

Review

How to Conduct Human-Centric Building Design? A Review of Occupant Modeling Methods and Applications

Rui Sun ^{1,2,3} , Cheng Sun ^{1,2} , Rajendra S. Adhikari ^{3,*} , Dagang Qu ^{1,2,*} and Claudio Del Pero ³ 

¹ School of Architecture and Design, Harbin Institute of Technology, Harbin 150001, China; rui.sun@mail.polimi.it (R.S.); suncheng@hit.edu.cn (C.S.)

² Key Laboratory of Cold Region Urban and Rural Human Settlement Environment Science and Technology, Ministry of Industry and Information Technology, Harbin 150001, China

³ Department of Architecture, Built Environment and Construction Engineering, Politecnico di Milano, 20133 Milan, Italy; claudio.delpero@polimi.it

* Correspondence: rajendra.adhikari@polimi.it (R.S.A.); qudagang@hit.edu.cn (D.Q.)

Abstract

Occupant modeling has emerged as a critical component in human-centric building design and operation, offering detailed insights into energy performance, comfort optimization, and behavior-driven control strategies. This study systematically examines occupant modeling (OM) in building design through a review of 312 articles, highlighting critical gaps between theoretical frameworks and real-world applications. Key dimensions of occupant modeling, including methodological classification, data frameworks, application scenarios and model selection strategies, are examined. The interpretability, advantages and disadvantages of 5 modeling methods are demonstrated, and the tools, algorithms and applications are analyzed. In addition, common input, output and application scenarios are sorted out and the data streams are presented. Results have shown that hybrid models represent breakthroughs but require validation beyond idealized scenarios. Meanwhile, with 88.7% of output derived from simulated results, risking self-reinforcing biases despite empirical inputs. Standardized protocols for model validation and hybrid modeling frameworks are urgently needed. To support model selection, a decision-oriented framework is proposed, integrating modeling goals, data characteristics, behavioral complexity, and platform interoperability. Future priorities include merging high explanatory methods with powerful predictive methods, advancing BIM-IoT symbiosis for adaptive digital twin, expanding to interdisciplinary projects, and establishing ethical data governance to align technical advancements with equitable, occupant-centric design.

Keywords: occupant modeling; building performance simulation; building design



Academic Editor: Koen Steemers

Received: 15 September 2025

Revised: 20 October 2025

Accepted: 13 November 2025

Published: 15 November 2025

Citation: Sun, R.; Sun, C.; Adhikari, R.S.; Qu, D.; Del Pero, C. How to Conduct Human-Centric Building Design? A Review of Occupant Modeling Methods and Applications. *Buildings* **2025**, *15*, 4117. <https://doi.org/10.3390/buildings15224117>

Copyright: © 2025 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

For architects, the concept of human-centric design is crucial for building design as the feedback of occupants decides whether their work is successful or not. Buildings are designed to provide comfortable, healthy, usable and safe spaces for a variety of occupants, but the interaction between people and buildings remains one of the least mature aspects of building science. Although occupant behavior research began 40 years ago, there has been a surge in research related to occupants in buildings in the past decade, but the design and operation of buildings are still based on outdated or simplistic assumptions about occupants, which are increasingly proven to be misleading [1]. There is evidence that buildings designed according to regulations may not be comfortable and efficient and

difficult for occupants to adjust the built environment in terms of their preferences and presence [2]. Meanwhile, occupants in the built environment are a central driver of dynamic energy consumption, comfort and safety. Statistically, occupants' uncertainty could lead to building energy consumption prediction deviations of up to 30–300% [3,4] and directly affect the effectiveness of building design. Traditional Building Performance Simulation (BPS) tools (e.g., EnergyPlus, TRNSYS) rely on static schedules with homogeneous population assumptions, which makes it difficult to capture heterogeneity in real scenarios. For example, fixed lighting usage schedules may underestimate actual energy consumption by up to 42% [5]; air conditioning setpoint deviations lead to cooling energy consumption errors of more than 60% [6].

Therefore, the construction of high-precision and scalable occupant models has become an urgent need in the field of building design. The representative node of occupant behavior research was carried out in the framework of IEA EBC Annex 66, which formally defines experimental research methods, modeling and model verification, and occupant simulation methods [7]. IEA-EBC Annex 79 advocates for integrating occupancy and occupant behavior into both design processes and building operation and maintenance to enhance energy performance and occupant comfort, emphasizing the development of new scientific knowledge on adaptive occupant behavior driven by multiple interdependent indoor environmental parameters. A variety of occupant models have been developed, and scholars have classified them from various perspectives. However, despite two decades of accelerated research on occupant behavior modeling (2004–2024), the translation of theoretical advances into architectural design practice remains critically fragmented. Seminal reviews have laid essential foundations: The IEA-EBC Annex 66 [7] and many scholars [8,9] have tried dividing them into different categories a few years ago. O'Brien and Gunay first deconstructed contextual factors (e.g., control accessibility, social constraints) shaping adaptive behaviors, yet their framework lacked quantification of behavioral impacts [10]. Hong et al. responded by establishing the DNAs ontology (standardizing Drivers–Needs–Actions-system interactions) and cataloging modeling advances, but these studies remained siloed in simulation domains, overlooking spatial design applications [11]. Concurrently, Delzendeh et al. and Jia et al. quantified the “performance gap” caused by static occupant assumptions, while exposing the neglect of psychological and cultural factors in energy tools [12,13]. The second wave confronted modeling-technical gaps: Abraham et al. pioneered Agent-Based Modeling (ABM) for dynamic behavior simulation, yet acknowledged its detachment from design workflows [14]. Jin et al. and Kanthila et al. systematized occupancy detection and prediction techniques but identified data scarcity and privacy barriers [4,15]. Uddin et al. pointed out that 90% of models relied on Global North office data [16]. Moreover, Azar et al. revealed a pivotal disconnect: despite proposing “occupant-centric design” (OCD) principles, fewer than 15% of behavior studies informed architectural spatial strategies [17]. Recent reviews intensified calls for integration: Ding et al. mapped occupancy prediction to control optimization but offered no design guidelines [18]. Ahmed et al. synthesized behavior modeling workflows and confirmed limited interoperability [19]. Wang et al. explicitly linked behavior models to building optimization but found none of the cases validating frameworks in real design projects [20], while Caballero-Pena et al. and Ebuy et al. highlighted voids in non-Western behavioral data and socio-technical coupling [21,22].

It can be seen that a paradigm shift is demanded for these gaps: (1) spatial-design irrelevance, (2) lack of practicality, and (3) tool-chain disintegration. For architectural designers, it is still a challenging problem to know what kind of scenarios these various models satisfy and how to create new models to serve architectural design. Therefore, it is essential to review the literature on occupant models that can be used in architectural

design, analyzing the types, data and applications of existing models at a finer granularity to provide support for designers, and make suggestions for model selection and application in the architectural design phase. The review focuses more on an occupant-driven spatial design framework than just model improvements. This research intends to provide a path for buildings that dynamically balance energy, comfort and cultural factors by transforming occupants from passive variables to active co-drivers of design.

Based on a systematic review of existing research on occupant models for building design, this paper aims to address the following scientific questions and subsequently proposes future research directions accordingly:

- Geographical distribution and targeting building typologies: What patterns exist in their spatial deployment and architectural applications?
- Typology: What are the prevalent categories, along with their respective strengths and limitations?
- Data framework: What constitutes the predominant input-output data structures, and to what extent is real-world data integrated into current practices?
- Application scenarios: What is the prevalence ratio between collective models and individual models, and how do their application scenarios differ?
- How can architects choose the right model and apply it in practice? And how to build a better iterative cycle?

This structured inquiry seeks to establish an evidence-based foundation for advancing occupant modeling in building performance research, while identifying critical gaps between theoretical frameworks and operational realities. In addition, by answering the questions above, we hope to establish an occupant-driven spatial design framework to guide architects. To ensure conceptual consistency across studies, this review adopts the following terminology: occupancy refers to the presence or absence of people in a given space; occupant behavior denotes observable actions and interactions with the environment (e.g., lighting, window, or thermostat adjustments), and occupant modeling describes the computational or conceptual frameworks used to represent and predict such behaviors. These terms are used consistently throughout the paper.

2. Materials and Methods

2.1. Literature Search

A comprehensive literature search was performed in the Web of Science database using a combination of controlled terms and free-text keywords related to occupant behavior and architectural design. The main query string was: TS = (“occupant modeling” OR “occupant model” OR “occupancy model”) AND (“architectural design” OR “building design”). Synonyms such as occupancy, presence, and behavior were included through Boolean operators to ensure comprehensive coverage. As depicted in Figure 1, the initial search found 4148 records with no time limitation (last updated 25 March 2025). Before screening, 2101 records were automatically excluded as they were identified as non-architectural studies (e.g., robotics, software engineering), and 38 records were removed for other reasons (e.g., editorials, book reviews, extended abstracts, and non-peer-reviewed materials, non-English language, inaccessible formats). A total of 1500 records underwent abstract screening. Among these, 728 records were excluded as irrelevant to occupant behavior research. The remaining 772 full-text reports were retrieved for eligibility assessment. During the full-text screening stage, 322 studies were retained based on the following predefined inclusion criteria: (1) the study explicitly addressed occupant behavior modeling methods or their applications in architectural design, or (2) the study was a review article synthesizing progress in this field. Studies were excluded for the following reasons: 189 studies focused exclusively on mechanical systems without any architectural linkage; 26 studies addressed

human–computer interaction outside the built environment context; 235 studies did not integrate occupant modeling with architectural design objectives. These criteria ensured that only studies with direct relevance to occupant modeling within architectural workflows were included in the synthesis. Ultimately, 312 studies were retained for analysis, as 10 studies were omitted due to multiple publications from individual projects. It should be noted that the selection procedure was made to the best of the authors’ knowledge, and other search methods may contribute to different statistics. The initial screening was performed by the first author based on predefined inclusion and exclusion criteria, and subsequently reviewed by other co-authors to ensure alignment with the review protocol. Although independent parallel screening was not conducted and inter-rater reliability was therefore not assessed, all inclusion decisions were verified through cross-checking to reduce potential bias.

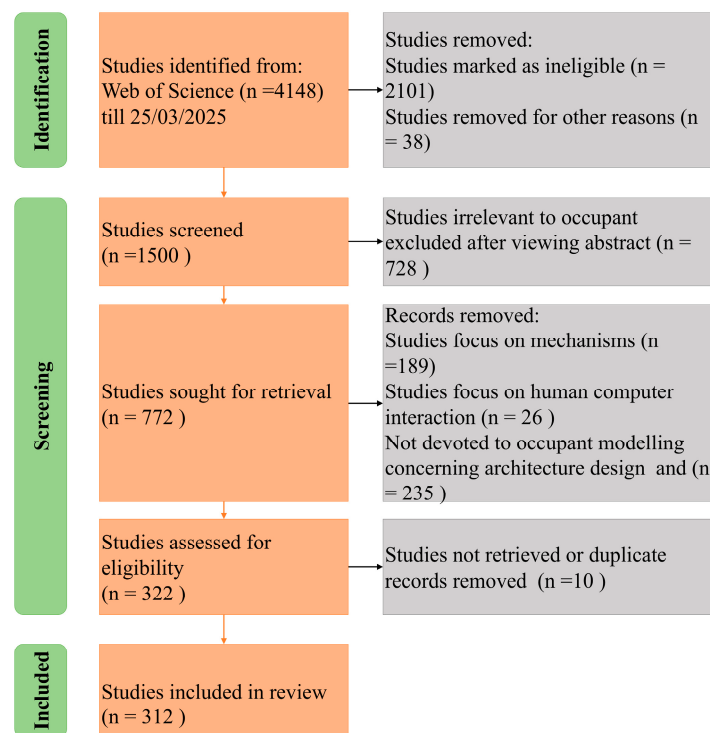


Figure 1. Procedure of the literature search.

2.2. Overview of the Existing Review Articles

The trend of publications is shown in Figure 2. The analyzed studies span 25 years (till 25 March 2025), revealing distinct evolutionary patterns in occupant modeling research in architectural concepts. Annual publication output demonstrates three phases: (1) Incubation phase (2001–2014): gradual growth from 1 to 11 articles annually, reflecting early conceptual explorations of behavior-model integration in architectural workflows. (2) Acceleration phase (2015–2021): exponential growth trajectory, with publications surging from 16 (2016) to 27 (2022), driven by advancements in IoT sensing, machine learning, and regulatory demands for performance-based design. (3) Consolidation phase (2022–present): peak output in 2022 followed by a modest decline (maintaining > 20 annual publications), suggesting maturation of core methodologies and shifting focus toward hybrid modeling approaches. These three phases were identified based on observed changes in publication volume and thematic development across the reviewed studies.

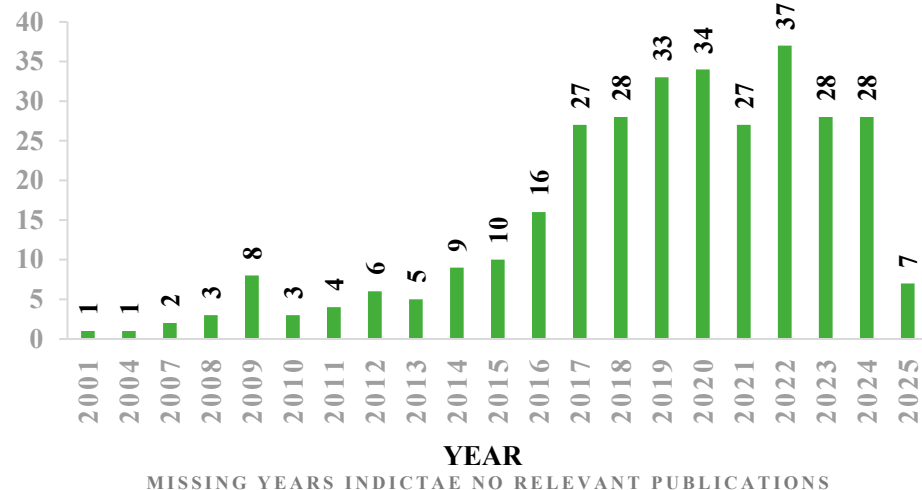


Figure 2. Trend of publications.

This trend aligns with the adoption of digital twin technologies in the Architecture, Engineering, and Construction industries and growing emphasis on occupant-centric building codes (e.g., WELL Standard, ASHRAE 55-2020). The observed plateau may indicate research pivoting from model development to validation studies and cross-domain integrations.

Prominent scholars, including T. Hong, D. Yan, and W. O'Brien, have shaped the trajectory of occupant modeling research, as illustrated in Figure 3 (analyzed via VOSviewer_1.6.20). Some authors [1,11,23] pioneered data-driven frameworks for energy behavior prediction and co-developed a standardized schema (e.g., Occupant Behavior Extensible Markup Language—obXML). Others [24] adopted Agent-Based modeling for spatial occupancy patterns and human-building interaction protocols. Moreover, they [7] also spearheaded hybrid models integrating machine learning with physical simulations for HVAC optimization. As shown in Figure 3, the co-authorship network reveals three structurally and temporally distinct clusters. The left cluster, centered around authors such as O'Brien, Mahdavi, and Hong, is closely linked to IEEE sources and reflects an engineering- and sensing-oriented modeling community. This cluster emerged during the mid-development phase (approximately 2018–2020), as indicated by the dominant blue-green coloration. The middle cluster, anchored by ASHRAE and researchers like Rijal and Clarke, displays a darker purple hue, signifying its earlier emergence (2016–2018) and strong foundation in standards, HVAC system design, and normative occupant models. The third cluster on the right exhibits a distinct yellow coloration, indicating a recent surge of activity since 2022. This cluster highlights a more localized and simulation-integrated modeling focus, suggesting emerging research themes in adaptive and context-specific occupant modeling.

A systematic review of articles reveals three distinct methodological phases in architectural occupant modeling. For the foundational phase (2001–2010), studies focus on thermal adaptation models, emergency egress simulations, and rule-based behavioral algorithms. These are mainly related to probabilistic thermal comfort frameworks [25], the quantification of adaptive window-opening behaviors across climates [26] and the development of stochastic models for occupant-driven energy demand [27]. The second phase was contributed by computational expansion (2011–2017) due to the emergence of IoT-enabled data mining and ABM for crowd dynamics. A couple of studies [11,23] introduced occupant-centric building energy modeling (OBEM) frameworks. Some research works [28,29] bridged cognitive science with spatial behavior prediction. Deep reinforcement learning (DRL), federated learning, and digital twin integration drive the

third intelligent phase that has lasted until now. These include deployment of transfer learning for cross-building behavior generalization [6,30], integration of graph neural networks (GNNs) with energy flexibility modeling [31,32] and the development of vision-language models for multi-modal occupant feedback analysis [33]. This progression mirrors the AEC industry's transition from deterministic assumptions to context-aware, self-calibrating models, driven by advances in edge computing and participatory design paradigms. The current phase emphasizes human-in-the-loop systems, with growing emphasis on ethical AI and privacy-preserving data strategies in occupant-centric design.

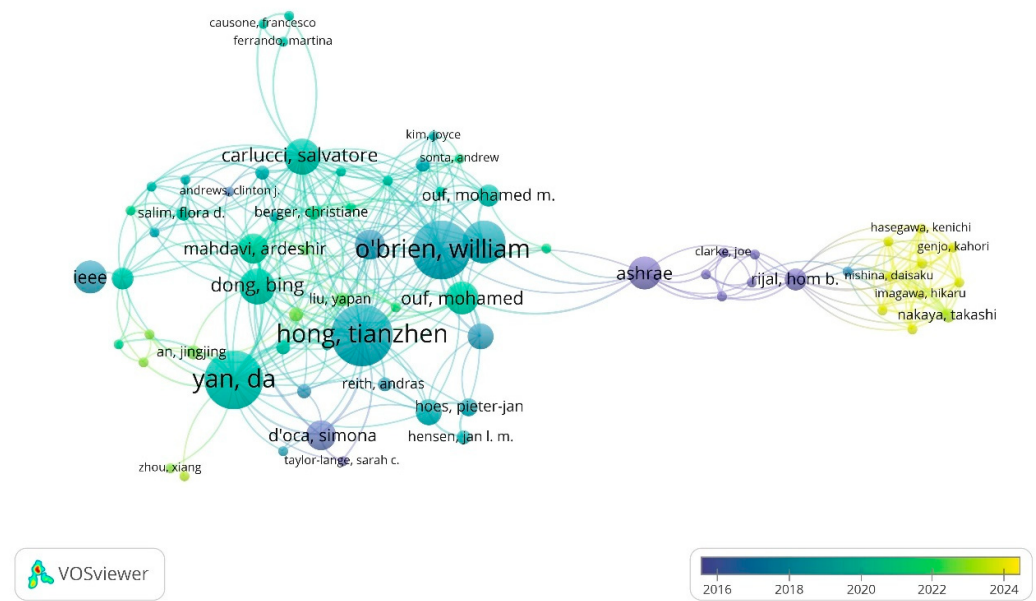


Figure 3. Representative authors and connections.

3. Methods and Applications of Occupant Modeling

3.1. Distributions of Current Models

As shown in Figure 4a, analysis of institutional affiliations reveals concentrated scholarly productivity across three primary regions: Asia (37 studies, 35%), North America (26 studies, 24%), and Europe (26 studies, 22%), with emerging contributions from Oceania, Africa, and South America. Notably, 7% International studies [4,34,35] demonstrate growing cross-border collaborations, particularly in climate-responsive modeling frameworks. This geographical skew highlights persistent asymmetries in research capacity, with Chinese and U.S. institutions collectively accounting for 45% of total publications. The limited representation from Global South regions underscores unmet needs for equitable knowledge co-creation in occupant modeling research. It is important to note that among the 312 studies reviewed, only 106 explicitly indicated the geographical region of their research. This limitation arises because many studies on occupant behavior modeling are conducted through technical simulations or framework discussions without contextualizing them in a specific regional setting. Therefore, more geographically attributed studies were encouraged in the future.

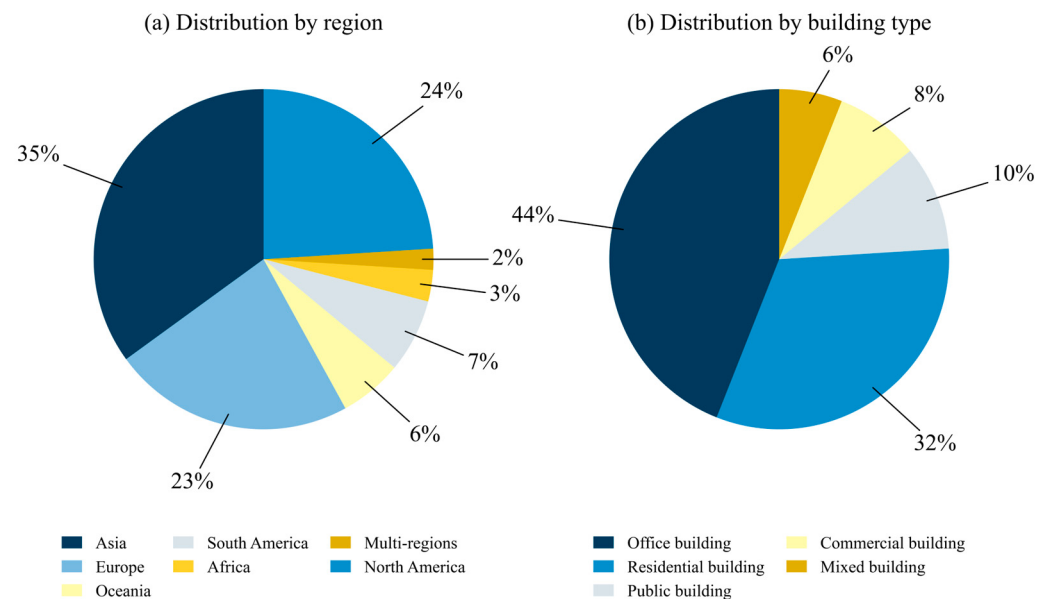


Figure 4. Regions and building types.

Figure 4b demonstrates pronounced application biases toward specific building types, with office buildings and residential structures constituting the primary domains for occupant modeling research (241 studies reported). High-rise/mid-rise offices [36,37], academic/commercial office complexes [11,38], private workspaces and activity-based flexible layouts [39,40] are all considered as office buildings (107 studies, 22.9% of total). Methodology focuses on office building occupancy scheduling, plug-load management, and meeting room utilization patterns. Residential Buildings (77 studies, 16.7%) include high-density apartments [41,42], single-family/multi-family dwellings [43], university dormitories and senior housing [44,45]. The thermostat adjustment behaviors, appliance usage patterns, and window operation frequency are the most concerned issues. Limited exploration of cross-typology behavioral synergies (e.g., live-work-play interactions) has been observed. Challenges in modeling dynamic occupancy overlaps between residential/commercial functions. This typological skew reflects methodological constraints in scaling occupant models across functional boundaries, with 83% of studies focusing on single-occupancy typologies. Emerging research priorities include parametric modeling of hybrid-use buildings.

3.2. Categories of Models

The diversity of occupant models reflects distinct methodological strengths and limitations, shaping their applicability across architectural design. Occupant models play a fundamental role in building intelligence, energy consumption prediction, and human comfort optimization. Based on a systematic analysis of 220 studies, occupant models in current research can be categorized into five types: statistical models, machine learning models, agent-based models, hybrid models, and probabilistic models. These methods exhibit significant heterogeneity across dimensions such as behavioral modeling capability, interpretability, data dependency, generalization capability, support for behavioral complexity, and modeling difficulty, as illustrated in Figure 5, making them suitable for different data conditions and task objectives in modeling scenarios. Table 1 lists representative algorithms, advantages and limitations of different models.

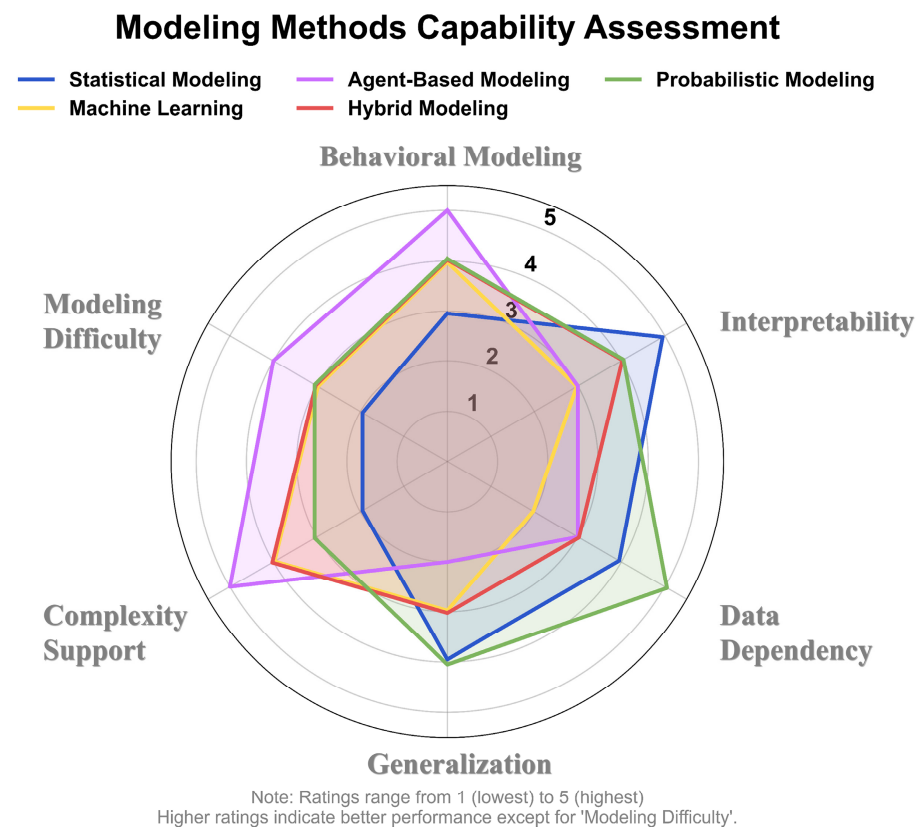


Figure 5. Modeling methods capability assessment.

Table 1. Features of different models.

Model Type	Representative Algorithm	Strengths	Limitations	Proportion
Statistical Models	Linear Regression, Logistic Regression, ANOVA	• High interpretability • Suitable for variable exploration and causal analysis • Low data requirements	• Unsuitable for nonlinear relationships and complex behaviors • Lacks temporal/interactive modeling capabilities	21.9%
Probabilistic Models	Markov Chains, Bayesian Networks, Monte Carlo (MC) Simulation	• Supports uncertainty modeling • Suitable for state-switching behaviors (e.g., on/off)	• Difficult to capture behavioral causality • Requires sufficient sampling • Hard to interpret	17.2%
Agent-Based Models (ABM)	Agent-Based Modeling, BDI Models	• Supports individual modeling and interaction • Suitable for complex scenario simulation • Facilitates context construction	• High modeling cost (relies on expert knowledge) • Lacks universal platform standards	18.1%

Table 1. Cont.

Model Type	Representative Algorithm	Strengths	Limitations	Proportion
Machine Learning (ML) Models	SVM, Random Forest, XGBoost, RNN	• Handles nonlinear and high-dimensional data • High prediction accuracy • Suitable for behavior recognition/classification	• Strong “black-box” nature (poor interpretability) • High sample quality requirements	33.9%
Hybrid Models	ABM + Optimization, Statistics + ML, MC + Rule Systems	• Integrate advantages of multiple methods • Balances accuracy and interpretability • Suitable for multi-objective/adaptive scenarios	• Complex structure (high implementation barrier) • Combination lacks standards; models are hard to transfer	8.9%

3.2.1. Statistical Models

Statistical models account for 21.9% of this area and typical algorithms include fixed statistical functions, regression analysis, analysis of variance (ANOVA) [46,47], sensitivity analysis [46] and structural equation modeling [48]. Their primary advantage lies in the clear model structure and parameters, with explicit physical or psychological interpretability [49–51]. By quantifying the extent to which different factors affect behavior or perception, they are suitable for tasks such as variable screening and causal inference, as modeling the causal structure between the built environment (building characteristics and the quality of indoor environment) and satisfaction [3,52]. As shown in Figure 5, they exhibit the strongest performance in terms of interpretability, even in multi-factor situations [53,54], while also offering the advantages of low modeling complexity and minimal data requirements. Their high interpretability and low modeling threshold make them the preferred choice for early researchers. Such methods are particularly valuable in variable selection, experimental design, and attributing behavioral trends, and are suitable for scenarios with clear problem structures and analyzable relationships between variables, such as impact factor assessment, lighting pattern [34] and modeling indoor thermal/light comfort preferences [46,48]. However, they have limited capabilities in handling high-dimensional data, nonlinear mechanisms, and time-dependent behaviors, with moderate behavioral modeling capabilities that support dynamic behavioral processes and behavioral complexity. This limits their applicability in dynamic systems.

Statistical models are commonly evaluated using R^2 and p -values, especially in regression-based frameworks. From the subset of 41 statistical modeling papers, p value tests whether a correlation exists, while the high or low R^2 value needs to be judged in combination with the specific methods, scenarios, observations, etc. In complex systems, due to the interaction of multiple factors and the influence of unobservable variables, R^2 is generally low (often below 0.3). At this time, a value exceeding 0.5 is considered to have a high explanatory power (for example, the 57.6% explanatory power of personal/environmental/physical parameters on lighting satisfaction [55] is already significant); in the thermal comfort area, R^2 is often close to 0.8 or above (such as 0.83 in reference [56]). Therefore, it is meaningless to judge R^2 in isolation from the field benchmark. The key is whether it has statistical and practical significance in similar studies. Statistical modeling tools remain foundational in occupant behavior studies, particularly for early-stage exploration of variable relationships and hypothesis-driven model devel-

opment. Commonly used tools include R, SPSS, MATLAB and Python. These tools offer robust support for regression analysis, logistic models, ANOVA, and multivariate statistics. R stands out for its package diversity, while Python offers reproducibility and integration with scripting-based workflows.

3.2.2. Probabilistic Models

Based on discrete choice, default scenario, and state transition theory, probabilistic models (accounting for 17.2%) were usually applied to binary behavior selection, state switching, and other scenarios, focusing on the inherent uncertainty of occupant behavior, with representative forms including Markov chains, Bayesian networks, and Monte Carlo simulations [26,57–59]. These models exhibit strong capabilities in behavioral modeling and support high levels of behavioral complexity, enabling the modeling of behavioral uncertainty and short-term state prediction, and effectively capturing the patterns of behavioral transitions. These models excel in expressing moderate interpretability and data dependency, making them particularly suitable for predicting discrete and switching behaviors [59,60]. However, it lacks the ability to describe causal structures and exogenous driving mechanisms, making it difficult to construct a complete explanatory framework. Its generalization ability is relatively weak, and the modeling difficulty is moderate. It is suitable for embedding into control systems or building operation simulation frameworks as a behavioral trigger module [39].

Typically, models with $R^2 \geq 0.7$ and RMSE close to zero are considered to have good performance, indicating that their predicted values are closer to reality, with a high proportion of explained variance and small prediction errors. However, model evaluation criteria must be combined with the uncertainty and data characteristics of specific scenarios: different scenarios (such as the complexity of influencing factors and noise levels) naturally have varying expectations for model performance; even within same subjects, performance may vary significantly across different prediction targets (e.g., kitchen occupancy rate prediction $R^2 = 0.84$ vs. living room $R^2 = 0.97$ [59]). Probabilistic modeling tools are typically implemented in general-purpose programming environments, notably Python, R, and MATLAB. These platforms support the construction of Bayesian Networks, Markov Chains, Hidden Markov Models, and Monte Carlo simulations—methods particularly suited to quantify behavioral randomness in equipment usage but struggle with real-time adaptability. Probabilistic models are often decoupled from building simulation engines, operating as upstream modules for scenario generation or behavior estimation, like EnergyPlus, OpenStudio [61,62] and Modelica [63].

3.2.3. Agent-Based Models

The advent of IoT sensing and computational advancements catalyzed a paradigm shift, with agent-based simulations gaining traction to quantify behavioral uncertainty and interpersonal interactions. Agent-based models account for 18.1% of all methods and have attracted widespread attention from researchers studying complex behavior, due to their powerful behavioral process modeling capabilities and group interaction simulation characteristics [64,65]. They focus on the interaction between individuals and their environment, using rule-driven agents to respond to situations. ABM possesses robust capabilities for expressing behavioral hierarchies and simulating interactions, making it particularly suitable for scenarios such as shared space behavior conflicts and collective response prediction [65,66]. Such methods perform best in supporting behavioral complexity, with full marks for behavioral modeling capabilities, moderate interpretability, and strong modeling accuracy control, such as evacuation [67]. However, their modeling rules heavily rely on expert experience, resulting in poor model generalizability. Additionally,

simulation granularity and computational complexity grow exponentially, posing significant challenges. These methods have high modeling costs, typically requiring manual rule design and individual behavior logic definition, and lack a unified modeling platform and standardized semantic specifications.

ABMs rarely explicitly report scalar accuracy metrics, as these models are based on simulations. Instead, when assessing behavioral realism and model conformity using qualitative methods, three dimensionless error metrics are typically examined: the coefficient of variation in the root mean square error (CV(RMSE)), the mean bias error (MBE), and the coefficient of determination (R^2). Typical calibration acceptance values for CV(RMSE) and MBE are 20% and $\pm 5\%$, respectively, while an acceptable R^2 value should be greater than 75% [65]. These ranges are context-specific and are typically associated with aggregated simulation outputs such as energy consumption or occupancy schedules. Agent-Based Modeling, mainly facilitated by AnyLogic [68,69], EnergyPlus and BIM-integrated platforms [70], could simulate crowd dynamics [71] and conduct social behavior analysis [72], although constrained by computational scalability. Software environments such as GAMA, AnyLogic, and NetLogo provide specialized capabilities for modeling autonomous decision-making, multi-agent interactions, and emergent collective behavior. These tools offer graphical or rule-based interfaces to define agent behaviors, decision drivers, and spatial navigation routines. Researchers increasingly couple ABMs with EnergyPlus or TRNSYS, either through intermediate data exchange (e.g., CSV, API) or co-simulation bridges. Despite these advances, calibration remains a persistent challenge due to the lack of standardized behavior ontologies and the difficulty of mapping agent-level behavior to real-world dynamics. Furthermore, few ABM tools support the stochastic parameter calibration or optimization required for large-scale simulations, which limits their scalability.

3.2.4. Machine Learning Models

Moreover, machine learning models account for 33.9% of the current literature, primarily including supervised learning (e.g., Support Vector Machines (SVM), Random Forests (RF), eXtreme Gradient-Boosted trees (XGBoost)) and deep neural networks (e.g., Long Short-Term Memory (LSTM), Recurrent Neural Networks (RNN)) strategies [73–76]. These methods can handle nonlinear relationships and high-dimensional multivariate inputs, demonstrating superior performance in tasks such as behavior prediction, personalized model construction, and behavior classification and recognition. Their behavioral modeling and generalization capabilities significantly outperform traditional methods and can support moderately complex behavioral modeling. However, despite their high accuracy, machine learning models score lowest in terms of interpretability, making them a typical example of “black-box models.” Additionally, they have high requirements for sample size and label quality, with relatively moderate to high modeling difficulty. Researchers are attempting to introduce explainable AI technologies or integrate them with rule-based systems to improve transparency [76].

Machine learning models report the widest variety of evaluation metrics, with the most frequent being R^2 and accuracy. Typical scenarios include thermal preference estimation [75,77,78] and energy consumption prediction [79,80]. Among 65 machine learning papers containing accuracy metrics, we observe that the R^2 range and accuracy range were usually higher than 0.80. For example, the spatial occupancy recognition accuracy of the CNN model based on multi-sensor data fusion can reach 92% [79], reflecting a strong capacity for quantitative prediction. The primary tool ecosystem of Machine learning (ML) is Python, particularly through libraries such as scikit-learn for conventional models (e.g., Random Forests, SVM), and TensorFlow or PyTorch for deep learning. These frameworks

support classification, regression, time-series forecasting, and unsupervised clustering. ML tools are model-centric but simulation-agnostic. However, although they excel in data handling and prediction, they do not natively interface with Building Performance Simulation (BPS) engines. For integration, ML models are often coupled via co-simulation platforms such as BCVTB or embedded in optimization loops through EMS (Energy Management Systems) or Modelica FMU (Functional Mock-up Unit) exports. Python makes it easy to bring in data and work with features smoothly, but the tricky part is making sure the results make sense physically and can be used effectively in decision-making systems that rely on simulations.

3.2.5. Hybrid Models

Recent years have seen accelerated innovation in hybrid models, which is increasingly becoming a research hotspot, although it accounts for the lowest proportion (8.84%) in current occupant modeling [81,82]. This method achieved complementary performance by combining multiple modeling strategies (such as ABM + optimization, statistics + ML, Monte Carlo + rule systems) and striking a good balance between accuracy, interpretability, and system behavior expression [83,84]. The greatest advantage of hybrid methods lies in their high flexibility and adaptability, making them suitable for modeling high-complexity systems (such as occupant–controller interactions). However, this type of method has the highest modeling difficulty, with its combination structure lacking universal patterns and standardized implementations, which makes it hard to turn into real-world applications. The modeling process relies on expert experience and has high transfer costs, making it an important direction for future methodological research.

Hybrid models combine machine learning, simulation, or rule-based methods. Papers in this group relevant to machine learning frequently reported high-performance metrics across different tasks: Accuracy was usually higher than 0.9. These models balance predictive performance and behavioral interpretation, often used in dynamic simulation [82,85] and adaptive occupant modeling [86]. Their composite architecture enhances robustness, particularly in multi-objective settings. Hybrid modeling strategies often require multi-platform toolchains, integrating rule-based systems, simulation engines, and machine learning pipelines. These models typically combine Python, EnergyPlus, and Building Information Modeling (BIM) in various configurations. In such setups, Python may be used for behavioral inference and feature selection, while EnergyPlus manages agent simulations and BIM handles physical modeling. Tools are heavily dependent on platform interoperability, frequently relying on co-simulation frameworks to facilitate communication between agents and building physics. The advantage lies in their ability to handle complex scenarios with multiple modeling resolutions. However, hybridization also introduces structural complexity and raises reproducibility concerns: tool configuration, synchronization timing, and data handoff mechanisms must be carefully controlled to avoid performance degradation or semantic inconsistencies.

To show interpretability and flexibility in occupant modeling, a minimal hybrid model template is summarized. Rule-based reasoning and data-driven learning through a layered structure were integrated. The behavioral layer captures occupant intent using predefined rules or heuristics (e.g., comfort thresholds, schedules), while the data-driven layer models behavioral variability based on sensor data or historical logs using techniques such as decision trees or neural networks. A fusion module combines both outputs to generate a unified control action (e.g., window opening or HVAC adjustment) at each simulation timestep. Inputs include environmental variables (temperature, humidity, CO₂, illumination) and occupancy status; outputs are time-stamped occupant actions aligned with Building Performance Simulation (BPS) needs. The model is calibrated with empirical

data and validated on out-of-sample datasets using standard metrics (e.g., MAE, CVRMSE). It is designed for seamless integration with platforms like EnergyPlus via co-simulation or API interfaces, and can ingest geometry and schedule data from BIM. This template supports reproducibility through transparent parameter disclosure, standard interface protocols, and documented model assumptions, offering a pragmatic baseline for future hybrid model development in design-driven scenarios.

In summary, although performance metrics reported across the studies were highly heterogeneous and not standardized, preventing a consistent statistical aggregation, a comparative review of representative studies suggests that machine learning and hybrid models generally outperform statistical and probabilistic approaches in prediction accuracy, while agent-based models exhibit greater variability due to contextual complexity. Meanwhile, according to widely recognized industry guidelines, ASHRAE Guideline 14 (2014) and IBPSA best-practice recommendations, acceptable thresholds for model validation are typically $|MBE| \leq 10\%$ and $CV(RMSE) \leq 30\%$ for hourly datasets. A review of the validation results reported in the included studies showed that most models fall within these accepted thresholds, particularly for occupancy prediction, window operation, and thermal comfort modeling. This observation highlights both the general reliability of existing occupant models and the ongoing need for more guidance for choosing the most appropriate modeling method.

3.3. Data Stream

Occupant behavior models rely on heterogeneous data inputs and outputs, each reflecting specific modeling objectives, data acquisition constraints, and validation paradigms. A critical synthesis of the existing literature reveals distinct trends and mismatches across input types, output focuses, data sources, and empirical grounding. Proportions of common input and output data of occupant models are demonstrated in Figure 6.

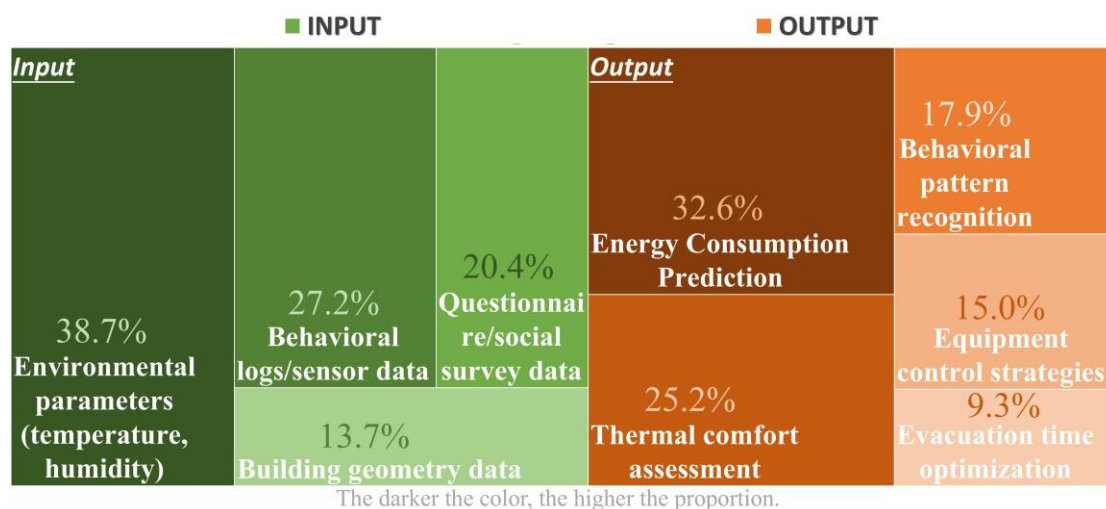


Figure 6. Proportions of input and output data of occupant models.

3.3.1. Input and Output

To illustrate the structural relationships between input categories, modeling approaches, and output targets, a Sankey diagram is shown in Figure 7. Through systematic classification and frequency analysis, several notable patterns emerge that reveal both the methodological preferences of researchers and the focus of current modeling efforts.

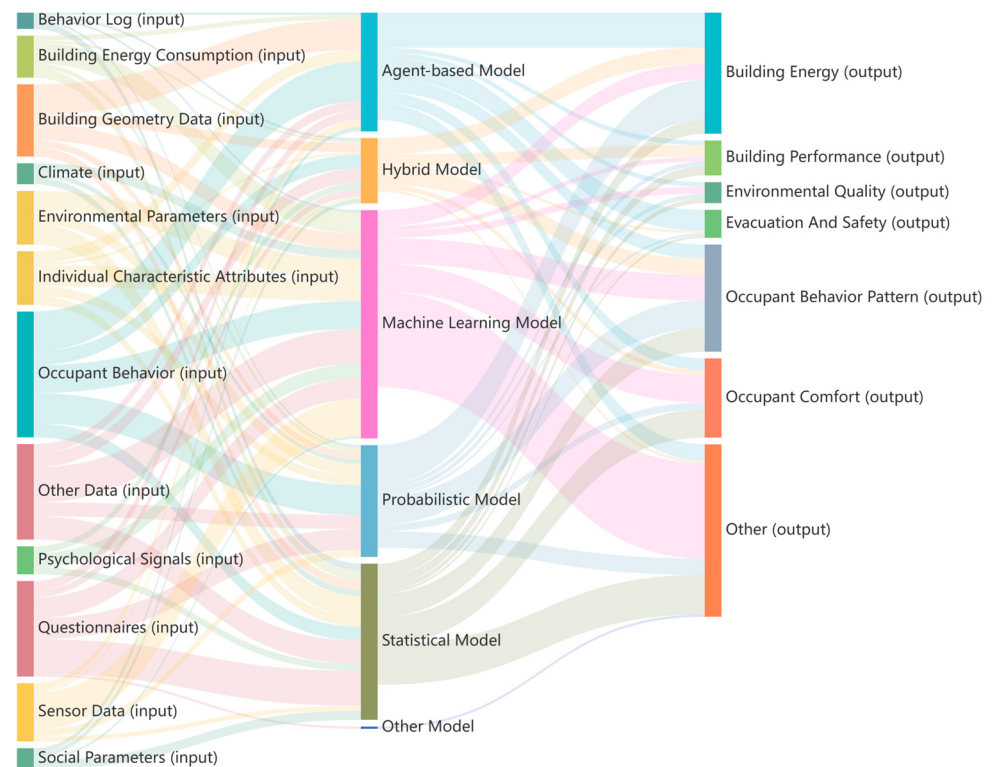


Figure 7. Data stream of occupant model.

First, input data sources are heavily dominated by occupant behavior, questionnaires, building geometry data, environmental parameters, and sensor data. Among these, “occupant behavior” appears as the most frequently cited input category, reflecting the central role of behavioral dynamics in contemporary models [41,73,87–91]. This prominence suggests a research orientation that prioritizes individualized modeling over purely physical or static inputs, aligning with recent shifts toward human-in-the-loop or occupant-centric building simulations. While questionnaires [52,92–94] provide a psychosocial context, but face scalability limitations due to subjective reporting biases. Building geometry data [68,95–97] and spatial configuration data [58,77–79] underpin spatial behavior analyses but remain underutilized due to complexity. Meanwhile, inputs for occupant modeling are also notably dominated by physical–environmental parameters, particularly temperature, humidity, and lighting [25,38,60,82,98,99], which are also highly represented, often in conjunction with sensor data, highlighting the increasing deployment of IoT-enabled data collection for model calibration and real-time prediction. These variables are essential for thermal comfort assessment and energy simulation, serving as foundational elements in both mechanistic and occupant models. Overall, input data exhibit a strong skew toward measurable environmental metrics, with 72% of data stemming from physical sources, often at the expense of underexplored but behaviorally informative psychosocial and geometric inputs.

In terms of outputs, categories concentrate around three main domains: building energy, occupant comfort, and occupant behavior patterns. Energy consumption prediction dominates applications (98 studies) [57,100,101], reflecting the industry’s compliance with regulatory decarbonization targets. Thermal comfort assessment, documented in 76 studies [74,102,103], which increasingly integrates physiological metrics with adaptive controls. This is especially evident in studies using environmental parameters or questionnaires as input, mapped onto statistical or machine learning models. Beyond energy and comfort, behavior pattern recognition (54 studies) [58,104,105] and evacuation time or path optimization (28 studies) [106,107] highlight safety-critical applications but often

rely on oversimplified, idealized scenarios that limit real-world transferability. However, the distribution of “other data” as input and “other” as output, though relatively minor, raises important considerations about the under-reported or emergent data types—such as psychological signals or social parameters—that remain rare in the current modeling ecosystem. Their low representation may point to methodological challenges, such as difficulty in data acquisition, lack of standardized metrics, or limited model compatibility.

In summary, the Sankey structure reveals a coherent but unbalanced ecosystem: data-rich, behavior-focused inputs are overwhelmingly mapped to machine learning outputs centered on energy and comfort, while interpretive or emergent domains—such as social factors, psychological responses, and integrative performance indices—remain under-explored. These patterns suggest both the maturity and the blind spots of the current research landscape, underscoring the need for more multimodal and context-sensitive modeling frameworks.

3.3.2. Linkages Between Data Stream

Distinct linkages exist between data types and their modeling applications. When mapping these inputs onto modeling approaches, machine learning models dominate as the most frequently applied category. Their prevalence underscores the field’s gravitation toward data-driven paradigms capable of capturing complex, nonlinear interactions between occupants and building systems. Particularly, the pairing of occupant behavior or sensor-derived data with machine learning models suggests a methodological tendency to exploit large datasets for predictive tasks such as energy consumption forecasting, comfort prediction, or behavior classification. In contrast, statistical models maintain a secondary but substantial presence, typically in scenarios where transparency and interpretability are prioritized, such as regression-based comfort evaluation or explanatory modeling of behavior patterns based on subjective evaluations. Agent-based models (ABMs) appear with moderate frequency, generally fed by individual-level characteristics and behavior data, especially in studies focused on evacuation and safety. The lower occurrence of probabilistic and hybrid models may reflect both the complexity of their implementation and a lack of standardized data workflows. For instance, questionnaire data (21.1% of studies) predominantly support psychosocial analyses, such as satisfaction assessments [108] and behavioral driver identification [109]. Similarly, physiological parameters (e.g., heart rate, skin temperature) [110] in 10.1% of studies enable personalized comfort prediction [111] and real-time HVAC control [112]. Wearable devices [112] and reinforcement learning [113] are redefining personalization, moving beyond aggregate comfort metrics to individualized thermal adaptation.

Meanwhile, analysis of data sourcing reveals divergent methodological practices. Nearly half of the literature (49.8%) employs empirical data for model calibration, primarily sourced from sensor logs [22], questionnaires [114], and equipment usage records. Hybrid datasets (empirical + simulated or literature-based) [115] constitute 20.6% [94], often used to fill data gaps or cross-validate assumptions. Conversely, simulated or literature-only data are used sparingly (9.4%) for hypothesis testing or supplementary analysis.

3.3.3. Reinforcement Caused by Simulated Data

Notably, most models utilize no more than two types of input data to produce a maximum of two outputs, highlighting the practical limitations associated with developing and maintaining complex, multi-dimensional models. However, a critical discrepancy emerges: while 67.1% of input data are empirically sourced, 88.7% of model outputs rely on simulated data, such as virtual scenarios, Monte Carlo simulations, and parametric BIM inputs. This heavy reliance on simulated outputs—despite empirical inputs—exposes a

validation gap, where models risk reinforcing theoretical assumptions rather than capturing real-world occupant variability. This conflict highlights the need for standardized benchmarking protocols that reconcile empirical inputs with empirical outputs, particularly in hybrid-use contexts where simulated data inadequately represent behavioral complexities. Future work should prioritize longitudinal, multi-modal datasets to bridge the current “empirical-synthetic divide” in occupant-centric building research. IoT, digital twin [116], and immersive virtual environments (IVEs) [33] could offer a unique opportunity and alternative for studying occupant energy behavior because of their potential to provide realistic and virtual experiences to participants and elicit their behavioral responses.

3.4. Application Scenarios of Occupant Modeling

The literature demonstrates a pronounced dichotomy between collective simulations (focused on group patterns) and individualized modeling (targeting occupant-specific dynamics), each addressing distinct scales of human-building interaction. Figure 8 displays studies about collective simulations are almost two times more than individualized modeling. This imbalance, where collective models account for 72% of simulations compared to individualized approaches (28%), reflects the historical emphasis on building-scale efficiency over occupant well-being. While collective methods benefit from computational tractability and regulatory alignment, their homogenized assumptions (e.g., average occupancy profiles) often misrepresent behavioral diversity. Conversely, individualized models, though more human-centric, require high-resolution data and face scalability challenges. Bridging this gap through hybrid frameworks, such as coupling agent-based crowd simulations with personalized comfort algorithms, could unlock synergies between energy efficiency and occupant agency, requiring advances in modeling methods and privacy-preserving data techniques.

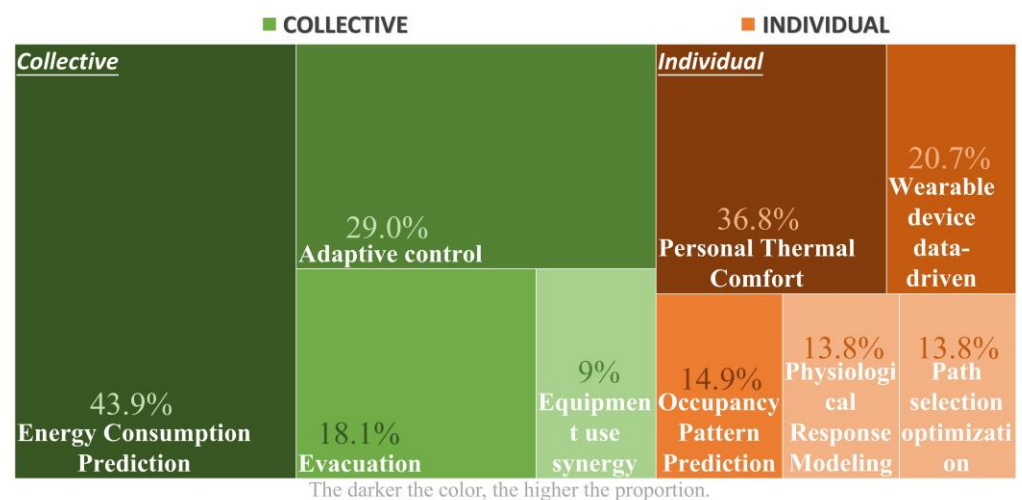


Figure 8. Proportions of application scenarios of occupant models.

Collective simulations dominate applications, with group energy consumption prediction (68 studies) [117,118] emerging as the most prevalent category, driven by regulatory demands for decarbonization and utility-driven load forecasting. Evacuation modeling (28 studies) [107,119] employs agent-based frameworks to simulate crowd dynamics, though its reliance on idealized spatial assumptions limits real-world applicability. Adaptive control interactions, especially window opening (45 studies) [53,120], contribute to optimizing HVAC and lighting for aggregate comfort but often overlook interpersonal variability, while equipment usage synergy (14 studies) [82,121] explores plug-load synchronization in offices, highlighting underexplored opportunities in demand-flexibility research.

In contrast, individualized modeling prioritizes occupant-centric adaptability, although with smaller sample sizes. Personal thermal comfort modeling (32 studies) with machine learning to adjust microclimate, but struggles with generalizability across demographics. Occupancy pattern prediction (13 studies) [59,98,122] and wearable-device-driven analytics (18 studies) [78,103] leverage real-time biometric data to refine personalized schedules but face ethical and privacy barriers. Physiological response modeling (12 studies) [55,111] and path-selection optimization (12 studies) [84,85] represent growing frontiers, emphasizing human factors in spatial navigation and health-centric design.

4. Discussion

4.1. Suggestions for Building the Most Suitable Occupant Model

To improve energy consumption and ensure occupant comfort, choosing the most suitable modeling method during the building design phase should be discussed. However, the diversity of modeling methods, tools, and behavioral contexts in occupant research has led to fragmentation in methodological choice. Architects often face ambiguity in selecting appropriate models due to inconsistent reporting standards, overlapping tool functionalities, and varying behavioral abstraction levels. Moreover, it should be noted that modeling approaches are not necessarily better if they are more complex; it is critical to select methods tailored to the specific problem. Different experimental objectives, parameters, experimental conditions, and contextual backgrounds require corresponding modeling strategies, often demanding comparative trial and error. However, most current studies rely on qualitative analyses. Li et al. [123] attempted to propose a set of selection criteria encompassing model complexity, computational time, variable selection, flexibility, integration capability, and expertise requirements, but lacked the comparison of hybrid models. Another significant breakthrough in supporting the selection of the most appropriate modeling method is the Fit-For-Purpose Occupant Behavior modeling (FFP-OBm) methodology [124]. Specifically, if the potential impact of occupant behavior (OB) aspects on simulation outcomes is low, those aspects should be modeled with minimal complexity. If the potential impact is identified as high but the associated uncertainty level is low, available knowledge should be utilized to model that specific OB aspect. The FFP-OBm methodology [124] remains the only proposed quantitative approach to date. However, it is too complex for architects to practice, and its demonstration has been limited to office buildings, heating and cooling demand estimation, and virtual experiments rather than real-world case studies, needing more validation.

To address this, we propose a decision-oriented standard that guides model selection based on scenario-specific requirements. This multi-layered framework provides recommendations, from data collection scenarios and types to model selection, bridging immersive spatial experiments, multidimensional data collection, and goal-driven design applications. The collection environment can be selected based on the novelty of the target space. Personalized spatial designs can be customized using virtual reality (VR), augmented reality (AR), or mixed reality (MR). Universally applicable spaces can be studied using a digital twin of existing similar spaces, actual environments, or modular laboratories. These environments support varying degrees of behavioral realism and environmental control, enabling customized assessments of occupant responses. At the data collection level, a two-dimensional taxonomy is employed, mapping data types onto axes of individual or interactive behavior and active or passive behavior. This includes subjective assessments, task performance, physiological and behavioral signals, social interaction, and passive device engagement (e.g., window or lighting controls). Different data collection methods are selected based on the desired complexity and sophistication. When selecting a modeling method, the design goals must be clearly defined, allowing evaluation across

the four dimensions: modeling objective, data availability and type, behavioral complexity, and platform integration requirements. As summarized in Figure 9, it serves as the basis for selecting an appropriate modeling approach.

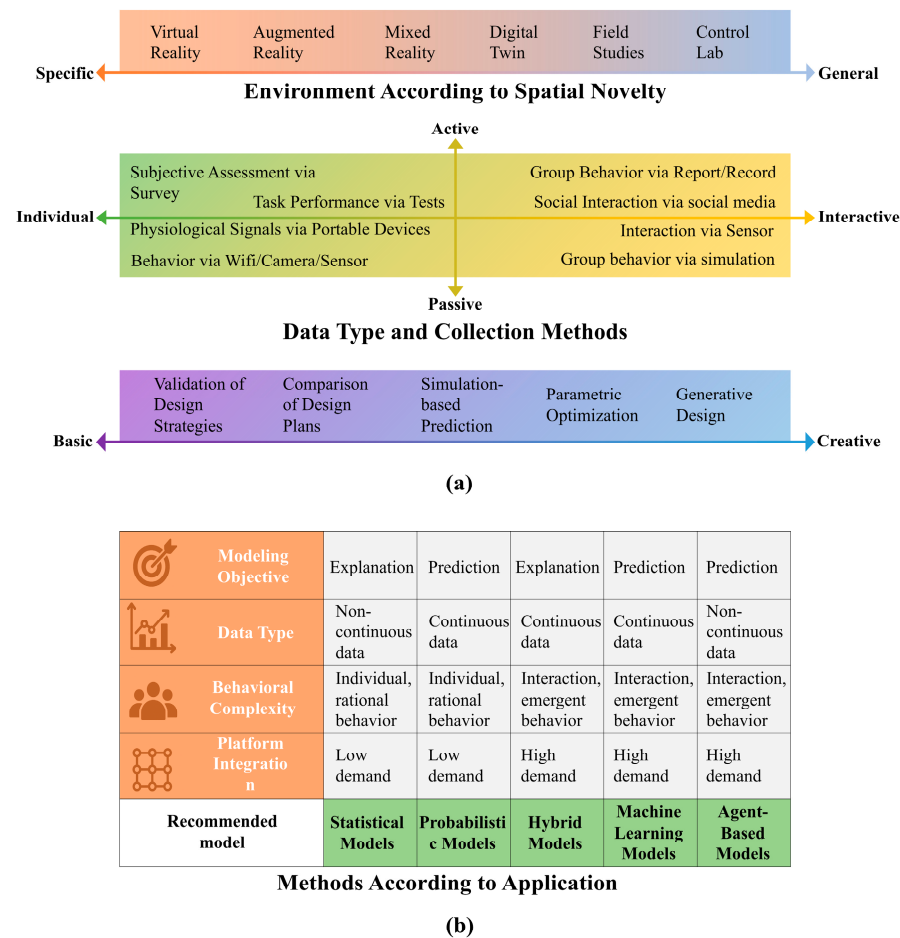


Figure 9. Modeling method selection reference: (a) four dimensions for modeling selection, (b) quantitative summary table.

For four dimensions: First, the modeling objective, whether focused on explaining phenomena or predicting outcomes, determines methodological orientation. For explanatory purposes, particularly when the goal is to uncover causal or psychological mechanisms behind occupant actions, statistical models and hybrid approaches are preferred due to their interpretability and capacity to integrate theoretical constructs. In contrast, predictive objectives, such as real-time window state estimation or energy demand forecasting, often necessitate machine learning, probabilistic, or agent-based models, which can effectively learn from large-scale data and capture temporal or emergent dynamics. Second, data characteristics and availability are key constraints in model selection. Statistical models are typically suited for discrete datasets, often derived from surveys, whereas machine learning and hybrid models are more adapted to high-frequency, continuous data streams from sensors or IoT platforms [91]. Agent-based models, though sometimes operating on discrete inputs, often require rich contextual datasets to calibrate individual agent rules and environmental interactions. Third, the behavioral complexity inherent in the system under study further narrows model choices. When behavior is largely individual and rational, statistical or probabilistic models—including Markov chains and Bayesian networks—are often sufficient to capture behavioral variability [60]. However, in cases involving interaction-rich, emergent, or adaptive behaviors, such as social influence, group

dynamics, or behavioral feedback loops, agent-based and hybrid models become essential. These models are capable of representing decentralized decision-making and capturing system-wide emergent outcomes [72,85,125]. Last, the requirement for integration with simulation platforms, such as EnergyPlus, TRNSYS, or BIM environments, directly affects the applicability of the model. Statistical and machine learning models are typically low-demand in terms of coupling complexity, often operating as pre- or post-processors via APIs. In contrast, hybrid and agent-based models often require tight coupling or co-simulation frameworks, especially when real-time control or dynamic simulation of occupant–system interaction is needed. While hybrid models, such as those coupling ABM with discrete event simulation [126], show significant promise in bridging occupant modeling with spatial and energy simulations [70,127], their practical adoption remains limited by integration complexity and a lack of standardized protocols such as obXML [128]. In summary, this framework emphasizes that modeling methods should not be selected based solely on disciplinary preference or data availability, but rather through a structured assessment of objective, data form, behavioral complexity, and platform needs. Mismatches between problem context and modeling paradigm can lead to either underpowered explanations or over-engineered predictive systems with limited generalizability. Therefore, the decision-guided modeling process could help to balance accuracy, interpretability, scalability, and operational feasibility as a prior reference.

To demonstrate how the proposed framework can support practical decision-making, an example of façade design optimization for natural ventilation and thermal comfort in an open-plan office is presented. (1) Identify modeling goal: The primary goal of this task is predictive, aiming to estimate occupants' window operation and indoor comfort responses under different façade porosity and shading configurations. Since the objective is not to explain behavioral mechanisms but to forecast behavioral outcomes in response to design changes, probabilistic and machine learning (ML) models are more suitable than purely statistical or explanatory ones. (2) Evaluate data characteristics: The input data include continuous environmental sensor streams (temperature, CO₂ concentration, air velocity) and interaction logs (window states). These high-frequency, contextual datasets require models capable of handling nonlinear and temporal dependencies. Hence, ML models such as random forest regressors, gradient boosting, or recurrent neural networks can be integrated with probabilistic layers to capture uncertainty in occupant response. (3) Analyze behavioral complexity: Window operation behavior exhibits moderate individual variation and context sensitivity but limited social interdependence. Consequently, a full agent-based approach is unnecessary. Instead, a probabilistic–ML hybrid model provides an efficient balance between behavioral variability representation and computational tractability. (4) Determine platform integration needs: To support the optimization, the predictive model must interact dynamically with Building Performance Simulation platforms to assess comfort and energy outcomes under varying façade parameters. In EnergyPlus, the coupling can be achieved via Python EMS or API-based data exchange, periodically updating occupant-related control variables. And occupant behavior parameters can be linked through semantic interoperability standards such as obXML or extended IFC nodes to be embedded into BIM environments, providing feedback during the design phase. For façade and ventilation optimization, BIM-based parametric tools (e.g., Dynamo, Rhino Inside Revit, Grasshopper with Ladybug/Honeybee) can directly communicate with EnergyPlus via Python API or standardized data exchange, enabling iterative performance feedback without additional co-simulation overhead. This approach achieves efficient coupling between design geometry, occupant behavior models, and building performance metrics while maintaining practical implementation feasibility for design teams. The overall analysis indicates that a probabilistic–machine learning hybrid model best satisfies this

task's requirements for predictive accuracy, contextual interpretability, and integration flexibility. This example demonstrates how the framework's four dimensions, goal, data, behavioral complexity, and platform integration, collectively guide the selection of suitable modeling approaches for design-oriented applications.

It is important to note that the decision-oriented framework proposed in this study is currently conceptual and descriptive. While it synthesizes existing modeling paradigms and links them to design variables, its validation in real-world design workflows or through controlled simulations remains a key direction for future research.

4.2. How Do Occupant Models Contribute to Building Design?

Traditional architectural design is predominantly driven by designers' subjective judgments, while building operation and maintenance are determined by occupants, leading to a lack of feedback between the two phases and resulting in deviations between design intent and actual operational performance. Occupant behavior models serve as a critical tool for bridging this gap during the design phase by enabling accurate prediction about operational performance through Building Performance Simulation. Although the diversity of occupant models determines that their guidance on architectural design must be tailored to local conditions, it is a scientific method currently widely used in the industry to evaluate the effectiveness of building design schemes. For example, probabilistic window opening models can help define operable facade zones, while agent-based simulations of movement patterns may inform layout zoning or egress pathways. Occupant preferences for daylight or acoustic privacy, when modeled and aggregated, can guide spatial adjacencies or facade aperture distributions. By integrating behavior-informed simulations early in the design process, architects can test multiple configurations and adjust spatial elements to better align with anticipated occupant needs. This bridges the gap between predictive modeling and actionable design interventions. The ontology based on drivers–needs–actions–systems framework is implemented through an Extensible Markup Language schema, titled 'occupant behavior XML', which is a practical implementation of OB models that can be integrated into BPS programs. As occupants are a major uncertainty factor in the actual operation and maintenance phase, the evaluation accuracy can be improved through the comprehensive application of occupant models and Building Performance Simulation. For example, one of the studies [129] proposed an Agent-Based Modeling method to evaluate the thermal comfort of subway stations in New Delhi, India. Other one [130] designed a framework that integrates the DNAS framework with Semantic Trajectories in Dynamic Environments data model to incorporate the dynamicity of building environments. Another work [128] presents a library of 52 OB models represented in the standardized obXML schema format, providing ready-to-use examples for BPS users to employ more accurate occupant representation in their energy models.

These models could help to explain the complex, bidirectional interactions between occupants and buildings [2]. However, current occupant-centric research primarily focuses on the prediction of the operational phase. Most models prioritize accuracy through black-box approaches, whose lack of interpretability prevents the extracted occupant-building interactions from directly informing design strategies beyond the model's runtime. Additionally, limitations in data availability constrain these models to specific scenarios or user groups, hindering their broader applicability to new construction projects. In addition, over 75% of models ultimately target energy savings or comfort optimization, reflecting the AEC industry's decarbonization mandates. However, fewer than 10% of studies explicitly address equity in resource distribution across socioeconomic groups. This lacks one most important functions of occupant models, humanistic care for people. Consequently, occupant-centric design optimization remains in its infancy. Future advancements may

require highly interpretable, real-time, updatable occupant models to efficiently support architectural decision-making.

At the same time, while high accuracy of occupant models is important for architectural design, it is even more essential to explore the nature of humans. Because of their adaptation and complexity, occupants can both widen the gap between design outcomes and potentially reduce the gap between different designs. Understanding and exploiting occupants is also an important step to provide more scientific support for active and passive design strategies, such as robust design, and facilitate the shift from an empirical to a data-driven architectural discipline. The move from 'BIM' (Building Information Modeling) to 'HIM' (Human Information Modeling) is the way forward [131], requiring a great deal of interdisciplinary collaboration. It is not only about obtaining relationships between occupants through social networks/sockets, etc., but also focusing on the traits of their emotions and health. At the same time, occupants are exposed to multiple environments with intertwined effects, as stated in the final report of ANNEX 79 [2], a strong need to base future studies on existing and advanced theories is required to bring the field of perception and behavior forward within interdisciplinary projects.

To support implementation and guide future research, a three-tier roadmap aligned with temporal and practical priorities was encouraged. In the near term (1–2 years), efforts should focus on establishing standardized validation benchmarks and interface protocols for integrating occupant models into BPS platforms. In the medium term (3–5 years), pilot studies in real architectural projects, particularly during early design phases, should test occupant-informed simulations in façade design, spatial zoning, and adaptive control. In the long term (>5 years), integration with adaptive Digital Twin frameworks should enable continuous behavior-informed optimization during both design and building operation, with feedback loops between sensors, simulations, and control systems. This roadmap provides a structured outlook for maturing occupant modeling from academic study to scalable design practice.

5. Conclusions

Occupant modeling has undergone a paradigm shift from empirical hypothesis to multidisciplinary integration. Based on a systematic analysis of 312 papers from 2001 to 2025, the present study reveals the methodologies and applications of current models. The diversity of occupant models reflects the advantages and limitations of different approaches in building design. The main results of the research are summarized as follows:

- Occupant modeling research has exhibited clear geographical and typological patterns. Most studies originate from technologically advanced regions such as North America, Western Europe, and East Asia. Regulatory pressures, digital infrastructure, and academic networks in these regions support occupant-centric innovation. Research has shifted from focusing on thermal comfort and adaptive behavior in office/residential environments to capturing dynamic behavior in mixed-use buildings. The rise in the Internet of Things (IoT), Agent-Based Modeling, and digital twin has reinforced this shift, enabling real-time data-driven applications in sensor-rich urban environments. Occupant modeling approaches are evolving from single, deterministic models to adaptive hybrid models fusing data-driven, agent-based simulation, stochasticity and physical principles, with the core driver being balancing computational complexity with human-centric design.
- Occupant models can be categorized into statistical models, machine learning models, agent-based models, probabilistic models, and hybrid models. Statistical models can provide powerful causal insights with minimal data requirements, but struggle with nonlinearity and temporal dynamics. Machine learning models offer high predictive

accuracy and scalability for complex behaviors, but they suffer from opacity and a reliance on large, high-quality datasets. Agent-based models excel at simulating individual interactions and emerging group dynamics, but they have limited generalization capabilities. Probabilistic models effectively capture uncertainty and state transitions but lack causal depth. Hybrid models, though emerging, integrate methodological advantages to balance accuracy and transparency, but face challenges in standardization and reproducibility. Overall, the diversity of occupant modeling types reflects the ongoing trade-off between behavioral fidelity, computational feasibility, and future directions may lie in hybrid and context-aware modeling frameworks.

- The current occupant modeling data framework exhibits significant imbalance: although input data increasingly originates from empirical sources—particularly occupant behavior and subjective data—output results remain highly dependent on simulation. Meanwhile, the adoption of parameters like psychosocial signals is minimal due to integration challenges. As a result, models risk reinforcing theoretical assumptions rather than capturing actual occupant variability. The application of physiological signals, immersive environments, and longitudinal datasets should be considered for the need for more comprehensive multimodal data strategies and standardized benchmarking protocols to balance empirical foundations with predictive complexity in occupant-centered modeling.
- Occupant modeling applications reveal a significant imbalance between collective and individualized approaches, reflecting an emphasis on building-scale efficiency and regulatory compliance, which tends to aggregate occupant behavior for energy prediction, evacuation planning, and adaptive control strategy simulation. The differences in application scenarios highlight methodological trade-offs: collective models align with system-level objectives but risk oversimplification, while individualized models prioritize behavioral realism but sacrifice broader applicability. Future research may need to bridge these paradigms through hybrid frameworks, potentially achieving a balance between efficiency and personalization, contingent on advancements in data integration and privacy-preserving modeling techniques.
- To determine the most appropriate occupant model, we propose a structured, decision-oriented framework that combines modeling strategies with specific design objectives, data availability, behavioral complexity, and platform integration requirements to guide method selection. At the same time, occupant models serve as a key tool for bridging the gap between design intent and operational performance. Although currently mainly applied in the operational phase, focusing on energy and comfort optimization. However, its transformative potential in early design stages remains underexplored, such as when integrated with a digital twin.

Occupant modeling is transitioning from siloed, deterministic paradigms to adaptive, human-in-the-loop systems. Future success depends on standardizing hybrid workflows, prioritizing equity in model applications, and fostering interdisciplinary collaborations to align technical advancements with societal needs. It is suggested that (1) develop hybrid models that balance accuracy and transparency with higher generalization ability; (2) advance BIM-IoT integrations to create adaptive virtual replicas that dynamically respond to occupant feedback; (3) expand from single domain applications to interdisciplinary simulations and occupant-centric design, leveraging IoT, digital twin and IVE for occupant resilience; (4) establish anonymized data-sharing mechanisms and consent frameworks for biometric datasets. It is necessary to promote the comprehensive application of occupant modeling in smart buildings. These efforts will position occupant modeling as a cornerstone of sustainable, human-centric building design in the new era.

Author Contributions: All authors contributed to the study conception and design. Conceptualization: R.S., C.D.P. and R.S.A.; Formal analysis and investigation: C.S., D.Q. and R.S.; Writing—original draft preparation: R.S.; Writing—review and editing: C.S., R.S.A. and D.Q.; Supervision: C.S., R.S.A. and C.D.P. The first draft of the manuscript was written by R.S. and all authors commented on previous versions of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding authors.

Acknowledgments: One of the authors (Rui Sun) was supported by the China Scholarship Council for an 18-month study at Politecnico di Milano. The authors used AI tools in the manuscript.

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

Abbreviations

The following abbreviations are used in this manuscript:

ABM	Agent-Based Modeling
AEC	Architecture, Engineering, and Construction
BIM	Building Information Modeling
BPS	Building Performance Simulation
RMSE	Root Mean Square Error
DNAs	Standardizing Drivers-Needs-Actions-system
DNN	Deep Neural Networks
DRL	Deep reinforcement learning
FMU	Functional Mock-up Unit
GNN	Graph Neural Networks
HIM	Human Information Modeling
LSTM	Long Short-Term Memory
OB	Occupant Modeling
OCD	Occupant-centric Design
OBEM	Occupant-centric Building Energy Modeling
RF	Random Forests
RNN	Recurrent Neural Networks
R ²	Coefficient of Determination
SEM	Structural Equation Modeling
SVM	Support Vector Machines
ML	Machine Learning
MC	Monte Carlo
MBE	Mean Bias Error
XGBoost	EXtreme Gradient-Boosted trees

References

1. Hong, T.; Yan, D.; D'Oca, S.; Chen, C.-F. Ten questions concerning occupant behavior in buildings: The big picture. *Build. Environ.* **2017**, *114*, 518–530. [[CrossRef](#)]
2. 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](#)]
3. Cao, B.; Ouyang, Q.; Zhu, Y.; Huang, L.; Hu, H.; Deng, G. Development of a multivariate regression model for overall satisfaction in public buildings based on field studies in Beijing and Shanghai. *Build. Environ.* **2012**, *47*, 394–399. [[CrossRef](#)]
4. Jin, Y.; Yan, D.; Chong, A.; Dong, B.; An, J. Building occupancy forecasting: A systematical and critical review. *Energy Build.* **2021**, *251*, 111345. [[CrossRef](#)]

5. Zhong, C.; Choi, J.-H. Development of a Data-Driven Approach for Human-Based Environmental Control. In Proceedings of the 10th International Symposium on Heating, Ventilation and Air Conditioning, ISHVAC 2017, Jinan, China, 19–22 October 2017; pp. 1665–1671.
6. Dziedzic, J.W.; Yan, D.; Sun, H.; Novakovic, V. Building occupant transient agent-based model—Movement module. *Appl. Energy* **2020**, *261*, 114417. [\[CrossRef\]](#)
7. 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. [\[CrossRef\]](#)
8. Gaetani, I.; Hoes, P.-J.; Hensen, J.L.M. Occupant behavior in building energy simulation: Towards a fit-for-purpose modeling strategy. *Energy Build.* **2016**, *121*, 188–204. [\[CrossRef\]](#)
9. Norouziasl, S.; Jafari, A.; Zhu, Y. Modeling and simulation of energy-related human-building interaction: A systematic review. *J. Build. Eng.* **2021**, *44*, 102928. [\[CrossRef\]](#)
10. O'Brien, W.; Gunay, H.B. The contextual factors contributing to occupants' adaptive comfort behaviors in offices—A review and proposed modeling framework. *Build. Environ.* **2014**, *77*, 77–87. [\[CrossRef\]](#)
11. Hong, T.; D'Oca, S.; Turner, W.J.N.; Taylor-Lange, S.C. An ontology to represent energy-related occupant behavior in buildings. Part I: Introduction to the DNAs framework. *Build. Environ.* **2015**, *92*, 764–777. [\[CrossRef\]](#)
12. Delzendeh, E.; Wu, S.; Lee, A.; Zhou, Y. The impact of occupants' behaviours on building energy analysis: A research review. *Renew. Sustain. Energy Rev.* **2017**, *80*, 1061–1071. [\[CrossRef\]](#)
13. Jia, M.; Srinivasan, R.S.; Raheem, A.A. From occupancy to occupant behavior: An analytical survey of data acquisition technologies, modeling methodologies and simulation coupling mechanisms for building energy efficiency. *Renew. Sustain. Energy Rev.* **2017**, *68*, 525–540. [\[CrossRef\]](#)
14. Abraham, Y.S.; Anumba, C.J.; Asadi, S. Exploring Agent-Based Modeling Approaches for Human-Centered Energy Consumption Prediction. In Proceedings of the Construction Research Congress 2018: Sustainable Design and Construction and Education, New Orleans, LA, USA, 2 April 2018; pp. 368–378.
15. Kanthila, C.; Boodi, A.; Beddiar, K.; Amirat, Y.; Benbouzid, M. Building Occupancy Behavior and Prediction Methods: A Critical Review and Challenging Locks. *IEEE Access* **2021**, *9*, 79353–79372. [\[CrossRef\]](#)
16. Uddin, M.N.; Wei, H.-H.; Chi, H.L.; Ni, M. Influence of Occupant Behavior for Building Energy Conservation: A Systematic Review Study of Diverse Modeling and Simulation Approach. *Buildings* **2021**, *11*, 41. [\[CrossRef\]](#)
17. Azar, E.; O'Brien, W.; Carlucci, S.; Hong, T.; Sonta, A.; Kim, J.; Andargie, M.S.; Abuimara, T.; El Asmar, M.; Jain, R.K.; et al. Simulation-aided occupant-centric building design: A critical review of tools, methods, and applications. *Energy Build.* **2020**, *224*, 110292. [\[CrossRef\]](#)
18. Ding, Y.; Han, S.; Tian, Z.; Yao, J.; Chen, W.; Zhang, Q. Review on occupancy detection and prediction in building simulation. *Build. Simul.* **2022**, *15*, 333–356. [\[CrossRef\]](#)
19. Ahmed, O.; Sezer, N.; Ouf, M.; Wang, L.; Hassan, I.G. State-of-the-art review of occupant behavior modeling and implementation in building performance simulation. *Renew. Sustain. Energy Rev.* **2023**, *185*, 113558. [\[CrossRef\]](#)
20. Wang, G.; Zhu, P.; Yao, S.; Yuan, J.; Hu, T. Occupant behavior model and its involvement in building optimization design: A review. *J. Asian Archit. Build. Eng.* **2025**, *24*, 2322–2338. [\[CrossRef\]](#)
21. Ebuy, H.T.; El Haouzi, H.B.; Benelmir, R.; Pannequin, R. Occupant Behavior Impact on Building Sustainability Performance: A Literature Review. *Sustainability* **2023**, *15*, 2440. [\[CrossRef\]](#)
22. Caballero-Pena, J.; Osma-Pinto, G.; Rey, J.M.; Nagarsheth, S.; Henao, N.; Agbossou, K. Analysis of the building occupancy estimation and prediction process: A systematic review. *Energy Build.* **2024**, *313*, 114230. [\[CrossRef\]](#)
23. Hong, T.; Taylor-Lange, S.C.; D'Oca, S.; Yan, D.; Corgnati, S.P. Advances in research and applications of energy-related occupant behavior in buildings. *Energy Build.* **2016**, *116*, 694–702. [\[CrossRef\]](#)
24. Gunay, H.B.; O'Brien, W.; Beausoleil-Morrison, I. Implementation and comparison of existing occupant behaviour models in EnergyPlus. *J. Build. Perform. Simul.* **2016**, *9*, 567–588. [\[CrossRef\]](#)
25. Kang, Z.J.; Xue, H.; Bong, T.Y. Modeling of thermal environment and human response in a crowded space for tropical climate. *Build. Environ.* **2001**, *36*, 511–525. [\[CrossRef\]](#)
26. Rijal, H.B.; Tuohy, P.; Humphreys, M.A.; Nicol, J.F.; Samuel, A.; Raja, I.A.; Clarke, J. Development of Adaptive Algorithms for the Operation of Windows, Fans, and Doors to Predict Thermal Comfort and Energy Use in Pakistani Buildings. *Ashrae Trans.* **2008**, *114*, 555–573.
27. Tuohy, P.; Rijal, H.B.; Humphrey, M.A.; Nicol, J.F.; Samuel, A.; Clarke, J. Comfort driven adaptive window opening behavior and the influence of building design. In Proceedings of the Building Simulation 2007: 10th Conference of IBPSA, Beijing, China, 27–30 July 2007; Volume 1–3, pp. 717–724.
28. Azar, E.; Menassa, C.C. Impact of Occupants Behavior on Building Energy Use: An Agent-Based Modeling Approach. In Proceedings of the 10th International Conference on Modeling and Applied Simulation, MAS, Rome, Italy, 12–14 September 2011; pp. 232–241.

29. Azar, E.; Menassa, C.C. Agent-Based Modeling of Occupants and Their Impact on Energy Use in Commercial Buildings. *J. Comput. Civ. Eng.* **2012**, *26*, 506–518. [\[CrossRef\]](#)
30. Dziedzic, J.; Yan, D.; Novakovic, V. Occupant migration monitoring in residential buildings with the use of a depth registration camera. In Proceedings of the 10th International Symposium on Heating, Ventilation and Air Conditioning, ISHVAC, Jinan, China, 19–22 October 2017; pp. 1193–1200.
31. 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. [\[CrossRef\]](#)
32. Dong, B.; Yan, D.; Li, Z.; Jin, Y.; Feng, X.; Fontenot, H. Modeling occupancy and behavior for better building design and operation—A critical review. *Build. Simul.* **2018**, *11*, 899–921. [\[CrossRef\]](#)
33. Zhu, Y.; Saeidi, S.; Rizzuto, T.; Roetzel, A.; Kooima, R. Potential and challenges of immersive virtual environments for occupant energy behavior modeling and validation: A literature review. *J. Build. Eng.* **2018**, *19*, 302–319. [\[CrossRef\]](#)
34. Das, A.; Paul, S.K. Artificial illumination during daytime in residential buildings: Factors, energy implications and future predictions. *Appl. Energy* **2015**, *158*, 65–85. [\[CrossRef\]](#)
35. Lu, C. Enhancing real-time nonintrusive occupancy estimation in buildings via knowledge fusion network. *Energy Build.* **2024**, *303*, 113812. [\[CrossRef\]](#)
36. Munoz, J.S.; Kelly, M.T.; Flores-Ales, V.; Caamano-Carrillo, C. Recognizing the effect of the thermal environment on self-perceived productivity in offices: A structural equation modeling perspective. *Build. Environ.* **2022**, *210*, 108696. [\[CrossRef\]](#)
37. Pinder, J.Z. Modelling the Utility and Occupancy Costs of Local Authority Office Buildings. Ph.D. Thesis, Sheffield Hallam University, Sheffield, UK, 2004. Available online: <http://shura.shu.ac.uk/20230/> (accessed on 17 October 2025).
38. Dutton, S.M.Z. Window Opening Behaviour and its Impact on Building Simulation: A Study in the Context of School Design. Ph.D. Thesis, University of Nottingham, Nottingham, UK, 2009.
39. Sadeghi, S.A.Z. Visual Preferences and Human Interactions with Shading and Electric Lighting Systems. Ph.D. Thesis, Purdue University, West Lafayette, IN, USA, 2018. Available online: https://docs.lib.purdue.edu/open_access_dissertations/1815/ (accessed on 17 October 2025).
40. Salem, D.; Elwakil, E.; Kandil, A. Fuzzy-Based Model for Predicting Lighting Efficiency in Institutional Buildings. In Proceedings of the 2015 Annual Meeting of the North American Fuzzy Information Processing Society DigiPen NAFIPS 2015, Redmond, WA, USA, 17–19 August 2015.
41. Zaraket, T.; Yannou, B.; Leroy, Y.; Minel, S.; Chapotot, E. An Occupant-Based Energy Consumption Model for User-Focused Design of Residential Buildings. *J. Mech. Des.* **2015**, *137*, 071412. [\[CrossRef\]](#)
42. 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. [\[CrossRef\]](#)
43. Mo, H.; Sun, H.; Liu, J.; Wei, S. Developing window behavior models for residential buildings using XGBoost algorithm. *Energy Build.* **2019**, *205*, 109564. [\[CrossRef\]](#)
44. Zhang, Y.; Zhang, B.; Hou, J. Simulation Study on Student Residential Energy Use Behaviors: A Case Study of University Dormitories in Sichuan, China. *Buildings* **2024**, *14*, 1484. [\[CrossRef\]](#)
45. Sharmin, T.; Gul, M.; Al-Hussein, M. A user-centric space heating energy management framework for multi-family residential facilities based on occupant pattern prediction modeling. *Build. Simul.* **2017**, *10*, 899–916. [\[CrossRef\]](#)
46. Gaetani, I.; Hoes, P.-J.; Hensen, J.L.M. On the sensitivity to different aspects of occupant behaviour for selecting the appropriate modelling complexity in building performance predictions. *J. Build. Perform. Simul.* **2017**, *10*, 601–611. [\[CrossRef\]](#)
47. Ji, L.; Hu, M.; Zhang, L.; Sun, Y. Evaluation Models for Luminous Environment Satisfaction in Green Office Buildings Integrating Environmental and Spatial Attributes Based on Massive Data Samples. In Proceedings of the 2021 2nd International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE 2021), Zhuhai, China, 24–26 September 2021; pp. 62–69.
48. Li, D.; Menassa, C.C.; Karatas, A. Energy use behaviors in buildings: Towards an integrated conceptual framework. *Energy Res. Soc. Sci.* **2017**, *23*, 97–112. [\[CrossRef\]](#)
49. Tuomaala, P.; Piippo, J.; Piira, K.; Airaksinen, M. Human Thermal Responses in Energy-efficient Buildings. In Proceedings of the First International Conference on Improving Construction and Use Through Integrated Design Solutions, Espoo, Finland, 10–12 June 2009; pp. 253–267.
50. Yun, G.Y.; Kim, J.T. Creating sustainable building through exploiting human comfort. In Proceedings of the 6th International Conference on Sustainability in Energy and Buildings, Cardiff, UK, 25–27 June 2014; pp. 590–594.
51. Rida, M.; Hoffmann, S. Using a Dynamic Clothing Insulation Model in Building Simulation—Impact on Thermal Comfort and Energy Consumption. In Proceedings of the Building Simulation 2019: 16th Conference of IBPSA, Rome, Italy, 2–4 September 2019; pp. 2302–2309.

52. Mostavi, E.; Asadi, S.; Ramaji, I.J. Completing the Missing Puzzle Piece of the Building Design Process: Modeling and Identifying Occupants' Satisfaction Level in Commercial Buildings. In Proceedings of the Construction Research Congress 2016: Old and New Construction Technologies Converge in Historic San Juan, San Juan, Puerto Rico, 31 May–2 June 2016; pp. 1112–1121.
53. Kim, A.; Wang, S.; Kim, J.-E.; Reed, D. Indoor/Outdoor Environmental Parameters and Window-Opening Behavior: A Structural Equation Modeling Analysis. *Buildings* **2019**, *9*, 94. [CrossRef]
54. Xu, L.; Zhang, Z. Effects of residential indoor environments on occupant satisfaction and performance. *J. Asian Archit. Build. Eng.* **2024**, *23*, 282–293. [CrossRef]
55. Fakhari, M.; Fayaz, R. Investigating the interrelation of influential parameters on indoor daylight quality using structural equation modeling. *Sol. Energy* **2023**, *256*, 179–190. [CrossRef]
56. Su, X.; Wang, Z.; Zhou, F.; Duanmu, L.; Zhai, Y.; Lian, Z.; Cao, B.; Zhang, Y.; Zhou, X.; Xie, J. Comfortable clothing model of occupants and thermal adaption to cold climates in China. *Build. Environ.* **2022**, *207*, 108499. [CrossRef]
57. Ao, J.; Du, C.; Jing, M.; Li, B.; Chen, Z. A Method of Integrating Air Conditioning Usage Models to Building Simulations for Predicting Residential Cooling Energy Consumption. *Buildings* **2024**, *14*, 2026. [CrossRef]
58. Genjo, K.; Nakanishi, H.; Oki, M.; Imagawa, H.; Uno, T.; Saito, T.; Takata, H.; Tsuzuki, K.; Nakaya, T.; Nishina, D.; et al. Development of Adaptive Model and Occupant Behavior Model in Four Office Buildings in Nagasaki, Japan. *Energies* **2023**, *16*, 6060. [CrossRef]
59. Zhang, R.; Zhou, T.; Ye, H.; Darkwa, J. Introducing a novel method for simulating stochastic movement and occupancy in residential spaces using time-use survey data. *Energy Build.* **2024**, *304*, 113854. [CrossRef]
60. Zhou, X.; Liu, T.; Yan, D.; Shi, X.; Jin, X. An action-based Markov chain modeling approach for predicting the window operating behavior in office spaces. *Build. Simul.* **2021**, *14*, 301–315. [CrossRef]
61. Stoppel, C.M.; Leite, F. Integrating probabilistic methods for describing occupant presence with building energy simulation models. *Energy Build.* **2014**, *68*, 99–107. [CrossRef]
62. Yang, Z.; Becerik-Gerber, B. Modeling personalized occupancy profiles for representing long term patterns by using ambient context. *Build. Environ.* **2014**, *78*, 23–35. [CrossRef]
63. Verbruggen, S.; Laverge, J.; Delghust, M.; Janssens, A. Stochastic Occupant Behaviour Model: Impact on residential energy use. In Proceedings of the Building Simulation 2019: 16th Conference of IBPSA, Rome, Italy, 2–4 September 2019; pp. 2310–2317.
64. Hassanpour, S.; Gonzalez, V.; Liu, J.; Zou, Y.; Cabrera-Guerrero, G. A hybrid hierarchical agent-based simulation approach for buildings indoor layout evaluation based on the post-earthquake evacuation. *Adv. Eng. Inform.* **2022**, *51*, 101531. [CrossRef]
65. Pedarla, L.P.; Khazaii, J. Modeling Effects of Occupants' Time-Off Behavior in Buildings on Load Calculation and Energy Modeling. In Proceedings of the ASME 2022 International Mechanical Engineering Congress and Exposition, IMECE 2022, Columbus, OH, USA, 30 October–3 November 2022; Volume 6.
66. Chapman, J.Z. Multi-Agent Stochastic Simulation of Occupants in Buildings. Ph.D. Thesis, University of Nottingham, Nottingham, UK, 2017. Available online: <https://eprints.nottingham.ac.uk/39868/> (accessed on 17 October 2025).
67. Oven, V.A.; Cakici, N. Modelling the evacuation of a high-rise office building in Istanbul. *Fire Saf. J.* **2009**, *44*, 1–15. [CrossRef]
68. Awada, M.; Hajj-Hassan, M.; Kiomjian, D.; Srouf, I.; Khoury. Energy Performance and Occupant Comfort in an Office Building: Co-simulation of an Agent-based Behavior Model with EnergyPlus. In Proceedings of the Third International Conference on Efficient Building Design—Materials and Hvac Equipment Technologies, Beirut, Lebanon, 4–5 October 2018; pp. 215–223.
69. Chen, Y.; Liang, X.; Hong, T.; Luo, X. Simulation and visualization of energy-related occupant behavior in office buildings. *Build. Simul.* **2017**, *10*, 785–798. [CrossRef]
70. Marzouk, M.; Mohamed, B. Multi-Criteria Ranking Tool for Evaluating Buildings Evacuation Using Agent-Based Simulation. In Proceedings of the Construction Research Congress 2018: Safety and Disaster Management, New Orleans, LA, USA, 2–4 April 2018; pp. 472–481.
71. Chu, M.L.; Parigi, P.; Law, K.; Latombe, J.-C. Modeling social behaviors in an evacuation simulator. *Comput. Animat. Virtual Worlds* **2014**, *25*, 375–384. [CrossRef]
72. Lian, X.; Zhu, H.; Zhang, X.; Jin, Y.; Zhou, H.; He, B.; Li, Z. Recognition of typical environmental control behavior patterns of indoor occupants based on temporal series association analysis. *Build. Environ.* **2023**, *234*, 110170. [CrossRef]
73. Asadi, N.; Moosavi, L. Investigation of window opening behavior during cold seasons through a non-intrusive sensor-based data-driven approach. *Energy Build.* **2024**, *317*, 114386. [CrossRef]
74. Cakir, M.; Akbulut, A. A Bayesian Deep Neural Network Approach to Seven-Point Thermal Sensation Perception. *IEEE Access* **2022**, *10*, 5193–5206. [CrossRef]
75. Boutahri, Y.; Tilioua, A. Machine learning-based predictive model for thermal comfort and energy optimization in smart buildings. *Results Eng.* **2024**, *22*, 102148. [CrossRef]
76. Yu, H.; Xu, X. Reinforcement learning for occupant behavior modeling in public buildings: Why, what and how? *J. Build. Eng.* **2024**, *96*, 110491. [CrossRef]

77. Fu, Y.; Zhou, T.; Lun, I.; Khayatian, F.; Deng, W.; Su, W. A data-driven approach for window opening predictions in non-air-conditioned buildings. *Intell. Build. Int.* **2022**, *14*, 329–345. [\[CrossRef\]](#)
78. Alsaleem, F.; Tesfay, M.K.; Rafaie, M.; Sinkar, K.; Besarla, D.; Arunasalam, P. An IoT Framework for Modeling and Controlling Thermal Comfort in Buildings. *Front. Built Environ.* **2020**, *6*, 87. [\[CrossRef\]](#)
79. Zhou, Y.; Wang, Y.; Li, C.; Ding, L.; Yang, Z. Energy-efficiency oriented occupancy space optimization in buildings: A data-driven approach based on multi-sensor fusion considering behavior-environment integration. *Energy* **2024**, *299*, 131396. [\[CrossRef\]](#)
80. Anik, S.M.H.; Gao, X.; Meng, N. Automation in Building Occupant Profile Development: A Machine Learning- and Persona-Enabled Approach. In Proceedings of the Construction Research Congress 2024: Advanced Technologies, Automation, And Computer Applications in Construction, Des Moines, IA, USA, 20–23 March 2024; pp. 41–49.
81. Huang, Z.; Gou, Z. Occupancy and equipment usage prototype schedules for building energy simulations of office building types in China. *J. Build. Perform. Simul.* **2025**, *18*, 56–75. [\[CrossRef\]](#)
82. Yi, H. Visualized Co-Simulation of Adaptive Human Behavior and Dynamic Building Performance: An Agent-Based Model (ABM) and Artificial Intelligence (AI) Approach for Smart Architectural Design. *Sustainability* **2020**, *12*, 6672. [\[CrossRef\]](#)
83. Roccotelli, M.; Rinaldi, A.; Fanti, M.P.; Iannone, F. Building Energy Management for Passive Cooling Based on Stochastic Occupants Behavior Evaluation. *Energies* **2021**, *14*, 138. [\[CrossRef\]](#)
84. Xu, Z.; Wei, W.; Jin, W.; Xue, Q.-R. Virtual drill for indoor fire evacuations considering occupant physical collisions. *Autom. Constr.* **2020**, *109*, 102999. [\[CrossRef\]](#)
85. Zhu, R.; Becerik-Gerber, B.; Lin, J.; Li, N. Behavioral, data-driven, agent-based evacuation simulation for building safety design using machine learning and discrete choice models. *Adv. Eng. Inform.* **2023**, *55*, 101827. [\[CrossRef\]](#)
86. Derbas, G.; Voss, K. Data-driven occupant-centric rules of automated shade adjustments: Luxembourg case study. In Proceedings of the Carbon-Neutral Cities—Energy Efficiency and Renewables in The Digital Era (CISBAT 2021), Lausanne, Switzerland, 8–10 September 2021.
87. Belafi, Z.; Hong, T.; Reith, A. Smart building management vs. intuitive human control-Lessons learnt from an office building in Hungary. *Build. Simul.* **2017**, *10*, 811–828. [\[CrossRef\]](#)
88. Naylor, S.; Gillott, M.; Herries, G. The Development of Occupancy Monitoring for Removing Uncertainty within Building Energy Management Systems. In Proceedings of the 2017 International Conference on Localization and GNSS (ICL-GNSS), Nottingham, UK, 27–29 June 2017.
89. Jia, M.; Srinivasan, R.S.; Ries, R.; Bharathy, G. Exploring the Validity of Occupant Behavior Model for Improving Office Building Energy Simulation. In Proceedings of the 2018 Winter Simulation Conference (WSC), Gothenburg, Sweden, 9–12 December 2018; pp. 3953–3964.
90. Sonta, A.J.; Simmons, P.E.; Jain, R.K. Understanding building occupant activities at scale: An integrated knowledge-based and data-driven approach. *Adv. Eng. Inform.* **2018**, *37*, 1–13. [\[CrossRef\]](#)
91. Hosamo, H.H.; Nielsen, H.K.; Kraniotis, D.; Svennevig, P.R.; Svidt, K. Improving building occupant comfort through a digital twin approach: A Bayesian network model and predictive maintenance method. *Energy Build.* **2023**, *288*, 112992. [\[CrossRef\]](#)
92. Goldsworthy, M. Towards a Residential Air-Conditioner Usage Model for Australia. *Energies* **2017**, *10*, 1256. [\[CrossRef\]](#)
93. Zhang, L.Z. Occupant-Aware Energy Management: Energy Saving and Comfort Outcomes Achievable Through Application of Cooling Setpoint Adjustments. Master's Thesis, University of Southern California, Los Angeles, CA, USA, 2017.
94. Ma, J.H.; Erdogmus, E.; Cha, S.H. Integration of a choice modeling approach with immersive virtual environments for accurate space utilization prediction. *J. Build. Eng.* **2023**, *76*, 107126. [\[CrossRef\]](#)
95. Chu, M.L.Z. A Computational Framework Incorporating Human and Social Behaviors for Occupant-Centric Egress Simulation. Ph.D. Thesis, Stanford University, Stanford, CA, USA, 2015. Available online: <https://purl.stanford.edu/jw835rf0798> (accessed on 17 October 2025).
96. Gilani, S.; O'Brien, W.; Gunay, H.B.; Carrizo, J.S. Use of dynamic occupant behavior models in the building design and code compliance processes. *Energy Build.* **2016**, *117*, 260–271. [\[CrossRef\]](#)
97. Schaumann, D.; Breslav, S.; Goldstein, R.; Khan, A.; Kalay, Y.E. Simulating use scenarios in hospitals using multi-agent narratives. *J. Build. Perform. Simul.* **2017**, *10*, 636–652. [\[CrossRef\]](#)
98. Zhou, H.; Yu, J.; Zhao, Y.; Chang, C.; Li, J.; Lin, B. Recognizing occupant presence status in residential buildings from environment sensing data by data mining approach. *Energy Build.* **2021**, *252*, 111432. [\[CrossRef\]](#)
99. Alzahrani, H.; Arif, M.; Kaushik, A.K.; Rana, M.Q.; Aburas, H.M.M. Evaluating the effects of indoor air quality on teacher performance using artificial neural network. *J. Eng. Des. Technol.* **2023**, *21*, 604–618. [\[CrossRef\]](#)
100. Çağlayan, İ.Z. Exploring Energy-Related Occupant Behavior in Office Buildings: An Interdisciplinary Study in the Context of Building Physics and Social Psychology. Ph.D. Thesis, Bilkent University, Ankara, Turkey, 2024.
101. Wu, Y.; An, J.; Qian, M.; Yan, D. Application-driven level-of-detail modeling framework for occupant air-conditioning behavior in district cooling. *J. Build. Eng.* **2023**, *70*, 106401. [\[CrossRef\]](#)

102. Uddin, M.N.; Lee, M.; Ni, M. The impact of socio-demographic factors on occupants' thermal comfort and sensation: An integrated approach using statistical analysis and agent-based modeling. *Build. Environ.* **2023**, *246*, 110974. [\[CrossRef\]](#)
103. Barone, G.; Buonomano, A.; Forzano, C.; Giuzio, G.F.; Palombo, A.; Russo, G. A new thermal comfort model based on physiological parameters for the smart design and control of energy-efficient HVAC systems. *Renew. Sustain. Energy Rev.* **2023**, *173*, 113015. [\[CrossRef\]](#)
104. Uddin, M.N.; Chi, H.-L.; Wei, H.-H.; Lee, M.; Ni, M. Influence of interior layouts on occupant energy-saving behaviour in buildings: An integrated approach using Agent-Based Modelling, System Dynamics and Building Information Modelling. *Renew. Sustain. Energy Rev.* **2022**, *161*, 112382. [\[CrossRef\]](#)
105. Osman, M.; Ouf, M. How Do Household Characteristics Affect Urban Occupancy Schedules? A Case Study Using Canadian Time Use Survey Data. *Ashrae Trans.* **2021**, *128*, 3–12.
106. Hassanpour, S.; Gonzalez, V.A.; Zou, Y.; Liu, J.; Wang, F.; Castillo, E.d.R.; Cabrera-Guerrero, G. Incorporation of BIM-based probabilistic non-structural damage assessment into agent-based post-earthquake evacuation simulation. *Adv. Eng. Inform.* **2023**, *56*, 101958. [\[CrossRef\]](#)
107. Tinaburri, A. Principles for Monte Carlo agent-based evacuation simulations including occupants who need assistance. From RSET to RiSET. *Fire Saf. J.* **2022**, *127*, 103510. [\[CrossRef\]](#)
108. Nimlyat, P.S. Indoor environmental quality performance and occupants' satisfaction [IEQ_{POS}] as assessment criteria for green healthcare building rating. *Build. Environ.* **2018**, *144*, 598–610. [\[CrossRef\]](#)
109. Li, D.; Xu, X.; Chen, C.-F.; Menassa, C. Understanding energy-saving behaviors in the American workplace: A unified theory of motivation, opportunity, and ability. *Energy Res. Soc. Sci.* **2019**, *51*, 198–209. [\[CrossRef\]](#)
110. Abdallah, M.; Cleverger, C.; Tam, V.; Anh, N. Sensing Occupant Comfort Using Wearable Technologies. In Proceedings of the Construction Research Congress 2016: Old and New Construction Technologies Converge in Historic San Juan, San Juan, Puerto Rico, 31 May–2 June 2016; pp. 940–950.
111. Youssef, A.; Caballero, N.; Aerts, J.-M. Model-Based Monitoring of Occupant's Thermal State for Adaptive HVAC Predictive Controlling. *Processes* **2019**, *7*, 720. [\[CrossRef\]](#)
112. Jayathissa, P.; Quintana, M.; Abdelrahman, M.; Miller, C. Humans-as-a-Sensor for Buildings-Intensive Longitudinal Indoor Comfort Models. *Buildings* **2020**, *10*, 174. [\[CrossRef\]](#)
113. Buonomano, A.; Forzano, C.; Gnecco, V.M.; Pigliautile, I.; Pisello, A.L.; Russo, G. Enhancing energy efficiency and comfort with a multi-domain approach: Development of a novel human thermoregulatory model for occupant-centric control. *Energy Build.* **2024**, *303*, 113771. [\[CrossRef\]](#)
114. Tan, C.Y.M.; Rahman, R.A.; Lee, Y.S. Modelling the WELL building concepts for office environments: PLS-SEM approach. *J. Eng. Des. Technol.* **2025**, *23*, 618–639. [\[CrossRef\]](#)
115. Donkers, A.; Yang, D.; de Vries, B.; Baken, N. Personal indoor comfort models through knowledge discovery in cross-domain semantic digital twins. *Build. Environ.* **2025**, *269*, 112433. [\[CrossRef\]](#)
116. Gnecco, V.M.; Vittori, F.; Pisello, A.L. Digital twins for decoding human-building interaction in multi-domain test-rooms for environmental comfort and energy saving via graph representation. *Energy Build.* **2023**, *279*, 112652. [\[CrossRef\]](#)
117. Li, X.; Yao, R. A machine-learning-based approach to predict residential annual space heating and cooling loads considering occupant behaviour. *Energy* **2020**, *212*, 118676. [\[CrossRef\]](#)
118. Liu, X.; Gou, Z. Occupant-centric HVAC and window control: A reinforcement learning model for enhancing indoor thermal comfort and energy efficiency. *Build. Environ.* **2024**, *250*, 111197. [\[CrossRef\]](#)
119. Kang, X.; Wu, Y.; Yan, D.; Zhu, Y.; Yao, Y.; Sun, H. A novel approach for occupants' horizontal and vertical movement modeling in non-residential buildings using Immersive Virtual Environment (IVE). *Sustain. Cities Soc.* **2022**, *87*, 104193. [\[CrossRef\]](#)
120. Vollmer, M.; Langer, M.; Banihashemi, F.; Harter, H.; Kierdorf, D.; Lang, W. Prediction of window handle state using machine learning. *Bauphysik* **2020**, *42*, 352–359. [\[CrossRef\]](#)
121. Sonta, A.J.; Jain, R.K. Inferring Occupant Ties Automated Inference of Occupant Network Structure in Commercial Buildings. In *Proceedings of the Buildsys'18: Proceedings of the 5th Conference on Systems for Built Environments, Shenzhen, China, 7–8 November 2018*; pp. 126–129.
122. Hou, H.; Pawlak, J.; Sivakumar, A.; Howard, B.; Polak, J. An approach for building occupancy modelling considering the urban context. *Build. Environ.* **2020**, *183*, 107126. [\[CrossRef\]](#)
123. Li, J.; Yu, Z.; Haghighat, F.; Zhang, G. Development and improvement of occupant behavior models towards realistic building performance simulation: A review. *Sustain. Cities Soc.* **2019**, *50*, 101685. [\[CrossRef\]](#)
124. Gaetani, I.; Hoes, P.-J.; Hensen, J.L.M. A stepwise approach for assessing the appropriate occupant behaviour modelling in building performance simulation. *J. Build. Perform. Simul.* **2020**, *13*, 362–377. [\[CrossRef\]](#)
125. Liu, Y.; Zhou, Y.; Yang, L.; Xin, Y. Simulating staff activities in healthcare environments: An empirical multi-agent modeling approach. *J. Build. Eng.* **2024**, *84*, 108580. [\[CrossRef\]](#)

126. Dorrah, D.H.; Marzouk, M. Integrated multi-objective optimization and agent-based building occupancy modeling for space layout planning. *J. Build. Eng.* **2021**, *34*, 101902. [[CrossRef](#)]
127. Hammad, A.W.A. Minimising the Deviation between Predicted and Actual Building Performance via Use of Neural Networks and BIM. *Buildings* **2019**, *9*, 131. [[CrossRef](#)]
128. Belafi, Z.D.; Hong, T.; Reith, A. A library of building occupant behaviour models represented in a standardised schema. *Energy Effic.* **2019**, *12*, 637–651. [[CrossRef](#)]
129. Sinha, K.; Rajasekar, E. Thermal comfort evaluation of an underground metro station in New Delhi using agent-based modelling. *Build. Environ.* **2020**, *177*, 106924. [[CrossRef](#)]
130. Arslan, M.; Cruz, C.; Gin hac, D. Understanding Occupant Behaviors in Dynamic Environments using OBiDE framework. *Build. Environ.* **2019**, *166*, 106412. [[CrossRef](#)]
131. Mahdavi, A. The trouble with ‘HIM’: New challenges and old misconceptions in human information modelling. *J. Build. Perform. Simul.* **2021**, *14*, 611–618. [[CrossRef](#)]

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.