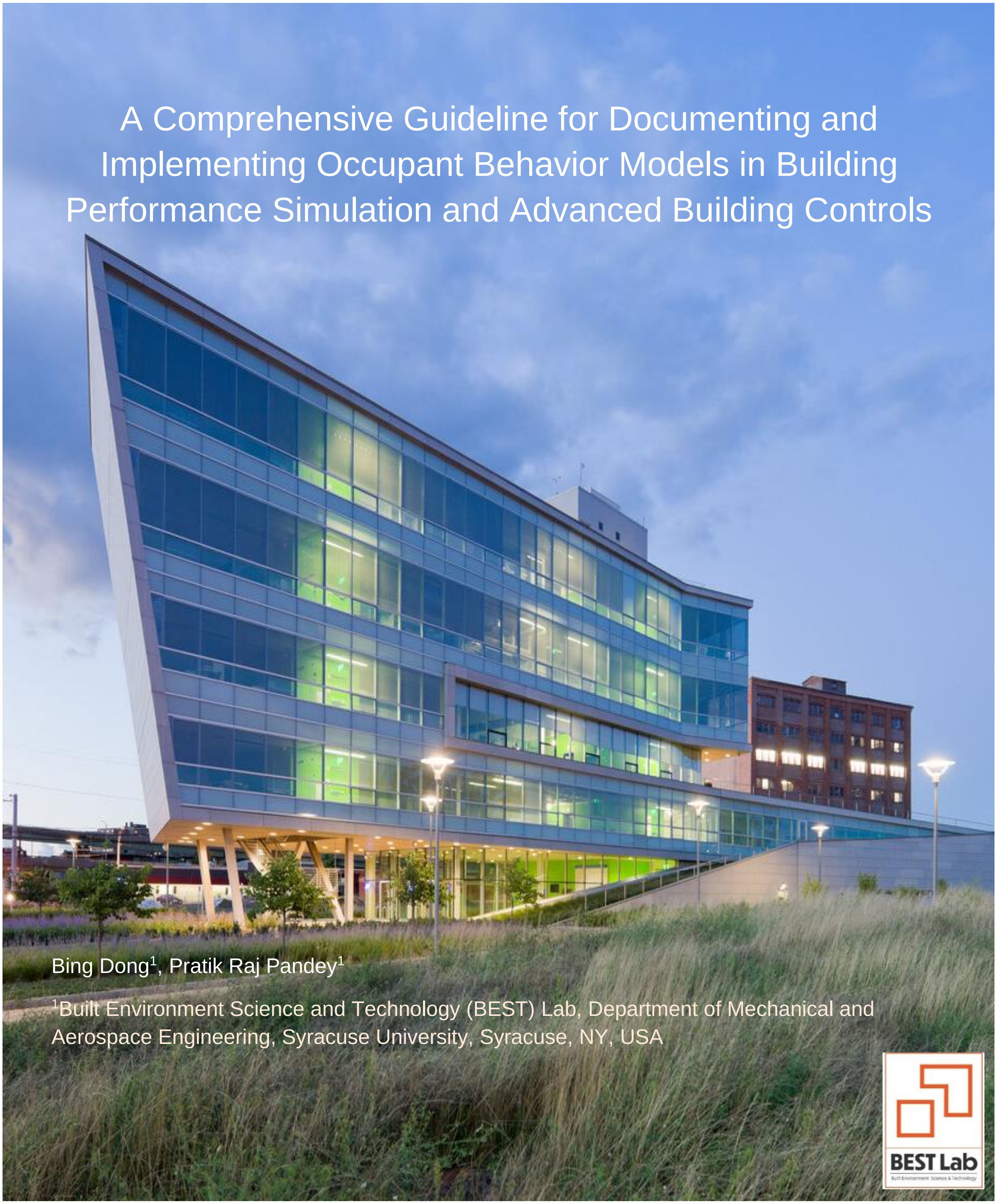

A Comprehensive Guideline for Documenting and Implementing Occupant Behavior Models in Building Performance Simulation and Advanced Building Controls

Bing Dong[1], Pratik Raj Pandey[1]

[1]Built Environment Science and Technology (BEST) Lab, Department of Mechanical and Aerospace Engineering, Syracuse University, Syracuse, NY, USA

Table of Contents

Abstract

Despite the advancements in sensor technologies, machine learning algorithms, and the broader Internet of Things (IoT), one critical component often remains neglected—the human element. Understanding Occupant Behavior (OB) is a complex but essential aspect of making smart systems truly "smart." It's not enough for buildings to be technologically advanced; they must also be capable of adapting to the needs and behaviors of the people who inhabit them. There have been many occupant behavior models developed for building design and controls, yet without a comprehensive framework or standard showing how those models are documented and implemented in fields. This document aims to fill this gap through:

1) developing a framework to document occupancy and occupant behavior models for building performance simulation, and emphasizing the importance of capturing the multidimensional aspects of human behavior. It consists of four blocks (description, development, evaluation, and implementation) and can be also regarded as a guideline to help researchers in the development, testing, implementation and transparent communication of their models.

2) developing a guideline to document occupant behavior models for advanced building controls by detailing how well-documented OB models can be operationalized to enhance building performance in real-time, and by presenting a model-evaluation schema that enables benchmarking of different models in field settings. Further, recommendations are given on how OB models are integrated into the building system. The framework is jointly developed by occupant behavior modelers and experts as part of the IEA EBC Annex 79 dedicated to occupant-centric building design and operation.

1. Introduction

1.1. Motivation

In the emerging field of intelligent building systems and smart cities, Occupancy and Occupant Behavior (OOB) and Occupant Behavior (OB) models play a critical role, serving as the backbone in both Building Performance Simulation (BPS) and Advanced Building Controls (ABC). As the research landscape expands, the gap between academic insight and real-world application is becoming increasingly evident. This disconnect primarily stems from the absence of standardized documentation and guidelines.

Standardized documentation acts like a universal language, simplifying the communication among researchers, developers, and practitioners. Without it, the reproducibility of scientific research is jeopardized. Researchers may find it challenging to build upon previous work, leading to inefficiencies such as duplicated efforts or the overlooking of promising avenues of research. This lack of a consistent framework can severely hamper the scientific process, causing delays in the development and refinement of OOB and OB models.

Moreover, inconsistent or absent documentation can severely undermine the confidence of practitioners and stakeholders, including architects, engineers, and policy-makers. If they cannot rely on a clear set of guidelines, it makes it difficult for them to assess the validity and utility of these models. This, in turn, affects their willingness to integrate such models into real-world projects, from individual intelligent building systems to larger smart city initiatives. This hesitancy is a significant barrier to the adoption of advanced technologies that could make buildings more efficient, sustainable, and responsive to human needs.

1.2. Standardization in OB Modeling and Documentation

The landscape of Occupancy and Occupant Behavior (OOB) and Occupant Behavior (OB) modeling is like a blend of pieces from various sets; they may be intriguing on their own but don't quite fit together. While there has been an influx of predictive data-driven OB models in the state-of-the-art, the missing link is a standardized framework for documenting these models. This omission makes it extremely challenging to plug these models into existing or emerging building control systems.

This lack of standardization has far-reaching implications. First, it throws a wrench into the replicability of models. Researchers coming onto the scene have to invest more time understanding each predecessor's unique approach, instead of focusing on innovation and improvement. This fragmented landscape is not just a bottleneck for academic progress but also a deterrent for newcomers to the field.

Additionally, practitioners in advanced building controls and Building Performance Simulations (BPS) face a similar issue. They may come across a potentially groundbreaking OB model but find themselves navigating a puzzle of inconsistencies, unclear methodologies, and absent documentation. This messiness undermines their confidence in adopting these models for real-world applications, which is counter-productive to the entire goal of advancing intelligent building systems and smart cities.

In essence, the absence of a unified documentation framework has become a significant roadblock, stalling both academic inquiry and practical application. It's like having cutting-edge tools but no manual on how to use them together; the potential is there, but the execution is hampered.

1.3. The Need for a Unified Approach

It's evident from prior studies that the lack of a unified approach for describing, developing, evaluating, and implementing OOB and OB models has resulted in their limited application in real-world settings. Researchers often struggle to communicate the nuances of their models transparently, making it difficult for other researchers to replicate or build upon their work. Practitioners are left with a fragmented understanding of which models are suitable for different applications, owing to inconsistent documentation practices. Thus, there is an unmet need for a comprehensive guideline that bridges this gap between academia and industry.

This led to the present guideline which is to serve as a comprehensive resource for documenting and implementing Occupant Behavior Models. The first part is specifically focusing on the Building Performance Simulation sector. The goal is to create a standard method for describing, developing, evaluating, and implementing Occupancy and Occupant Behavior (OOB) models. By standardizing these aspects, the guide aims to add rigor and transparency, making it easier for stakeholders to compare and implement these models in various simulation scenarios. The second part of the guideline delves into Advanced Building Controls. It provides directions on how to effectively document and implement OOB models that can be incorporated into real-time or predictive control systems. This section is designed to facilitate the integration of OB models into existing or new building automation systems, thereby enhancing their applicability in practical, real-world settings. Figure 1 shows the overall methodology used in the guideline document.

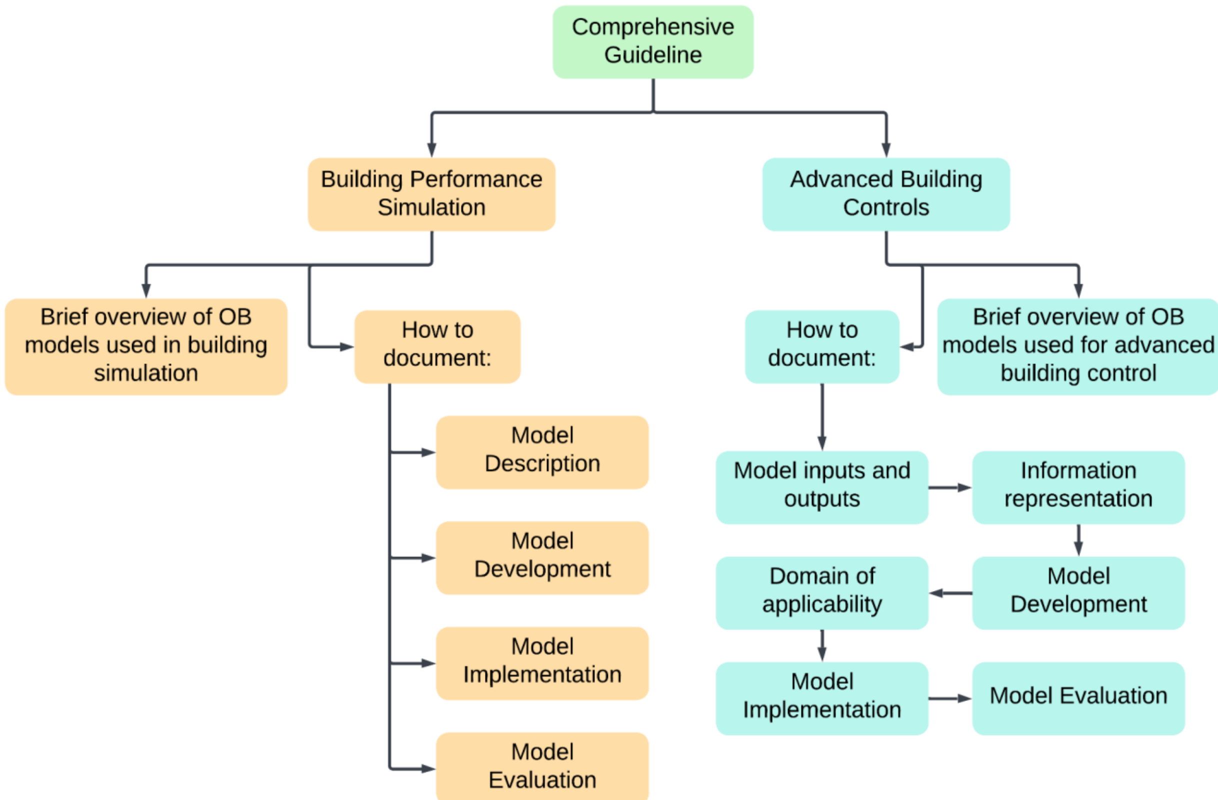

Figure 1: Flowchart describing methodology used in guideline document

2. Building Performance Simulation

2.1 Review of OB models for Building Performance Simulation

The importance of OB in building energy consumption is often ignored. OB is often oversimplified in Building Performance Simulation (BPS), leading to discrepancies between actual and simulated energy use. There are various OB model categories that are being used to improve the simulation:

Probabilistic or Stochastic Modeling: Includes Markov chain models, Bernoulli process, and survival analysis. These models are effective in capturing the dynamic nature of OB, particularly in long-term schedule formation.

Statistical Modeling: These models establish numerical relationships between OB and various parameters such as indoor and outdoor conditions. They are commonly used for light switching and window opening behaviors.

Data Mining Approaches: Useful for identifying patterns and information from large datasets. Techniques like decision trees, Bayesian networks, cluster analyses, and association rule mining are used to estimate and identify OB patterns.

Agent-Based Modeling (ABM): Involves independent actors ("agents") interacting based on predefined behavioral rules. ABM is effective in modeling diverse and dynamic energy usage trends among occupants.

There are various methods that can be employed to integrate these models to a simulation tool. Some of the methods include using user-defined profiles and rules approach, use customized functions or codes, and using co-simulation. In order to gather and evaluate information on description, a literature review was performed. In total 86 papers were selected for the review. There were very few studies published before 2005 but a noticeable increase was observed later.

In the domain of applicability, the literature review revealed significant gaps in the explicit delineation of both temporal and spatial scales. Starting with the temporal scale, nearly half of the studies (45%) do not specify the time-step at which their model operates. This omission hampers the ability to assess the model's adaptability to scenarios requiring different time resolutions. Additionally, more than half (57%) of the papers do not mention the temporal scale, which could range from seconds to years. This lack of information severely limits the model's potential for broader application and makes it difficult to evaluate its effectiveness in simulating real-world conditions over time. Moving to spatial scale, the data reveals a strong inclination towards smaller scopes. A significant 21% of the papers focus solely on a room-zone level, making these models less adaptable to comprehensive building or multi-building analyses. Similarly, 28% concentrate their efforts at the building level, which, while broader, still restricts the model's applicability to larger urban or district contexts. It is noteworthy that a meager 3.5% of the models venture into district or urban scale simulations. This limited foray into larger spatial contexts underscores a missed opportunity to explore more integrated, system-level interactions and their impact on building performance.

In terms of modeled outputs, most studies are singularly focused, with 67% covering only one output category like appliance use or window operation. Multi-category models are rare, accounting for just 33% of the total. These limitations suggest a compartmentalized approach that may overlook the interconnectedness of various occupant behaviors. The requirement for input variables shows an even more striking imbalance. Physical and contextual inputs dominate, being utilized in 60% of the reviewed models. However, less than 10% incorporate physiological or psychological parameters, which could be critical for more nuanced and realistic simulations.

In summary, the existing OB models in building performance simulation often suffer from a lack of specificity in purpose and applicability, a narrow focus on modeled outputs, and a limited range of input parameters. These gaps present both a challenge and an opportunity for future research to build more comprehensive and effective models.

2.2 Documentation of Occupant Behavior Models

2.2.1. Model Description

Introduction and Scope: Occupant Behavior (OB) models serve to capture and predict the actions and choices of individuals or groups in indoor environments. The primary objective of these models may range from enhancing energy efficiency and sustainability to optimizing occupant comfort and health. This section should clearly state the goals and application range of the model, be it residential, commercial, or special-purpose buildings.

Key Variables and Assumptions: Critical to any model's success is the identification and description of key variables such as occupant density, location, activity type, window operations, HVAC settings, and lighting conditions. The key input variables can be categorized into multiple categories identified in Fabi’s review [1]. For example:

Physical variables: Indoor air temperature, indoor transmitted solar radiation, outdoor relative humidity and rain.

Physiological variables: Occupant age, gender.

Psychological variables: Occupant habits and attitudes.

Contextual variables: (the collection of the inputs related to context): arrival/departure, socio-demographics, time of day, type of day.

The assumptions underpinning the model, such as constant temperature or humidity, should also be stated explicitly for clarity.

2.2.2. Model Development

Data Collection: The data used for model development and validation should be clearly identified and sourced. Information regarding whether the dataset is external or developed in-house should be disclosed. If external data are used for validation, the source should be cited and its relevance to the model explained. Data related to physical variables are collected using sensors. This includes temperature, humidity, pollutant, solar radiation etc. Other types of data can be collected using surveys.

Feature Selection: After data collection process, it is important to determine the features based on certain criteria as described below. These criteria for variable selection has been explained by Heinze [2]. First the “Significance criteria” should be determined. This is the most popular criteria for variable selection and includes hypothesis tests like Wald test’s p-value etc. Another criterion is “Change-in-estimate criterion”, which examines the relative percentage change in the parameters of the remaining variables when one variable is removed. Another feature selection criterion is selecting a model from a set of models rather than variable selection. This is referred to as “Information criteria”. Other criteria include “penalized likelihood” which uses the techniques like lasso regression and “background knowledge” where variables are selected using intuition derived from expertise in a particular domain.

Methodology: There are several algorithms, computational methods, or statistical tools used in model development including stochastic methods, Markov Chains, or Monte Carlo simulations.

Constraints and Limitations: The model limitations include but are not limited to data inconsistencies, computational power requirements, and any assumptions that might not hold universally. These constraints can often impact the model's generalizability and robustness. Constraints can be introduced by limiting the scope of the applicability of the model. Some of the constraints are as follows:

Spatial scale: The simulation spatial extension (room-zone, floor, district, urban) [3]

Spatial resolution: The zonal destination of the model (room, household, floor, building) [3]

Climate constraints: heating or cooling design models defined for particular months

Building type: The dominant function of the building (example: residential, office, retail, educational, dormitory)

2.2.3. Model Implementation

Computational Environment: Given that only 17% of reviewed articles specify this, it becomes crucial to document the computational environment in which the model operates. Whether it's EnergyPlus, IDA ICE, MATLAB, or Python, the choice of simulation tool and programming language should be clearly stated [4].

Integration with Building Systems: It is important to understand how the OB model will interact with existing building systems. Whether it’s a plug-and-play solution or requires manual calibration, the integration process should be clearly described to aid end-users and researchers alike.

Performance Specifications: Despite being neglected in most literature, computational specifications such as hardware requirements, runtime, and resource utilization should be provided for practical implementation and further development.

2.2.4. Model Evaluation

Metric Type: The chosen metrics for model evaluation play a pivotal role in determining the effectiveness of the model. These metrics can vary in nature. Some are direct, such as the number of window operations, offering a straightforward measure of specific actions or features within the model. Others are indirect, like resultant energy consumption, which provides insight into the broader impacts or outcomes of the model's performance. The selection of these metrics is critical in painting an accurate picture of how the model behaves and performs in various scenarios.

Evaluation Type: The evaluation process varies in its temporal scope, from detailed minute-by-minute analyses to broader seasonal assessments. It includes different methods of comparison: aggregated for a comprehensive overview, and interval-by-interval for a detailed, momentary analysis. This approach determines the model's performance across varying time frames and provides insights into its predictive accuracy.

Statistical Tools: In model evaluation, various statistical tools are used, including Pearson's correlation coefficient, RMSE, and confusion matrices. Each tool is chosen for its relevance to the model's goals, like measuring linear relationships or quantifying prediction errors, thereby ensuring a comprehensive assessment of the model's performance.

Feedback Challenges: Model feedback loops, wherein the model's output influences its subsequent input, are a prevalent challenge in OB modeling. Figuring out how the model accounts for or plans to mitigate this feedback issue is crucial.

These methods of model evaluation are based on the work of Mahdavi & Tahmasebi [5].

3. Advanced Building Controls

3.1. Review of OB models for advanced building controls

Occupant Behavior (OB) models serve as a critical component in the development of advanced building control systems. These models focus on a range of behaviors, broadly categorized into six main areas:

Appliance Use: Models predicting appliance use are often integrated with smart plugs and home automation systems. The control algorithms use predicted states to selectively power on or off appliances, optimizing for both energy usage and user convenience. Primary inputs like historical plug-load and occupancy are crucial for fine-tuning these automated controls [6], [7], [8].

Lighting Operation: The predictions made by lighting operation models are translated into control actions through lighting management systems. For example, automated dimming or color adjustment can be executed based on the model's prediction, using input data on illuminance levels and occupancy to optimize for energy efficiency and comfort [9], [10], [11].

Occupancy Estimation and Prediction: Control systems use these models to determine HVAC settings, lighting, and even security features. For instance, if the model predicts low occupancy for a specific time of day, the control system might reduce HVAC output or dim the lights to conserve energy. Inputs like motion detection and CO_2 levels are invaluable for these control adjustments [12], [13], [14].

Thermostat Adjustment: These models inform the control systems of optimal setpoints for indoor temperature, which is then automatically adjusted by the HVAC system. Inputs like real-time electricity pricing can even allow the control system to optimize for cost-efficiency in addition to occupant comfort.

Shading Operation: In automated building systems, shading models can be coupled with motorized blinds or tinting windows. Based on the predictions, the blinds could automatically adjust to let in more or less natural light, employing inputs such as temperature and illuminance levels for optimized control [15].

Window Operation: These models can be tied to automated window systems to control opening and closing mechanisms. Based on predictive data about temperature, humidity, and wind speed, the control system may decide to open windows for natural ventilation, thereby saving on HVAC costs [16].

In a fully integrated smart building, these models don't operate in isolation. They are often part of a centralized building management system (BMS) that takes multiple data streams as inputs to control various subsystems optimally. By synthesizing data from each of these occupant behavior models, the BMS can make more informed and nuanced decisions, enhancing both energy efficiency and occupant comfort in a synergistic manner.

3.2 How should OB models be documented?

3.2.1. Information representation

The selection of a data structure or schema, such as IFC, gbXML, or Brick, is a critical decision that impacts the efficacy of handling various building elements [17]. These tools are chosen for their capability to effectively manage static elements, like building geometry, as well as dynamic elements, such as time-series temperature data. However, it's important to recognize the limitations or gaps present in these ontologies. A notable area of concern is the representation of occupant behavior, which often lacks standardization. Furthermore, the use of metadata and terminology requires careful consideration. Specialized terms, particularly those related to occupant behavior, need clear definitions to ensure clarity and understanding. The naming conventions across different schemas are not always consistent, posing challenges for data integration and interoperability. This inconsistency underscores the need for a unified naming guideline, which could be proposed to harmonize data and enhance the universal applicability of the findings.

3.2.2. Model inputs and outputs

Outlining the different behavioral categories that the occupant behavior models are designed to address, such as Appliance Use, Lighting Operation, Occupancy Estimation and Prediction, Thermostat Adjustment, Shading Operation, and Window Operation is essential. It is important to provide a summary of the types of variables considered for each category: independent variables as inputs and dependent variables as outputs. Table 1 shows the commonly used inputs and outputs for various OB models.

Table 1: Inputs and outputs of various OB models

OB Model Category	Inputs	Outputs
Appliance use	Plug-load energy, Space's occupancy status	Multi-state of appliances, Energy consumption levels
Lighting operation	Illuminance levels, Occupancy status, Power consumption	State of lighting (binary/multi-state), Operation time
Occupancy Estimation and Prediction	Historical occupancy patterns, Motion detection, Power usage, Indoor environmental measures (illuminance, temperature, relative humidity, CO2, VOC levels)	Presence status (binary), Number of occupants
Thermostat adjustments	Indoor/outdoor temperatures and humidity, Solar radiation, CO2 levels, Hour of the day, Electricity load and price	Temperature setpoint setting, Indoor temperature, Probability of adjusting thermostat settings, Energy consumption
Shading Operation	Indoor/outdoor temperature, Illuminance, Solar radiation	Shading state (binary/multi-state), Probability of blinds position, Portion of blinds position

Window Operation	Indoor/outdoor temperature and humidity, Wind speed and direction, Solar radiation, Rainfall, CO2 and particulate matter concentrations	Probability of window state, Probability of action (e.g. opening/closing), Portion of a window open

It's also suggested to use figures to enumerate and categorize these variables. For each behavioral category, the researcher needs to specify the common inputs and outputs. It is essential to understand why these variables are relevant and frequently appear across various models. Same is true for each behavioral category. For example, we know that for Occupancy Estimation, a wide range of inputs like motion detection, historical occupancy patterns, and environmental measures are used to predict the presence or number of occupants. This variety should be clearly highlighted to provide readers with an understanding of the range of variables that are considered relevant for each type of model. By structuring the documentation in this manner, readers will gain a comprehensive yet organized understanding of the inputs and outputs typically involved in occupant behavior models. This will be invaluable for both understanding existing models and potentially developing new ones.

3.2.3. Domain of applicability

When documenting the domain of applicability, it is imperative to cover both temporal and spatial dimensions in which the OB models operate. For example, when occupancy models are applied to building controls in the temporal dimension, addressing three key aspects: "Time Granularity," "Predictive Horizon," and "Control Horizon" is crucial. These factors along with temporal and spatial domain are defined as follows:

Control Horizon for OBC: This refers to the time span over which the control actions are optimized. In OBC, the control horizon is the duration for which the system plans its control strategies, like adjusting temperature or lighting. For example, an HVAC system might have a control horizon of a few hours, planning its operations based on the expected occupancy during that period.

Predictive Horizon: This is the time frame over which the occupancy-based (OB) model makes predictions about the occupancy status. In simpler terms, it's how far into the future the system predicts whether a space will be occupied or not. This horizon can
vary widely. Some systems might predict just a few minutes to an hour ahead, while others might look 24 hours into the future, especially in settings with predictable occupancy patterns.

Temporal Domain: This encompasses the aspects of time in which the OB models operate, like the time granularity, predictive horizon, and control horizon. Temporal domain is critical because it defines the time-scale at which the system operates and reacts.

Spatial Domain: This refers to the physical space or the area in which the OB models are applied. It's about where in the building the controls are being implemented, like specific rooms, floors, or zones. The spatial domain can influence how the system behaves – for instance, controls in a rarely used conference room will differ significantly from those in a constantly occupied lobby.

Statistical evidence should be provided, akin to a figure or a chart, showcasing the distribution of this aspect across the various models reviewed. For example, note that in models focused on lighting operation, time granularity might be less frequently documented. Next, it is important to determine the "Predictive Horizon" which is defined as the time frame over which the OB model aims to make predictions. After pointing out the variability in this factor, which can range from less than an hour to up to 24 hours, it is essential to specify that it is often dependent on the variable being controlled. Once again, supplementing this information with statistical data to show the frequency with which this parameter is documented in existing literature is advised.

For a practical example, an office building using an OB model for its HVAC system is considered. The temporal domain might include a predictive horizon of up to 24 hours, based on the typical work schedule and meetings planned in the building. The time granularity could be in intervals of 15 minutes, adjusting the predictions based on real-time occupancy data. The control horizon might be set for 3 hours, optimizing the temperature and air quality for the expected occupancy during that time.

In the spatial domain, different zones of the building, like executive offices, general work areas, and meeting rooms, would have distinct control strategies based on their unique occupancy patterns. This spatial variability is key in tailoring the system for maximum efficiency and comfort. Documenting these domains with statistical evidence, such as figures or charts, will vividly illustrate the variations and commonalities among different OB models.

3.2.4. Model Development

When developing Occupant Behavior (OB) models, a thorough and structured approach is paramount. The process begins with data preparation, which involves cleaning the data to remove anomalies [18], filling in missing data using techniques like moving averages or regression methods, and scaling the data through normalization methods such as min-max scaling [19], distribution-based standardization [20] or structure-based techniques [21]. This ensures the integrity and consistency of the data for accurate modeling.

The next step is to select an appropriate model formalism. This includes rule-based models for well-defined patterns, statistical models like regression or Markov chains for capturing data trends, and data-driven models (such as neural networks) for complex behaviors with substantial datasets [22]. It's crucial to assess the chosen model's suitability for the specific building behavior, ensuring it accurately reflects occupant behavior dynamics. The implementation of the model involves developing algorithms or mathematical formulations tailored to the chosen model type and prepared data. A robust validation process, using real-world data or benchmark datasets, is essential to test and refine the model.

Finally, transparent documentation of each step in the model development process is necessary. This should include detailing data preparation methods, algorithm development, and the validation process. Acknowledging limitations and identifying gaps in the current methodology is vital, as it guides future research towards enhancing the model's applicability and performance. This comprehensive approach ensures that OB models are not only accurately reflective of real-world scenarios but also replicable and adaptable for future advancements in the field.

3.2.5. Model Evaluation

The evaluation of Occupant Behavior (OB) models is a critical phase following their development, where the model's performance is systematically assessed. This process is vital in ensuring the accuracy and reliability of the model in predicting occupant behavior in building environments. The evaluation primarily employs a variety of performance metrics, each serving a distinct purpose in assessing different aspects of the model's functionality.

Performance Metrics play a central role in this evaluation process. Absolute Metrics such as precision, recall, and the F1 score quantitatively assess the model's predictive accuracy, crucial in cases of data imbalance. F1 score is the harmonic mean of precision and recall. Precision is the measure of false positives whereas recall is the measure of false negative. Absolute metrics is also defined as “the metrics that are based on the absolute error calculation” [23]. For example, in a lighting control model, these metrics would determine how precisely the model predicts light usage or occupancy. Domain-Specific Metrics delve deeper, leveraging an understanding of occupant behavior and building physics. They provide insights beyond basic accuracy, assessing the model's practical applicability in real-world scenarios. For instance, in models predicting window operation, these metrics evaluate not just the accuracy of window state predictions, but also the frequency and duration of window operations. Indirect Metrics, on the other hand, measure the model's impact on broader building performance objectives like energy efficiency and thermal comfort, essentially gauging how changes in predicted occupant behavior patterns translate into tangible improvements in building operations.

The Evaluation Process involves rigorous statistical validation, where the model is tested against known datasets or through controlled experiments to confirm its robustness. Scenario testing is also crucial, where the model is applied to various settings to evaluate its adaptability to different environmental conditions and occupant behaviors. Comparing the model with existing benchmark models provides a relative understanding of its effectiveness and areas needing improvement. Finally, comprehensive documentation of each step in the evaluation, including the rationale behind chosen metrics and the results, is essential. This ensures transparency, replicability, and provides a foundation for further research and development in the field.

Through this structured and comprehensive approach, the evaluation of OB models not only verifies their accuracy but also their practical utility in real-world building environments, guiding improvements and innovations in future model development.

3.2.6. Model Implementation

Computational Environment: When documenting the implementation of OB models, especially in the context of building control systems, the computational environment should be specified at the beginning. It is important to be informed about the operating system (OS) on which the model was developed and tested. Furthermore, it is crucial to document information on the programming languages used and any libraries or software dependencies. Sometime the discrepancies in the versions of package can lead to errors and there are many modifications in methods. Thus, inclusion of version numbers for all of these elements will contribute to the reproducibility of the model.

Experimental Setup: In this segment, it is important to discuss the types of sensors that were employed, their implementation locations, and the rationale behind these choices. Table 2 shows types of sensors, its use and implementation location. Tables should be used to organize this information, making it easy for future building operators to understand and implement.

Table 2. Various sensor types and their use and implementation location

Sensor Type	Use in OB Model	Implementation Location
Temperature Sensors	Monitor indoor and outdoor temperatures for HVAC control models	Placed both indoors and outdoors to capture environmental variations
Occupancy Sensors	Predict occupancy patterns and lighting controls, including motion detectors and infrared sensors	Installed in strategic locations like entry/exit points and common areas
Light Sensors	Measure illuminance levels for lighting operation models	Positioned in areas with natural light variability
Humidity Sensors	Gauge indoor and outdoor humidity levels, influencing HVAC operation models	Distributed within and outside the building
CO2 Sensors	Assess indoor air quality and ventilation control	Located in areas with variable occupancy rates, such as meeting rooms or lounges
Energy Consumption Meters	Track power usage of appliances, lighting, and HVAC systems	Connected to key power-consuming systems for detailed energy usage data

Integration into Model Predictive Control (MPC): It is important to understand how the OB models integrate into MPC systems, particularly for HVAC control which is a common use case. It is vital to comprehend how the OB models contribute to setpoint/reference scheduling and how they can be used to include measurable and predictable disturbances. Also, it is imperative to be aware of how they shape constraints in the MPC, such as different upper and lower indoor air temperature bounds during occupancy hours. The requirements needed for the model's effective use in advanced optimal control methods like MPC should be clear. It is important to realize the necessity for the model to provide a forecast of occupancy behavior over the length of the prediction horizon, which is typically between 1-24 hours. Finally, the quality of presence and OB forecasts should be addressed while emphasizing that the most accurate predictions are obtained from dedicated sensor data such as PIR and cameras.

4. Conclusion

The comprehensive guideline outlined in this paper presents an approach to integrating Occupant Behavior (OB) models into Building Performance Simulation (BPS) and Advanced Building Controls (ABC). It addresses a critical gap in the field of intelligent building systems and smart cities, offering a unified framework for documenting and implementing these models. The guideline emphasizes the importance of considering human behavior in the design and operation of smart systems, ensuring that technological advancements in buildings are not just innovative but also adaptable and responsive to the people who inhabit them. The first part of the guideline focuses on developing a framework for documenting occupancy and occupant behavior models in building performance simulation. This framework comprises four essential blocks: description, development, evaluation, and implementation. By standardizing these aspects, the guideline enhances the rigor, transparency, and reproducibility of OB models, making it easier for stakeholders to compare and implement them in various simulation scenarios. The second part provides a detailed approach to documenting OB models for advanced building controls. It outlines how well-documented OB models can be operationalized to enhance building performance in real-time. The guideline introduces a model-evaluation schema for benchmarking different models in field settings and offers recommendations on integrating OB models into building systems. This ensures that the models are not only theoretically sound but also practical and applicable in real-world settings. The paper also reviews the existing OB models, highlighting the need for a more comprehensive approach that covers various aspects of human behavior and considers both temporal and spatial scales. It underscores the importance of having a diverse range of input parameters to capture the multidimensional aspects of human behavior accurately. In conclusion, this guideline serves as a vital resource for researchers, developers, and practitioners in the field of intelligent building systems. It bridges the gap between academic research and real-world application, enabling the development of smarter, more efficient, and occupant-centric buildings.

References

[1] V. Fabi, R. V. Andersen, S. Corgnati, and B. W. Olesen, “Occupants’ window opening behaviour: A literature review of factors influencing occupant behaviour and models,” *Build. Environ.*, vol. 58, pp. 188–198, Dec. 2012, doi: 10.1016/j.buildenv.2012.07.009.

[2] G. Heinze, C. Wallisch, and D. Dunkler, “Variable selection – A review and recommendations for the practicing statistician,” *Biom. J.*, vol. 60, no. 3, pp. 431–449, 2018, doi: 10.1002/bimj.201700067.

[3] B. Gunay, I. Beausoleil-Morrison, R. Goldstein, S. Breslav, and A. Khan, “Coupling stochastic occupant models to building performance simulation using the discrete event system specification formalism,” *J. Build. Perform. Simul.*, vol. 7, Nov. 2014, doi: 10.1080/19401493.2013.866695.

[4] T. Hong, Y. Chen, Z. Belafi, and S. D’Oca, “Occupant behavior models: A critical review of implementation and representation approaches in building performance simulation

programs," *Build. Simul.*, vol. 11, no. 1, pp. 1–14, Feb. 2018, doi: 10.1007/s12273-017-0396-6.
[5] A. Mahdavi and F. Tahmasebi, "On the quality evaluation of behavioural models for building performance applications," *J. Build. Perform. Simul.*, vol. 10, no. 5–6, pp. 554–564, Nov. 2017, doi: 10.1080/19401493.2016.1230148.
[6] Y. Jin *et al.*, "Appliance use behavior modelling and evaluation in residential buildings: A case study of television energy use," *Build. Simul.*, vol. 13, no. 4, pp. 787–801, Aug. 2020, doi: 10.1007/s12273-020-0648-8.
[7] K. X. Perez, M. Baldea, and T. F. Edgar, "Integrated smart appliance scheduling and HVAC control for peak residential load management," in *2016 American Control Conference (ACC)*, Jul. 2016, pp. 1458–1463. doi: 10.1109/ACC.2016.7525122.
[8] X. Zhang *et al.*, "IEHouse: A non-intrusive household appliance state recognition system," in *2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, Aug. 2017, pp. 1–8. doi: 10.1109/UIC-ATC.2017.8397510.
[9] Y. Jin, D. Yan, X. Zhang, J. An, and M. Han, "A data-driven model predictive control for lighting system based on historical occupancy in an office building: Methodology development," *Build. Simul.*, vol. 14, no. 1, pp. 219–235, Feb. 2021, doi: 10.1007/s12273-020-0638-x.
[10] D. Yan, X. Feng, Y. Jin, and C. Wang, "The evaluation of stochastic occupant behavior models from an application-oriented perspective: Using the lighting behavior model as a case study," *Energy Build.*, vol. 176, pp. 151–162, Oct. 2018, doi: 10.1016/j.enbuild.2018.07.037.
[11] T. Sanjeev Kumar, C. Kurian, and S. Shetty, "A data-driven approach for the control of a daylight–artificial light integrated scheme," *Light. Res. Technol.*, vol. 52, no. 2, pp. 292–313, Apr. 2020, doi: 10.1177/1477153519841104.
[12] B. Dong, D. Yan, Z. Li, Y. Jin, X. Feng, and H. Fontenot, "Modeling occupancy and behavior for better building design and operation—A critical review," *Build. Simul.*, vol. 11, no. 5, pp. 899–921, Oct. 2018, doi: 10.1007/s12273-018-0452-x.
[13] A. Mirakhorli and B. Dong, "Occupancy behavior based model predictive control for building indoor climate—A critical review," *Energy Build.*, vol. 129, pp. 499–513, Oct. 2016, doi: 10.1016/j.enbuild.2016.07.036.
[14] S. D'Oca and T. Hong, "Occupancy schedules learning process through a data mining framework," *Energy Build.*, vol. 88, pp. 395–408, Feb. 2015, doi: 10.1016/j.enbuild.2014.11.065.
[15] Z. Li, H. Zhu, B. Dong, and X. Xu, "Development of a systematic procedure to establish customized shading behavior identification model," *Energy Build.*, vol. 239, p. 110793, May 2021, doi: 10.1016/j.enbuild.2021.110793.
[16] S. D'Oca and T. Hong, "A data-mining approach to discover patterns of window opening and closing behavior in offices," *Build. Environ.*, vol. 82, pp. 726–739, Dec. 2014, doi: 10.1016/j.buildenv.2014.10.021.
[17] M. B. Kjærgaard *et al.*, "Current practices and infrastructure for open data based research on occupant-centric design and operation of buildings," *Build. Environ.*, vol. 177, p. 106848, Jun. 2020, doi: 10.1016/j.buildenv.2020.106848.
[18] J. Kim, S. Schiavon, and G. Brager, "Personal comfort models – A new paradigm in thermal comfort for occupant-centric environmental control," *Build. Environ.*, vol. 132, pp. 114–124, Mar. 2018, doi: 10.1016/j.buildenv.2018.01.023.

[19] M. Ashouri, B. C. M. Fung, F. Haghighat, and H. Yoshino, “Systematic approach to provide building occupants with feedback to reduce energy consumption,” *Energy*, vol. 194, p. 116813, Mar. 2020, doi: 10.1016/j.energy.2019.116813.
[20] X. Yu, S. Ergan, and G. Dedemen, “A data-driven approach to extract operational signatures of HVAC systems and analyze impact on electricity consumption,” *Appl. Energy*, vol. 253, p. 113497, Nov. 2019, doi: 10.1016/j.apenergy.2019.113497.
[21] S. D’Oca and T. Hong, “A data-mining approach to discover patterns of window opening and closing behavior in offices,” *Build. Environ.*, vol. 82, pp. 726–739, Dec. 2014, doi: 10.1016/j.buildenv.2014.10.021.
[22] S. Dabirian, K. Panchabikesan, and U. Eicker, “Occupant-centric urban building energy modeling: Approaches, inputs, and data sources - A review,” *Energy Build.*, vol. 257, p. 111809, Feb. 2022, doi: 10.1016/j.enbuild.2021.111809.
[23] M. Shcherbakov, A. Brebels, A. Tyukov, T. Janovsky, and V. Anatol, “A Survey of Forecast Error Measures,” 2013. Accessed: Jan. 04, 2024. [Online]. Available: https://www.semanticscholar.org/paper/A-Survey-of-Forecast-Error-Measures-Shcherbakov-Brebels/435eb6e05fcda264c8c1e7fbcdee7116bb5b1424