

3.1.4. Network Connectivity

Network connectivity data, such as Wi-fi logs or Call Detail Records (CDRs), provide insights into the number and types of devices connected to networks, connection durations and timings, and overall network-traffic patterns.

CDRs are metadata files that telecom operators generate for every telecommunication transaction that occurs within their network. Among others, they typically include information on the date and time when the data session started and ended, a unique identifier of each user, and an identifier for, and location coordinates of, the cell towers or base stations that managed the data session. In their study, Barbour et al. [34] started from CDR to investigate users' stay points in the Boston area. to estimate building occupancy. Through the application of an agent-based modelling framework developed for transportation applications, they managed to derive occupancy profiles for residential, industrial, retail and mixed-use buildings.

Hou et al. [28] and Abolhassani et al. [77] decided to employ Wi-Fi sensing data to infer building occupancy in urban applications. Wi-Fi sensing data are derived from the interaction of devices (such as smartphones, laptops, and IoT devices) with Wi-Fi networks. Similarly to CDRs, they includes information on the unique identifier of the device connecting to the Wi-Fi network, date and time of connection/disconnection occurrences and identifiers of Wi-Fi access points. While in UBEM Wi-Fi sensing data have been primarily used to model occupancy, especially in the case of interdependent buildings [28], in single-building analyses they has also proven useful for investigating occupant energy-related actions [98].

The advantages of network connectivity data include high spatial and temporal resolution and the ability to capture real-time occupancy patterns. Their integration into large-spatial-scale applications for the modelling of other OB attributes needs further investigation. However, privacy concerns and data security issues are significant limitations that need to be addressed.

Overall, each of the presented data sources encompasses advantages and disadvantages. Therefore, leveraging diverse data sources for urban OB modelling can be the best solution to obtain a deep understanding of complex interactions within urban spaces. The increasing availability of open-source datasets offers promising opportunities for enhancing UBEM simulations, though privacy and data-integration challenges remain significant barriers.

3.2. Modelling Techniques

From their initial application in single BPSs, OB modelling approaches have been adapted and expanded for UBEM use. Choosing the appropriate approach depends on a comprehensive exploration of the trade-offs between simplicity, accuracy, and computational demands, as well as on the research question and the data availability [20]. This section reviews the modelling formalism chosen by previous researchers in their urban OB modelling studies (Table 1), distinguishing between deterministic, stochastic and agent-based modelling (ABM), and investigating the statistical and mathematical techniques that support them.

For those interested in modelling techniques from other fields (e.g., transportation, risk management, epidemiology) which can support OB modelling, the work of Dong et al. [36] is recommended.

3.2.1. Deterministic Models

Deterministic models use fixed schedules or rule-based schemes to describe OB in buildings. Derived from statistical trends or observed data, these models have evolved from standard schedules to be more robust and context-specific. Typically, they take the form of 24 h normalized profiles that correlate OB attributes to their peak value (e.g., occupancy density, electric power density, etc.) or follow simple rules based on environmental triggers (e.g., switching lights on in the absence of sufficient illumination).

Out of the 39 reviewed studies, 15 propose data-driven deterministic models, primarily for occupancy but also for thermostat control, lighting control, electric appliance control, and HVAC control. Table 2 details the deterministic models, their mathematical or statistical techniques, and data sources.

Table 2. Data sources and modelling-technique details of the reviewed studies proposing deterministic OB models.

Ref.	OB Attributes	OB Data Source	OB Modelling Technique
[49]	O	Travel survey	k-means clustering
[51]	O	Google Popular Times	Direct use of Google Popular Times
[54]	O, L, HVAC, EA	Smart-meter registrations	k-means clustering
[59]	O	Private mobility data	Count of users in buildings at different hours
[35]	O	Google Popular Times	Direct use of Google Popular Times
[61]	T	Smart-thermostat registrations	k-means clustering
[63]	O	Smart-meter registrations	Shape-based clustering and change-point detection
[62]	O	French TUS, smart-thermostat registrations	Hierarchical agglomerative clustering
[64]	O	Network connectivity data	k-means clustering
[5]	O, EA	Smart-meter registrations	k-means clustering
[68]	T, HVAC	VRF sub-meter registrations	k-means clustering
[69]	O	Dedicated survey	k-means clustering
[75]	O	Smart-thermostat registrations	Rule-based model
[77]	O	Wi-Fi sensing	Random forest classification
[73]	O	UK TUS	k-mode clustering

O = occupancy, T = thermostat control, L = lighting control, EA = electric appliances control, HVAC = HVAC system control.

Clustering, especially k-means clustering, is the most adopted technique. k-means clustering is an unsupervised machine learning technique used to partition a dataset into a set number of distinct groups of similar data points. This technique is simple and robust, making it a common choice for clustering metered time-series data [91]. For instance, Ferrando et al. [21] implemented a k-means algorithm to cluster similar residential electricity-use patterns in Milan, to extract occupancy profiles. The same procedure subsequently allowed for the detection of variations in residential energy-use profiles experienced in Milan before, during, and after the COVID-19 emergency [99]. Ueno et al. [61] obtained 40 thermostat-adjustment schedules from smart-thermostat registrations through the k-means algorithm application. Furthermore, the k-means algorithm proved useful for extracting occupancy schedules from survey data, as seen in the works of Rakha et al. [49] and Liu et al. [69].

In contrast, Sood et al. [73], dealing with Time Use Survey (TUS) categorical data, used the k-mode algorithm, an extension of k-means for categorical data. Other deterministic OB modelling techniques used in the literature include hierarchical agglomerative clustering [62], shape-based clustering [63], and random forest classification [77].

Deterministic models are the simplest and least data-intensive, but they lack the stochasticity of human behaviour, resulting in fixed profiles without any variation. They usually assume the form of space-based profiles, meaning that their focus is on the aggregated impact of OB attributes on a specific space (i.e., a certain category archetype, a certain building or a single thermal zone), rather than on the individual movement and activities within the urban fabric. The space-based approach is advantageous for its simplicity and focus on building performance, but may lose individual behaviour detail and may limit integration with other urban domains (e.g., mobility).

Thus, deterministic models are well-suited for applications with large spatiotemporal resolutions [20], and are often preferred when data are insufficient for more detailed models.

3.2.2. Stochastic Models

Stochastic models use probabilistic approaches to account for the randomness and variability of human behaviour. They can generate sequences of activities based on the probability of occurrence and duration of events (e.g., occupancy of a space, switch on/off events, etc.), and can be informed by both historical and monitored data.

These models became popular among researchers to overcome the oversimplifications introduced by deterministic models [9]. Seventeen reviewed case studies propose stochastic modelling, covering the topics of occupancy, thermostat adjustment, lighting control, electric appliances control, HVAC system control, DHW usage and windows operations. Table 3 reports the details of data sources and modelling techniques employed.

Table 3. Data sources and modelling-technique details of the reviewed studies proposing stochastic OB models.

Ref.	OB Attributes	OB Data Source	OB Modelling Technique
[40]	O	UK TUS	First-Order Markov Chain model
[41]	L	Active occupancy from [40], CREST irradiance database	First-Order Markov Chain model for O combined with probabilistic switch-on event for L
[42]	EA	Active occupancy from [40], UK TUS, statistical appliance-ownership data	First-Order Markov Chain model for O combined with probabilistic switch-on event for EA
[43]	O	in situ measurements	First-Order Markov Chain model
[44]	O, EA, DHW, T	Belgian TUS, household budget survey	Survival model
[52]	O, T, L, HVAC, W	Dedicated survey	First-Order Markov Chain model for O combined with probabilistic switch-on event T, HVAC, L, W
[53]	O, HVAC	UK TUS, Census data	First-Order Markov Chain model
[56]	O	American TUS, dedicated survey	First-Order Markov Chain model
[57]	O	in situ measurements	Gaussian Mixture Model
[46]	EA	UK TUS, English Housing survey Household electricity survey	First-Order Markov Chain model
[28]	O	Wi-Fi sensing	Hazard-based model combined with copula approach
[66]	DHW, L, EA, O	American TUS	First-Order Markov Chain model combined with probabilistic sampling
[67]	O	American TUS	First-Order Markov Chain model
[70]	O	Location-based service application, in situ measurements	First-Order Markov Chain model and Bayesian Network
[71]	O	Smart-thermostat registrations	First-Order Markov Chain model
[74]	O	Dedicated survey	First-Order Markov Chain model
[76]	O, HVAC	VRF sub-meter registrations	First-Order Markov Chain model for O, three-parameter Weibull cumulative function for HVAC

O = occupancy, T = thermostat control, L = lighting control, EA = electric appliances control, HVAC = HVAC system control, DHW = domestic hot-water Usage, W = window operations.

A widely adopted technique is the First-Order Markov Chain (MC), which creates event sequences based on state transition probabilities. With First- Order MC, the probability of a state occurring is based only on the previous state; therefore, they are incapable of capturing long-term dependencies in the time-series data [100]. Higher-Order MCs can overcome this limitation; however, the trade-off between complexity and detail in modelling large-scale OB attributes is unclear. Richardson et al. [40–42] used the First-Order MC to model household occupancy, lighting use, and electric appliances, differentiating schedules by household size. Baetens et al. [44] used the First-Order MC and survival models for occupancy, thermostat adjustment, electric appliances control, and DHW usage. In particular, the transition between states is defined using the MC and the duration of each state in a survival model. Survival models, predicting the time before an event occurs, have been employed by Hou et al. [28] to model occupancy patterns within neighbourhoods, while Wang et al. [57] proposed a dynamic occupancy model for commercial buildings using Gaussian Mixture Models (i.e., generating data based on Gaussian distributions).

By incorporating randomness, stochastic modelling can offer a higher level of realism compared to deterministic models, making the generated OB profiles suitable for an application that necessitates high spatiotemporal resolution (e.g., demand response, grid flexibility, etc.). They can either have a space-based approach, as used by Buttitta et al. [46] for the creation of occupancy-integrated archetypes, or they can capture individual behaviour detail though a person-based approach, allowing for the modelling of individuals or groups of individuals with similar behaviour. An example of person-based stochastic modelling is found in the work of Chen et al. [66], where four typologies of occupants have been defined based on their daily presence/absence schedule at home. A person-based approach can improve the capabilities of the simulation, increasing the heterogeneity of OB and opening other topics, for example, the exploration of dedicated behavioural policies on districts' energy use (e.g., the adoption of remote working for a share of the population) or the integration of urban domains (e.g., mobility, public transportation, etc.). However, the complexity and data requirements of stochastic models can be significantly higher, necessitating comprehensive data collection and processing efforts.

3.2.3. Agent-Based Models

ABM simulates actions and interactions among autonomous agents and their environment [101]. Agents are intended as individual entities having a certain set of attributes and rules that govern their behaviour. Based on the level of detail of the model, agents can anticipate future consequences based on environmental sensing, learn from experience and develop a group behaviour [20].

In OB modelling, agents can be individuals or groups with similar characteristics, with their state and its evolution being deterministic or stochastic. A few examples, summarized in Table 4, are proposed in the literature for ABM applications with respect to OB, with novel models or framework adaptations from the transportation domain.

Mosteiro et al. [60] adapted the MATSim [102], originally for transportation analyses, to UBEM purposes, creating a synthetic population of agents distributed in the study area with spatial position and activity (i.e., work, study, and lunch) evolving in time. Barbour et al. [34] used the Time-Geo Framework [103] and CDRs to generate urban mobility patterns with stay durations of activities, number of visited locations per day, and daily mobility networks and eventually extract occupancy patterns to be inputted in UBEM simulation. Berres et al. [55] generated human mobility patterns and vehicle trips with TRANSIMs [104] to eventually map agents to nearby buildings, based on their final destinations. Then, the arrival and departure times of agents were used to create hourly building occupancy schedules.

Mahmood et al. [58] developed a three-layer ABM for estimating residential electricity consumption, representing building stock, occupants, and appliance agents, interacting dynamically to simulate aggregate electricity consumption. Schumann et al. [65] designed an ABM with autonomous agents capable of coordination among household members, for simulating daily occupant activities and electricity use, based on survey and census data, and TUS data. Yu et al. [72] proposed a combined ABM for predicting community heating loads by simulating occupant transfer behaviour, temperature preferences and thermostat adjustments.

Occupant agents in ABM can encompass a great variety of characteristics, depending on the model's purpose and granularity. The typical characteristics emerging from the reviewed studies are the following:

- Demographic attributes such as household size and employment status;
- Behavioural patterns such as daily routines or working hours;
- Energy usage habits such as preferences for heating/cooling, appliance usage patterns, and responses to environmental changes;
- Mobility patterns such as commuting habits and travel frequency;
- Environmental sensitivity such as thermal-comfort perception of temperature preferences;
- Interaction rules (i.e., how agents interact with each other), such as collective behaviour in a household.

These characteristics allow ABM to simulate the complex and dynamic interactions between occupants and their environments, leading to more detailed predictions of building performance and energy usage. ABM models can adapt to changes in the environment and agent behaviours, providing simulations that account for a wide range of scenarios. Nonetheless, ABM requires substantial data and computational resources, which can be a barrier to large-scale or highly detailed models. Ensuring the accuracy and validity of ABM simulations requires rigorous validation against real-world data, which can be complex and time-consuming. Developing and implementing ABM is complex, requiring expertise in modelling techniques and domain-specific knowledge (e.g., building energy performance). The level of detail of the ABM should be accurately chosen, based on the research question, to avoid the addition of unnecessary complexity. Thus, despite these challenges, the dynamic, adaptive, and data-driven nature of ABM makes it a valuable tool for researchers and practitioners in sustainable building performance and its integration into available UBEM tools should be investigated.

Table 4. Data sources and modelling-technique details of the reviewed studies proposing OB ABM.

Ref.	OB Modelled Attributes	OB Data Source	OB Modelling Technique
[50]	O	in situ measurement, historical data from the literature, Commercial Building Energy Consumption Survey	Tailored Agent-Based Model
[55]	O	National Household Travel Survey	TRANSIMS [104]
[34]	O	Call-detail Records	TimeGeo Framework [103]
[58]	O, EA	Smart-meter registrations	Tailored Agent-Based Model
[60]	O, EA, DHW	Employee registers and course-enrolment data	Tailored Agent-Based Model
[62]	O, T	French TUS, smart-thermostat registrations	Tailored Agent-Based Model
[65]	O, T, EA	French TUS	Tailored Agent-Based Model
[72]	O, EA	Dedicated survey	Tailored Agent-Based Model

O = Occupancy, T = thermostat control, EA = electric appliances control, DHW = domestic hot-water usage.

4. Flexibility of UBEM Tools in Incorporating Advanced OB Models

This section provides an in-depth analysis of the occupant-related features available in various UBEM tools, focusing on how they model occupancy, thermostat control, lighting control, electric appliances control, HVAC control, windows operation and DHW usage. The tools analysed are the ones proposed in Section 2.2. Table 5 summarizes the key findings of the proposed analysis.

It is worth noting that, independently of the specific input and modelling features, UBEM tools currently implement only space-based OB inputs. This implies that the tools do not model occupants as individuals moving and acting in the urban environment, but rather account for their impact on a specific building or thermal zone.

Table 5. Summary table of OB attributes in the analysed UBEM tools.

	Occupancy	Thermostat Control	Lighting Control	Electric Appliances Control	HVAC Control	Windows Operation	DHW Usage
CitySim	Det and Sto	Det	Det and Sto [1]	Det and Sto [1]	Det [1]	Det [1]	Det [1]
umi	Det	Det	Det	Det	Det	Det	Det
CityBES	Det	Det	Det	Det	Det	NS	Det
OpenIDEAS	Det and Sto	Det and Sto [1,2]	Det and Sto [1,2]	Det and Sto [1,2]	Det and Sto [1,2]	Det	Det and Sto [1,2]
CEA	Det and Sto	Det	Det	Det	Det	NS	Det
IES-iCD	Det [3]	Det [3]	Det [3]	Det [3]	Det [3]	NS	Det [1,3]

Det = deterministic, Sto = stochastic, NS = not supported; [1]: based on occupancy states, [2]: limited to residential buildings, [3]: non-customizable.

4.1. CitySim

CitySim [79] offers support for both deterministic and stochastic hourly-occupancy models. Deterministic models provide consistent patterns for weekdays, Saturdays, and Sundays without seasonal variability, and can be either predefined or customizable by the user. To add randomness to occupancy profiles, CitySim [79] offers the possibility of generating stochastic schedules based on the Page et al. [43] model. This model is built upon the First-Order MC, in which a fixed mobility parameter lightly varies transition probabilities over time, allowing for the creation of yearly profiles. The probabilistic modelling of long absences is also included. This model is informed by occupancy data deriving from standards, lacking the capability of capturing context-specific or real-time details. Creating space-based occupancy profiles, it does not allow for individual heterogeneity.

CitySim [79] also includes deterministic and stochastic models for window and blinds operation, and usage of lights and electric appliances. These behaviours are triggered based on occupancy states generated by the presence model.

Besides schedules, the tool offers the setting of standardized or customizable values of occupancy density, electric-equipment power density, and lighting power density. Thermostat control in CitySim [79] is deterministic, with fixed temperature setpoints. HVAC control is also deterministic, with on/off control based on occupancy schedules. DHW usage follows a deterministic schedule.

Another advantage that CitySim [79] can offer is the integration (still under development) with the MATSim toolkit [102], an interactive multi-agent transport model that could open up to the simulation of individuals' movement in the urban environment.

4.2. umi

umi [80] is designed for detailed and customizable occupancy scheduling but lacks stochastic modelling capabilities. It allows users to set daily, weekly, and yearly schedules for occupancy, lighting, electric appliances, DHW usage, natural ventilation, mechanical ventilation, window openings and HVAC control. Similarly, users can set customized values for occupancy densities, lighting and electric-appliance power densities and DHW peak flow for each defined schedule. Thermostat control in umi [80] is deterministic, with customizable setpoints for heating and cooling based on a user-defined hourly schedule.

Despite the possibility of defining detailed yearly schedules for each building in the simulation, the inputting process is not automized and is highly time-consuming.

Additionally, the tool offers the possibility to adopt predefined OB standard schedules, differentiated by building-use type.

4.3. CityBES

CityBES [81] offers the possibility of using predefined schedules. Default occupancy schedules are fixed and derived from predefined datasets, modelled on an hourly time step with occupancy density differentiated by building use and based on standard data. Both peak sensible- and latent-heat load related to people occupancy can be set up. Thermostat control follows fixed schedules with predefined setpoints, also on an hourly time step with standard values. However, CityBES [81] offers the possibility, working with the developers' team, of customizing the schedules of each building in the model.

HVAC control in CityBES [81] uses, by default, predefined schedules, modelled on an hourly time step with on/off control. The tool, however, for more advanced models offers the possibility of changing the HVAC control via customized schedules. Similarly, the lighting control and DHW usage follow deterministic schedules, by default, deriving from standards but customizable with the support of the developers.

4.4. OpenIDEAS

OpenIDEAS [82] integrates detailed stochastic modelling for residential buildings using the StROBe module. This allows for high-resolution time-series modelling of occupancy and various occupant-related behaviours. StROBe employs MC to model the transition between three possible occupancy states (i.e., awake at home, asleep at home and away), while the duration of each state is determined through a survival model. Then, the activities are linked to occupancy states estimating the likelihood of occurrence. This model is detailed in the work of Baetens et al. [44].

Occupancy is modelled stochastically with a 1 min resolution, based on the Belgian Time-Use Survey. Thermostat control involves stochastic setpoints for space heating, also with a 1 min resolution, and varying based on occupancy and external conditions.

DHW usage is modelled stochastically, based on occupancy patterns, with a 1 min resolution and peak usage correlating with peak occupancy times. HVAC control supports both deterministic schedules and stochastic models, with an hourly time-step and on/off control based on occupancy and temperature setpoints. Windows operation is modelled deterministically with fixed schedules, while lighting control uses stochastic models linked to occupancy, both with a 1 min resolution.

However, StROBe is limited to residential OB attributes, and, for other building-use types, the user can only adopt deterministic schedules to model occupancy and occupant-related actions, using either standard or customized profiles.

4.5. CEA

CEA [83] includes built-in standard schedules based on existing databases of Swiss and Singaporean buildings and customizable deterministic schedules for occupancy, heating, cooling, electric appliances and lighting usage, DHW consumption and electromobility. Such hourly-based schedules can be defined for typical weekdays, Saturdays and Sundays with the possibility of accounting for seasonality through the setting of monthly probabilities. Occupancy density, electric appliances and lighting-power density, and heating and cooling setpoints are customizable. To account for both latent- and sensible-heat loads of people, CEA [83] allows the definition of peak values (i.e., Watts per person) for sensible- and latent-heat exchange.

Additionally, CEA [83] provides a stochastic model for occupancy, based on a First-Order MC. Similar to CitySim [79], the stochastic model implemented in CEA [83] is based on Page et al. [43], and is therefore a First-Order Markov Chain, with transitional probabilities adjusted via a fixed mobility parameter and probabilistic determination of long absences. The limitations of this model evidenced by CitySim [79] remain valid for CEA [83].

4.6. IES-iCD

IES-iCD [105] relies on predefined hourly schedules derived from the ASHRAE 90.1 standard [14] with limited customization options for occupancy, electric appliances control, thermostat control and lighting control. Schedules are defined for typical weekdays, Saturdays and Sundays, without accounting for seasonality, differentiated by building-use type. Customization options include the selection of fixed schedules from a limited list. Occupancy density, electric appliances and lighting-power density, and temperature setpoints, can be customized.

HVAC control in IES-iCD [105] uses fixed schedules which are on/off control-based, while the DHW usage profile, also fixed, is linked to occupancy states. The tool does not support window operation.

5. Current Challenges and Potentials for Development in UBEM Tools and OB Modelling

UBEM tools have made significant advancements in recent years, yet they still hold considerable potential for further development, particularly in improving the representation of OB in simulations, where several challenges persist. These challenges limit the effectiveness, flexibility, and overall applicability of UBEM tools in realistically simulating urban-energy dynamics.

One of the primary challenges is that OB in UBEM tools is often represented through deterministic, space-based schedules. This approach reduces the adaptability of simulations and can lead to unrealistic load-curve predictions, as it fails to capture the inherent variability and stochastic nature of human behaviour. While deterministic models provide a generalized representation of OB, they do not account for the randomness and heterogeneity of people's presence and actions across different locations and times, making it difficult to accurately reflect the dynamics of energy use in various contexts.

However, integrating more advanced stochastic models into UBEM tools presents another significant challenge, due to their high computational demands, including the need for complex algorithms, large datasets, and substantial processing power. These factors make the process resource-intensive and complex, leading to longer simulation times. Some tools, such as OpenIDEAS [82], CitySim [79], and CEA [83], have already incorporated stochastic models that offer the potential to better capture the unpredictability of human behaviour. Nonetheless, these models still have limitations, and require further refinement to ensure they are suitable for different building-use types and mixed-use case studies.

Another challenge lies in the limited flexibility of current UBEM tools to accommodate dynamic and customizable schedules that reflect real-time changes or seasonal variations. Although most analysed tools, except for IES-iCD [105], allow schedule customization, the process of associating schedules with buildings can be unintuitive and time-consuming, especially when managing a large number of schedules. This can lead to inefficiencies and inaccuracies in simulations.

Interoperability between UBEM tools and other urban-modelling platforms is also a pressing issue. The current lack of common data formats and interoperability standards hinders the integration of data across platforms, limiting the potential for a comprehensive analysis of all the urban domains.

Additionally, there is a clear need for standardized methods for inputting OB aspects within UBEM tools. The absence of such standards leads to inconsistencies across studies and tools, reducing the reliability of simulation results. The challenge is not only to develop these standards, but also to ensure they are adopted and implemented across the field.

Improving OB modelling itself is another crucial area where challenges remain. One major issue is the lack of comprehensive data that accurately represent diverse contexts, such as different cultural, climatic, and socio-economic environments. Even when data are available, ensuring that they are representative enough to capture the variability of occupant behaviour across these contexts may be challenging. Leveraging diverse data sources, such as IoT devices and smart technologies, offers promising avenues for capturing real-time information on occupancy and energy-usage patterns. However, integrating these diverse data streams into UBEM models requires sophisticated data processing techniques and raises concerns about data privacy and security.

Addressing these challenges could unlock substantial potential for the development of UBEM tools.

Integrating advanced stochastic models informed by reliable and representative data can enhance the adaptability of simulations, providing more realistic load-curve predictions. The proliferation of IoT and smart devices offers an unprecedented opportunity to incorporate dynamic, real-time OB data into UBEM simulations, which could significantly improve model accuracy and responsiveness.

Moreover, advancing person-based and ABM techniques can allow for a more detailed understanding of how occupants interact with the urban environment. Techniques such as MC and survival models, which simulate probabilistic transitions between different states of occupancy and predict the duration of each state, can add a layer of randomness and variability to OB. Meanwhile, ABM can provide a detailed and dynamic representation of occupants by simulating individuals with unique behaviours and interactions. Tools like CitySim [79] are beginning to explore this with their integration of the MATSim toolkit [102], but further advancements are needed to fully capture the complexity of occupant behaviour. Co-simulation of UBEM with ABM tools presents a promising development path for addressing this complexity by simulating individual occupants with unique behaviours and interactions.

Enhancing UBEM tools to support dynamic and customizable schedules can provide users with greater flexibility and precision in their models. Implementing user-friendly interfaces, such as drag-and-drop scheduling and templates for different building types or enabling the upload and automatic assignment of schedules through external files, can empower users to create more accurate and context-specific simulations.

Improving interoperability between UBEM tools and other urban-modelling platforms is also crucial, as it offers several key advantages. One significant benefit is the ability to holistically study urban areas by combining different domains such as buildings, energy, transportation, and more. Establishing common data formats and developing middleware solutions can enable seamless communication and data sharing across platforms, facilitating coordinated strategies for urban planning, energy management, and sustainability. This approach allows for a more comprehensive analysis, leading to better-informed decisions that consider the interconnectedness of various urban systems.

Finally, rigorous validation and calibration of OB models against real-world data are essential for ensuring their accuracy and reliability. Developing a validation protocol can help ensure the robustness of OB models.

6. Conclusions

This paper provides a review of the integration of Occupant Behaviour (OB) into Urban-Building Energy Modelling (UBEM), addressing current practices, challenges, and potential improvements. The scope of this study includes an examination of existing OB modelling techniques, data sources, and UBEM tools, aiming to guide future research and tool development towards more accurate and reliable urban-energy simulations.

The review of available studies proposing advanced OB models for urban applications revealed a predominant focus on the residential sector and occupancy attributes, underlying a critical need to explore a broader range of OB attributes and building uses. This expansion is essential for developing comprehensive and nuanced models suitable for mixed-use case studies. The analysis also highlighted significant geographic regional disparities in OB research, with regions like South America, Africa, and parts of Asia, as well as tropical and arid climates, being notably under-represented. Investigating diverse contexts is crucial, given the varying climatic and cultural influences on OB.

Insights from the analysis of data sources emphasize the increasing use of modern technologies such as LBS applications and network connectivity data. These sources offer high spatial and temporal resolutions, enhancing the granularity of occupancy modelling. However, their suitability for the modelling of other OB attributes needs further investigation. Furthermore, they present challenges related to accessibility, hindered by privacy constraints. Traditional data sources, such as in situ measurements and surveys, while reliable, face limitations in scalability. A balanced approach to data integration is therefore required to gather comprehensive information about all aspects of OB in buildings.

The assessment of modelling methods reveals a shift towards advanced stochastic and agent-based models, which better capture the variability and unpredictability of human behaviour. Despite this progress, the implementation of these sophisticated models is often constrained by the current capabilities of UBEM tools, which predominantly rely on deterministic approaches. These limitations hinder the ability to account for real-time dynamics and individual heterogeneity. Furthermore, a standardized validation protocol of OB models is needed to ensure accuracy in simulation results.

Current UBEM tools offer varying degrees of customization and flexibility in OB modelling. However, they are primarily deterministic and space-based, and lack the ability to simulate variable individual movements within the urban environment. Future development should focus on integrating person-based occupant modelling, co-simulation with agent-based frameworks, and the adoption of dynamic, easily customizable schedules. Enhancing interoperability between UBEM tools and other urban-modelling platforms is also crucial for providing a comprehensive view of urban dynamics.

In conclusion, while significant advancements have been made in integrating OB into UBEM, further progress is needed to incorporate diverse data sources, improve modelling techniques, and enhance tool capabilities. Addressing these challenges will enable more accurate and reliable urban-energy simulations, contributing to more effective and sustainable urban planning and energy management strategies. Moreover, fostering collaboration between researchers and tool developers is essential to bridge the gap between advanced OB modelling and practical UBEM applications, ultimately facilitating the creation of more occupant-centric urban-energy systems.

Funding: The study was developed partially within the framework of the project URBEM (PRIN, code: 2020ZWKXKE), which has received funding from the Italian Ministry of University and Research (MUR), and partially under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.3—Call for tender No.1561 of 11 October 2022 of Ministero dell'Università e della Ricerca (MUR), funded by the European Union–NextGenerationEU. Award Number: Project code PE0000021, Concession Decree No.1561 of 11 October 2022 adopted by Ministero dell'Università e della Ricerca (MUR), CUPC93C22005230007, according to attachment E of Decree No.1561/2022, Project title "Network 4 Energy Sustainable Transition—NEST".

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

References

1. United Nations. *The Sustainable Development Goals Report*; United Nations: New York, NY, USA, 2024.
2. Official Journal of the European Union. *DIRECTIVE (EU) 2024/1275 of the European Parliament and of the Council of 24 April 2024 on the Energy Performance of Buildings*; Official Journal of the European Union: Luxembourg, 2024; Available online: http://data.europa.eu/eli/dir/2024/1275/oj (accessed on 27 July 2024).
3. Ferrando, M.; Causone, F.; Hong, T.; Chen, Y. Urban building energy modeling (UBEM) tools: A state-of-the-art review of bottom-up physics-based approaches. *Sustain. Cities Soc.* **2020**, *62*, 102408. [CrossRef]
4. Carnieletto, L.; Ferrando, M.; Teso, L.; Sun, K.; Zhang, W.; Causone, F.; Romagnoni, P.; Zarrella, A.; Hong, T. Italian prototype building models for urban scale building performance simulation. *Build. Environ.* **2021**, *192*, 107590. [CrossRef]
5. Ferrando, M.; Causone, F. An overview of urban building energy modelling (UBEM) tools. In Proceedings of the Building Simulation Conference Proceedings, International Building Performance Simulation Association, Rome, Italy, 2–4 September 2019; pp. 3452–3459. [CrossRef]
6. Reinhart, C.F.; Davila, C.C. Urban building energy modeling—A review of a nascent field. *Build. Environ.* **2016**, *97*, 196–202. [CrossRef]
7. Hong, T.; Chen, Y.; Luo, X.; Luo, N.; Lee, S.H. Ten questions on urban building energy modeling. *Build. Environ.* **2019**, *168*, 106508. [CrossRef]
8. Carlucci, S.; Causone, F.; Biandrate, S.; Ferrando, M.; Moazami, A.; Erba, S. On the impact of stochastic modeling of occupant behavior on the energy use of office buildings. *Energy Build.* **2021**, *246*, 111049. [CrossRef]
9. Banfi, A.; Ferrando, M.; Causone, F. Occupant behaviour (OB) modelling to support urban building energy simulation: An overview. In Proceedings of the 2023 Building Simulation Conference Proceedings, International Building Performance Simulation Association, Shanghai, China, 4–6 September 2023; pp. 3003–3010.
10. Carlucci, S.; De Simone, M.; Firth, S.K.; Kjærgaard, M.B.; Markovic, R.; Rahaman, M.S.; Annaqeeb, M.K.; Biandrate, S.; Das, A.; Dziedzic, J.W.; et al. Modeling occupant behavior in buildings. *Build. Environ.* **2020**, *174*, 106768. [CrossRef]
11. Yan, D.; Hong, T.; Dong, B.; Mahdavi, A.; D'oca, S.; Gaetani, I.; Feng, X. IEA EBC Annex 66: Definition and simulation of occupant behavior in buildings. *Energy Build.* **2017**, *156*, 258–270. [CrossRef]

12. O'Brien, W.; Wagner, A.; Schweiker, M.; Mahdavi, A.; Day, J.; Kjærgaard, M.B.; Carlucci, S.; Dong, B.; Tahmasebi, F.; Yan, D.; et al. Introducing IEA EBC annex 79: Key challenges and opportunities in the field of occupant-centric building design and operation. *Build. Environ.* **2020**, *178*, 106738. [CrossRef]
13. Dong, B.; Liu, Y.; Mu, W.; Jiang, Z.; Pandey, P.; Hong, T.; Olesen, B.; Lawrence, T.; O'neil, Z.; Andrews, C.; et al. A Global Building Occupant Behavior Database. *Sci. Data* **2022**, *9*, 369. [CrossRef]
14. ASHRAE. ASHRAE 90.1-2022 (I-P) Standard 90.1-2022 (I-P Edition)—Energy Standard for Sites and Buildings Except Low-Rise Residential Buildings (ANSI Approved; IES Co-sponsored). 2022. Available online: https://www.ashrae.org/technical-resources/bookstore/standard-90-1 (accessed on 27 July 2024).
15. Deru, M.; Field, K.; Studer, D.; Benne, K.; Griffith, B.; Torcellini, P.; Liu, B.; Halverson, M.; Winiarski, D.; Rosenberg, M.; et al. Commercial Reference Building Models of the National Building Stock. 2011. Available online: http://www.osti.gov/bridge (accessed on 1 August 2024).
16. Ente Italiano di Normazione, UNI 13790—Prestazione energetica degli edifici—Calcolo del fabbisogno di energia per il riscaldamento e il raffrescamento. 2018. Available online: https://store.uni.com/uni-en-iso-13790-2008 (accessed on 1 August 2024).
17. NCM. National Calculation Methodology (NCM) modelling guide (for buildings other than dwellings in England). 2021. Available online: https://www.uk-ncm.org.uk/filelibrary/NCM_Modelling_Guide_2021_Edition_England_15Dec2021.pdf (accessed on 1 August 2024).
18. National Research Council of Canada (NRC). *NECB, National Energy Code of Canada for Buildings*; National Research Council of Canada (NRC): Ottawa, ON, Canada, 2017.
19. Happle, G.; Fonseca, J.A.; Schlueter, A. A review on occupant behavior in urban building energy models. *Energy Build.* **2018**, *174*, 276–292. [CrossRef]
20. Malik, J.; Azar, E.; Mahdavi, A.; Hong, T. A level-of-details framework for representing occupant behavior in agent-based models. *Autom. Constr.* **2022**, *139*, 104290. [CrossRef]
21. Ferrando, M.; Ferroni, S.; Pelle, M.; Tatti, A.; Erba, S.; Shi, X.; Causone, F. UBEM's archetypes improvement via data-driven occupant-related schedules randomly distributed and their impact assessment. *Sustain. Cities Soc.* **2022**, *87*, 104164. [CrossRef]
22. Mohammadi, N.; Taylor, J.E. Urban infrastructure-mobility energy flux. *Energy* **2017**, *140*, 716–728. [CrossRef]
23. Salim, F.D.; Dong, B.; Ouf, M.; Wang, Q.; Pigliautile, I.; Kang, X.; Hong, T.; Wu, W.; Liu, Y.; Rumi, S.K.; et al. Modelling urban-scale occupant behaviour, mobility, and energy in buildings: A survey. *Build. Environ.* **2020**, *183*, 106964. [CrossRef]
24. Dabirian, S.; Panchabikesan, K.; Eicker, U. Occupant-centric urban building energy modeling: Approaches, inputs, and data sources—A review. *Energy Build.* **2022**, *257*. [CrossRef]
25. 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*. [CrossRef]
26. Schweiker, M.; Carlucci, S.; Andersen, R.K.; Dong, B.; O'Brien, W. Occupancy and occupants' actions. In *Exploring Occupant Behavior in Buildings: Methods and Challenges*; Springer International Publishing: Berlin/Heidelberg, Germany, 2017; pp. 7–38. [CrossRef]
27. Hong, T.; Yan, D.; D'Oca, S.; Chen, C.-F. Ten questions concerning occupant behavior in buildings: The big picture. *Build. Environ.* **2016**, *114*, 518–530. [CrossRef]
28. 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. [CrossRef]
29. Balasubramanian, S.; Balachandra, P. Effectiveness of demand response in achieving supply-demand matching in a renewables dominated electricity system: A modelling approach. *Renew. Sustain. Energy Rev.* **2021**, *147*, 111245. [CrossRef]
30. Strbac, G. Demand side management: Benefits and challenges. *Energy Policy* **2008**, *36*, 4419–4426. [CrossRef]
31. Dashdorj, Z.; Sobolevsky, S.; Lee, S.; Ratti, C. Deriving human activity from geo-located data by ontological and statistical reasoning. *Knowl.-Based Syst.* **2018**, *143*, 225–235. [CrossRef]
32. Shahroz, M.; Ahmad, F.; Younis, M.S.; Ahmad, N.; Boulos, M.N.K.; Vinuesa, R.; Qadir, J. COVID-19 digital contact tracing applications and techniques: A review post initial deployments. *Transp. Eng.* **2021**, *5*, 100072. [CrossRef]
33. Su, R.; McBride, E.C.; Goulias, K.G. Pattern recognition of daily activity patterns using human mobility motifs and sequence analysis. *Transp. Res. Part C Emerg. Technol.* **2020**, *120*, 102796. [CrossRef]
34. Barbour, E.; Davila, C.C.; Gupta, S.; Reinhart, C.; Kaur, J.; González, M.C. Planning for sustainable cities by estimating building occupancy with mobile phones. *Nat. Commun.* **2019**, *10*, 3736. [CrossRef]
35. Happle, G.; Fonseca, J.A.; Schlueter, A. Context-specific urban occupancy modeling using location-based services data. *Build. Environ.* **2020**, *175*, 106803. [CrossRef]
36. Dong, B.; Liu, Y.; Fontenot, H.; Ouf, M.; Osman, M.; Chong, A.; Qin, S.; Salim, F.; Xue, H.; Yan, D.; et al. Occupant behavior modeling methods for resilient building design, operation and policy at urban scale: A review. *Appl. Energy* **2021**, *293*, 116856. [CrossRef]
37. Ren, J.; Zhou, X.; Jin, X.; Ye, Y.; Causone, F.; Ferrando, M.; Li, P.; Shi, X. A systematic review of occupancy pattern in urban building energy modeling: From urban to building-scale. *J. Build. Eng.* **2024**, *95*, 110307. [CrossRef]
38. Doma, A.; Ouf, M. Modelling occupant behaviour for urban scale simulation: Review of available approaches and tools. *Build. Simul.* **2022**, *16*, 169–184. [CrossRef]

39. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, 71. [CrossRef]
40. Richardson, I.; Thomson, M.; Infield, D. A high-resolution domestic building occupancy model for energy demand simulations. *Energy Build.* **2008**, *40*, 1560–1566. [CrossRef]
41. Richardson, I.; Thomson, M.; Infield, D.; Delahunty, A. Domestic lighting: A high-resolution energy demand model. *Energy Build.* **2009**, *41*, 781–789. [CrossRef]
42. Richardson, I.; Thomson, M.; Infield, D.; Clifford, C. Domestic electricity use: A high-resolution energy demand model. *Energy Build.* **2010**, *42*, 1878–1887. [CrossRef]
43. 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]
44. Baetens, R.; Saelens, D. Modelling uncertainty in district energy simulations by stochastic residential occupant behaviour. *J. Build. Perform. Simul.* **2015**, *9*, 431–447. [CrossRef]
45. Rouleau, J.; Ramallo-González, A.P.; Gosselin, L.; Blanchet, P.; Natarajan, S. A unified probabilistic model for predicting occupancy, domestic hot water use and electricity use in residential buildings. *Energy Build.* **2019**, *202*, 109375. [CrossRef]
46. Buttitta, G.; Turner, W.J.; Neu, O.; Finn, D.P. Development of occupancy-integrated archetypes: Use of data mining clustering techniques to embed occupant behaviour profiles in archetypes. *Energy Build.* **2019**, *198*, 84–99. [CrossRef]
47. Auffhammer, M.; Mansur, E.T. Measuring climatic impacts on energy consumption: A review of the empirical literature. *Energy Econ.* **2014**, *46*, 522–530. [CrossRef]
48. Stokes, M.; Rylatt, M.; Lomas, K. A simple model of domestic lighting demand. *Energy Build.* **2004**, *36*, 103–116. [CrossRef]
49. Rakha, T.; Rose, C.M.; Reinhart, C.F. A Framework for Modelling Occupancy Schedules and Local trips based on Activity Based Surveys. In Proceedings of the 2014 ASHRAE/IBPSA-USA Building Simulation Conference, Atlanta, GA, USA, 10–12 September 2014.
50. Azar, E.; Nikolopoulou, C.; Papadopoulos, S. Integrating and optimizing metrics of sustainable building performance using human-focused agent-based modeling. *Appl. Energy* **2016**, *183*, 926–937. [CrossRef]
51. Parker, J.; Hardy, A.; Glew, D.; Gorse, C. A methodology for creating building energy model occupancy schedules using personal location metadata. *Energy Build.* **2017**, *150*, 211–223. [CrossRef]
52. 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]
53. Tröndle, T.; Choudhary, R. Occupancy based thermal energy modelling in the urban residential sector. In *WIT Transactions on Ecology and the Environment*; WIT Press: Southampton, UK, 2017; pp. 31–44. [CrossRef]
54. El Kontar, R.; Rakha, T. Profiling Occupancy Patterns to Calibrate Urban Building Energy Models (UBEMs) Using Measured Data Clustering. *Technol. Archit. Des.* **2018**, *2*, 206–217. [CrossRef]
55. Berres, A.; Im, P.; Kurte, K.; Allen-Dumas, M.; Thakur, G.; Sanyal, J. A Mobility-Driven Approach to Modeling Building Energy. In Proceedings of the 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 9–12 December 2019; pp. 3887–3895. [CrossRef]
56. Koupaei, D.M.; Hashemi, F.; Tabard-Fortecoëf, V.; Passe, U. Development Of A Modeling Framework for Refined Residential Occupancy Schedules in An Urban Energy Model. In Proceedings of the Building Simulation 2019: 16th Conference of IBPSA, IBPSA, Rome, Italy, 2–4 September 2020; pp. 3377–3384. [CrossRef]
57. Wang, C.; Wu, Y.; Shi, X.; Li, Y.; Zhu, S.; Jin, X.; Zhou, X. Dynamic occupant density models of commercial buildings for urban energy simulation. *Build. Environ.* **2019**, *169*, 106549. [CrossRef]
58. Mahmood, I.; Ain, Q.T.; Nasir, H.A.; Javed, F.; A Aguado, J. A hierarchical multi-resolution agent-based modeling and simulation framework for household electricity demand profile. *Simulation* **2020**, *96*, 655–678. [CrossRef]
59. Wu, W.; Dong, B.; Wang, Q.; Kong, M.; Yan, D.; An, J.; Liu, Y. A novel mobility-based approach to derive urban-scale building occupant profiles and analyze impacts on building energy consumption. *Appl. Energy* **2020**, *278*, 115656. [CrossRef]
60. Mosteiro-Romero, M.; Hischier, I.; Fonseca, J.A.; Schlueter, A. A novel population-based occupancy modeling approach for district-scale simulations compared to standard-based methods. *Build. Environ.* **2020**, *181*, 107084. [CrossRef]
61. Ueno, T.; Meier, A. A method to generate heating and cooling schedules based on data from connected thermostats. *Energy Build.* **2020**, *228*, 110423. [CrossRef]
62. Vellei, M.; Martinez, S.; Le Dréau, J. Agent-based stochastic model of thermostat adjustments: A demand response application. *Energy Build.* **2021**, *238*, 110846. [CrossRef]
63. Panchabikesan, K.; Haghighat, F.; El Mankibi, M. Data driven occupancy information for energy simulation and energy use assessment in residential buildings. *Energy* **2020**, *218*, 119539. [CrossRef]
64. Kang, X.; Yan, D.; An, J.; Jin, Y.; Sun, H. Typical weekly occupancy profiles in non-residential buildings based on mobile positioning data. *Energy Build.* **2021**, *250*, 111264. [CrossRef]
65. Schumann, M.; Reynaud, Q.; Sabouret, N.; Sempé, F.; Haradji, Y.; Charrier, B.; Albouys, J.; Inard, C. Multi-Agent based simulation of human activity for building and urban scale assessment of residential load curves and energy use. In Proceedings of the Building Simulation Conference Proceedings, International Building Performance Simulation Association, Bruges Belgium, 1–3 September 2022; pp. 3457–3464. [CrossRef]

66. Chen, J.; Adhikari, R.; Wilson, E.; Robertson, J.; Fontanini, A.; Polly, B.; Olawale, O. Stochastic simulation of occupant-driven energy use in a bottom-up residential building stock model. *Appl. Energy* **2022**, *325*, 119890. [CrossRef]
67. Koupaei, D.M.; Cetin, K.S.; Passe, U. Stochastic residential occupancy schedules based on the American Time-Use Survey. *Sci. Technol. Built Environ.* **2022**, *28*, 776–790. [CrossRef]
68. Wu, Y.; Zhou, X.; Qian, M.; Jin, Y.; Sun, H.; Yan, D. Novel approach to typical air-conditioning behavior pattern extraction based on large-scale VRF system online monitoring data. *J. Build. Eng.* **2023**, *69*, 106243. [CrossRef]
69. Liu, X.; Hu, S.; Yan, D. A statistical quantitative analysis of the correlations between socio-demographic characteristics and household occupancy patterns in residential buildings in China. *Energy Build.* **2023**, *284*, 112842. [CrossRef]
70. Zhu, M.; Pan, Y.; Wu, Z.; Huang, Z.; Kosonen, R. Modelling method of inter-building movement for campus-scale occupancy simulation: A case study. *Build. Simul.* **2022**, *16*, 461–481. [CrossRef] [PubMed]
71. Jung, W.; Wang, Z.; Hong, T.; Jazizadeh, F. Smart thermostat data-driven U.S. residential occupancy schedules and development of a U.S. residential occupancy schedule simulator. *Build. Environ.* **2023**, *243*, 110628. [CrossRef]
72. Yu, Z.; Song, C.; Liu, Y.; Wang, D.; Li, B. A bottom-up approach for community load prediction based on multi-agent model. *Sustain. Cities Soc.* **2023**, *97*, 104774. [CrossRef]
73. Sood, D.; Alhindawi, I.; Ali, U.; McGrath, J.A.; Byrne, M.A.; Finn, D.; O'donnell, J. Simulation-based evaluation of occupancy on energy consumption of multi-scale residential building archetypes. *J. Build. Eng.* **2023**, *75*, 106872. [CrossRef]
74. Zhou, W.; Chen, Y.; Liu, Y.; Yu, Z.; Wang, D.; Zhou, W.; Chen, Y.; Liu, Y.; Yu, Z.; Wang, D.; et al. A new method for dynamic load simulation of urban building complexes' heating considering building level occupancy based on shoebox algorithm. *Build. Environ.* **2024**, *261*, 111710. [CrossRef]
75. Doma, A.; Prajapati, S.N.; Ouf, M.M. Developing a residential occupancy schedule generator based on smart thermostat data. *Build. Environ.* **2024**, *261*, 111713. [CrossRef]
76. Liu, Z.; Dou, Z.; Chen, H.; Zhang, C.; Wang, S.; Wu, Y.; Liu, X.; Yan, D. Exploring the impacts of heterogeneity and stochasticity in air-conditioning behavior on urban building energy models. *Sustain. Cities Soc.* **2024**, *103*, 105285. [CrossRef]
77. Abolhassani, S.S.; Zandifar, A.; Ghourchian, N.; Amayri, M.; Bouguila, N.; Eicker, U. Occupant counting model development for urban building energy modeling using commercial off-the-shelf Wi-Fi sensing technology. *Build. Environ.* **2024**, *258*, 111548. [CrossRef]
78. Köppen, W.P. Klassification der Klimate nach Temperatur, Niederschlag und Jahreslauf. 1918. Available online: https://koeppen-geiger.vu-wien.ac.at/pdf/Koppen_1918.pdf (accessed on 27 July 2024).
79. Robinson, D.; Perez, D.; Rasheed, A. CITYSIM: Comprehensive Micro-Simulation of Resource Flows for Sustainable Urban Planning. 2009. Available online: https://www.researchgate.net/publication/43652081 (accessed on 25 July 2024).
80. Reinhart, C.F.; Dogan, T.; Jakubiec, A.; Rakha, T.; Sang, A. Umi-an urban simulation environment for building energy use, daylighting and walkability. In Proceedings of the 13th Conference of International Building Performance Simulation Association, Chambéry, France, 26–28 August 2013.
81. Chen, Y.; Hong, T.; Piette, M.A. Automatic generation and simulation of urban building energy models based on city datasets for city-scale building retrofit analysis. *Appl. Energy* **2017**, *205*, 323–335. [CrossRef]
82. Baetens, R.; De Coninck, R.; Jorissen, F.; Picard, D.; Helsen, L.; Saelens, D. OpenIDEAS—An Open Framework for integrated District Energy Simulations. In Proceedings of the 2015 Building Simulation Conference, Bolzano, Italy, 4–6 February 2015; pp. 347–354.
83. Fonseca, J.A.; Nguyen, T.-A.; Schlueter, A.; Marechal, F. City Energy Analyst (CEA): Integrated framework for analysis and optimization of building energy systems in neighborhoods and city districts. *Energy Build.* **2016**, *113*, 202–226. [CrossRef]
84. IESVE. Available online: http://www.iesve.com/ (accessed on 7 April 2021).
85. Concordia University, Tool4Cities. (n.d.) Available online: https://www.concordia.ca/research/cities-institute/initiatives/tools4cities.html (accessed on 27 July 2024).
86. Wilson, E.; Christensen, C.; Horowitz, S.; Horsey, H. A High-Granularity Approach to Modeling Energy Consumption and Savings Potential in the U.S. Residential Building Stock: Preprint. 2022. Available online: www.nrel.gov/publications (accessed on 25 July 2024).
87. Mattsson, S.E.; Elmqvist, H. Modelica—An International Effort to Design the Next Generation Modeling Language. *IFAC Proc. Vol.* **1997**, *30*, 151–155. [CrossRef]
88. EnergyPlusTM. 2017. Available online: https://www.osti.gov/biblio/1395882 (accessed on 25 July 2024).
89. Yan, D.; Xia, J.; Tang, W.; Song, F.; Zhang, X.; Jiang, Y. DeST—An integrated building simulation toolkit Part I: Fundamentals. *Build. Simul.* **2008**, *1*, 95–110. [CrossRef]
90. Choi, H.; Um, C.Y.; Kang, K.; Kim, H.; Kim, T. Review of vision-based occupant information sensing systems for occupant-centric control. *Build. Environ.* **2021**, *203*, 108064. [CrossRef]
91. Causone, F.; Carlucci, S.; Ferrando, M.; Marchenko, A.; Erba, S. A data-driven procedure to model occupancy and occupant-related electric load profiles in residential buildings for energy simulation. *Energy Build.* **2019**, *202*, 109342. [CrossRef]
92. Ecobee. Donate-Your-Data (DYD). (n.d.) Available online: https://www.ecobee.com/en-us/donate-your-data/ (accessed on 29 July 2024).
93. Osman, M.; Ouf, M. A comprehensive review of time use surveys in modelling occupant presence and behavior: Data, methods, and applications. *Build. Environ.* **2021**, *196*, 107785. [CrossRef]

94. Centre for Time Use Research (CTUR). Multinational Time Use Study. (n.d.) Available online: https://www.timeuse.org/mtus (accessed on 14 December 2023).
95. Eurostat. Harmonized European TIme Use Survey (HETUS). (n.d.) Available online: https://ec.europa.eu/eurostat/web/microdata/harmonised-european-time-use-surveys (accessed on 29 July 2024).
96. Google Popular Times, Wait Times, and Visit Duration. (n.d.) Available online: https://support.google.com/business/answer/6263531?hl=en (accessed on 2 January 2023).
97. Abbasabadi, N.; Ashayeri, M. From Tweets to Energy Trends (TwEn): An exploratory framework for machine learning-based forecasting of urban-scale energy behavior leveraging social media data. *Energy Build.* **2024**, *317*, 114440. [CrossRef]
98. Chen, J.; Ahn, C. Assessing occupants' energy load variation through existing wireless network infrastructure in commercial and educational buildings. *Energy Build.* **2014**, *82*, 540–549. [CrossRef]
99. Ferrando, M.; Banfi, A.; Causone, F. Changes in energy use profiles derived from electricity smart meter readings of residential buildings in Milan before, during and after the COVID-19 main lockdown. *Sustain. Cities Soc.* **2023**, *99*, 104876. [CrossRef]
100. Tweedie, R.L. Markov chains: Structure and applications. In *Handbook of Statistics*; Elsevier: Amsterdam, The Netherlands, 2001; pp. 817–851. [CrossRef]
101. Malik, J.; Mahdavi, A.; Azar, E.; Putra, H.C.; Berger, C.; Andrews, C.; Hong, T. Ten questions concerning agent-based modeling of occupant behavior for energy and environmental performance of buildings. *Build. Environ.* **2022**, *217*, 109016. [CrossRef]
102. MATSim. Multi-Agent Transport Simulation. (n.d.) Available online: https://matsim.org/ (accessed on 29 July 2024).
103. Jiang, S.; Yang, Y.; Gupta, S.; Veneziano, D.; Athavale, S.; González, M.C. The TimeGeo modeling framework for urban mobility without travel surveys. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E5370–E5378. [CrossRef]
104. TRANSIMS. TRansportation ANalysis and SIMulation System. (n.d.) Available online: https://sourceforge.net/projects/transimsstudio/ (accessed on 29 July 2024).
105. Integrated Environmental Solutions, IES Virtual Environment (IES-VE). (n.d.) Available online: https://www.iesve.com/software/virtual-environment (accessed on 27 July 2024).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.