DTU Library

Introduction to Occupant Modeling

O’brien, William; Calì, Davide; De Simone, Marilena; Tabadkani, Amir; Azar, Elie; Rajus, Vinu Subashini; Agee, Philip; Schweiker, Marcel; Rysanek, Adam

Published in:
Occupant-Centric Simulation Aided Building Design

Link to article, DOI:
10.1201/9781003176985-6

Publication date:
2023

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
O’brien, W., Calì, D., De Simone, M., Tabadkani, A., Azar, E., Rajus, V. S., Agee, P., Schweiker, M., & Rysanek, A. (2023). Introduction to Occupant Modeling. In *Occupant-Centric Simulation Aided Building Design: Theory, Application, and Case Studies* (1st ed., pp. 104-144). Taylor & Francis. https://doi.org/10.1201/9781003176985-6

6 Introduction to Occupant Modeling

William O'Brien, Davide Calì, Marilena De Simone, Amir Tabadkani, Elie Azar, Vinu Subashini Rajus, Philip Agee, Marcel Schweiker and Adam Rysanek

Summary

In this chapter, we will provide an overview of occupant modeling, beginning with key definitions and a background on common occupant modeling approaches. Next, we will present more advanced modeling approaches, including data-driven stochastic models, agent-based models, and personas. Finally, we will discuss methods to implement occupant models into building performance simulation tools and methods to communicate occupant model characteristics.

6.1 Introduction

Computational modeling and simulation are powerful techniques to create a representation of buildings. In general, building performance modeling and simulation provide a deeper understanding of a given system to inform decision-making at any or all phases of the building life cycle, from early-stage design to operations and management. In the past two decades, occupant modeling has gained significant traction by researchers and practitioners due to the increasingly significant impact of occupants, interest in occupant well-being, and increased computational and simulation capabilities. Occupant modeling is a mathematical approach to characterize how people occupy and act in buildings. When integrated into building performance simulation (BPS), occupant modeling can be used to estimate how occupants might behave in buildings for a year or longer, and how building design and operation might affect occupants.

Ultimately, occupants can profoundly affect building performance relative to predictions. This impact has been evidenced in studies of architecturally similar spaces or buildings whose performance varies greatly as a result of their occupants (Dong *et al*., 2015; Iwashita and Akasaka, 1997). The so-called energy performance gap—the difference between predicted and measured energy use—tends to be even larger for high-performance buildings. For instance, as insulation levels and airtightness increase as a consequence of stricter regulations, occupants' control over building systems

DOI: 10.1201/9781003176985-6

and equipment will have higher relative effects on heat transfer and energy use (Carpino *et al*., 2017; Guerra Santin *et al*., 2009). Occupant control of windows and blinds can also significantly impact energy flows across the envelope (Hoes *et al*., 2009).

Failure to accurately characterize occupants in the building design process caries two risks: first, it may lead to a performance gap; second, and perhaps more critically, it may lead to poor design decisions (Gilani *et al*., 2016). For example, optimistic assumptions about how occupants will behave (e.g., in an energy-optimal way) or pessimistic assumptions about occupant density (e.g., very high values for HVAC equipment sizing) may lead to design decisions that impact a building's performance for life.

In the past decade, occupant modeling has been used extensively to support building design (discussed further in Chapter 8) and to close the gap between the predicted and actual energy performance (e.g., Goubran *et al*., 2021; Mahdavi *et al*., 2021). For example, occupant modeling can be used to assess the impact of occupant interactions with architectural features and technologies (e.g., adaptive facades) (Hong *et al*., 2017; O'Brien and Gunay, 2015; Luna-Navarro *et al*., 2020; Stopps and Touchie, 2021; Yan *et al*., 2015). Occupant modeling can also be used to design more comfortable and energy-efficient spaces and to avoid oversizing or undersizing equipment and spaces (e.g., O'Brien *et al*., 2019).

Aside from energy performance, occupant modeling can be used to better understand comfort and adaptive opportunities, such as adaptive facades, clothing, and thermostats (Deng and Chen, 2021). It can also be used to help develop strategies toward healthy indoor spaces, e.g., to control the transmission of COVID-19 and other pathogens (Li *et al*., 2021). Building models, for example, can be used in combination with various occupant scenarios to create profiles of individual heat exposure (Sailor *et al*., 2021) and analyses of occupant presence and behavior (Yan *et al*., 2021).

This first section introduces basic occupant modeling concepts and definitions, and subsequent sections delve into more details and more complex methods.

6.1.1 Occupancy and Occupant Behaviors

In this chapter and throughout the book, we distinguish between two major occupant characteristics: occupancy and behavior. Occupancy is used synonymously with presence and quantitatively defines the number of occupants or density of occupants in spaces. It can be defined as a binary state: occupied (at least one person present) or vacant (no occupants in space or building). It can also be distinguished by occupant types and groups (e.g., children, students, guests, staff). Accurate modeling of occupancy is important for estimating latent and sensible heat gains and air contaminant loads and to understand schedules and logic for controls and operations. Yet, one of the primary reasons to try to predict occupancy is to predict occupant behaviors and actions; except for cases of remote actions (e.g., smartphone-based thermostats), occupant presence is a necessary condition for actions to occur.

In contrast to occupancy, behaviors are actions that occupants take that affect building performance directly or indirectly (e.g., energy, indoor environmental conditions). In many instances, occupants are triggered to act by indoor environmental conditions (e.g., open a window in response to stale air). These are known as adaptive triggers. In turn, these behaviors affect indoor environmental quality (IEQ) and potentially building energy use. However, other behaviors (e.g., use of office and entertainment equipment) affect building performance but are not related to IEQ. These are known as non-adaptive triggers and may be a result of habits or tasks (e.g., occupant turns on computer when they arrive at work).

Occupant actions may be triggered by physical, physiological, psychological, or social phenomena. The relationship between triggers and actions is often moderated by contextual factors (e.g., office dress codes constrain opportunities to modify clothing levels) (O'Brien and Gunay, 2014). Figure 6.1 represents the relationships between actions, behaviors, and triggers (Schweiker *et al*., 2018).

6.1.2 Occupant Modeling Approaches

Following the terminology of Figure 6.1, occupants' presence and behavior can be modeled as actions (e.g., the action of turning on/off the heating/cooling system) and states (e.g., the state of light switch, state of windows opening, thermostat setpoint). An action changes the state, which then normally remains constant until a new action is taken, though interventions from mechanical and electrical systems may occur (e.g., overriding controls).

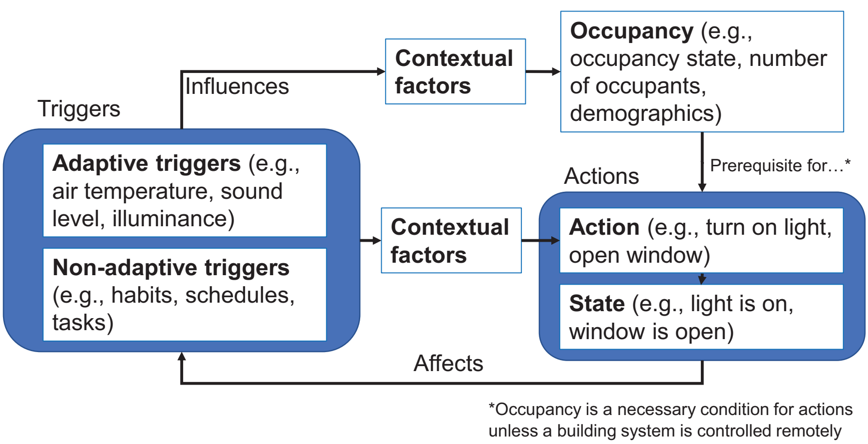

Figure 6.1 Relationships between actions, behaviors, and triggers in buildings.

A state can be defined by more than two levels and options, depending on the accuracy and targets of the modeling approach. For example, a window can have several states, including fully open/closed or half-open, percent open/closed, or lighting can be switched on/off or can include dimming (Schweiker *et al.*, 2018).

Ultimately, the objective of occupant behavior modeling is to predict either occupant actions/interactions with building systems or the resulting state of the building systems. It is generally accepted that accurately predicting individual actions is difficult, but predicting long-term trends is feasible if enough data is available to make generalized models. Defining generalizable predictors and model coefficients is challenging due to the diversity of available studies and the fact that many actions are contextually sensitive (e.g., climates, cultures, building types, systems) or differ for personal characteristics (Carlucci *et al.*, 2020; Schweiker and Shukuya, 2009). In addition, influencing variables for some domains, such as spatial movements or changes in body posture, can be important yet difficult to define and measure (Jakubiec and Reinhart, 2012; Schweiker *et al.*, 2018). Therefore, there is a strong need for researchers to collaborate on standard frameworks; this was one of the main motivations of initiating IEA EBC Annex 79 (O'Brien *et al.*, 2020).

While this book is broadly focused on buildings and building performance simulation, it should be noted that there are many other applications and domains for occupant modeling in the built environment. For example, human mobility and behavior modeling is of primary interest in scientific disciplines that explore topics such as evacuation in emergencies, pedestrian flow in public transportation, and motion in vehicles. Such models have the advantage of capturing occupants at the individual level while attaining realistic collective activities. For instance, people's velocities and buildings' structure (Lizhong *et al.*, 2003), occupants' health status and social influence (Liu *et al.*, 2020), and herd behaviors (Yang *et al.*, 2014) are some of the key factors affecting evacuation efficiency. Modeling frameworks to capture pedestrians' walking behaviors use a combination of concepts from the social force model, behavioral heuristics, and materials science (Porter *et al.*, 2018). Aircraft boarding models are implemented considering individual properties to explore the dynamics of passengers' motions (Tang *et al.*, 2012). With this background on occupant behavior and presence, the following section provides greater depth on traditional occupant modeling methods.

6.2 Traditional Occupant Modeling

To date, the dominant method to model occupants in building simulation is through relatively simple schedules, values, and simple rules. A survey of building simulation users indicated that the majority of them use occupant modeling approaches that are specified by building codes, in part to avoid the liability of making other assumptions that may prove to be incorrect

(O'Brien *et al*., 2016). In another study, practitioners were found to rely on default tool values, which also likely originated from codes and standards (Duarte *et al*., 2013). However, applying the same schedules and other values to all buildings neglects the impact of building design and people. This approach is akin to the way that weather files are imposed in building simulation, i.e., as a boundary condition; however, it fails to recognize that building design influences occupant behavior (O'Brien and Gunay, 2015).

It is relevant to precede our discussion of the state-of-the-art and future of occupant modeling by acknowledging *why* using schedules to represent occupants became a long-standing norm. The practice of assigning a single occupancy value to a modeled interior space for any simulated point in time dates back to at least the early 1980s and the first generations of building energy performance simulation tools, including but not limited to DOE-2 (Clarke, 2001; Diamond and Hunn, 1981; Norford, 1984; Vine *et al*., 1982). This was a time where 3D computer-aided design had yet to be introduced to the buildings industry. Interior building volumes simulated in the BPS tools of the day were prescribed numerically, using simplified metrics for building geometry such as wall area, window area, and interior volume. With respect to building heat transfer modeling, these volumes were represented as perfectly mixed indoor air spaces, with only a single value representing the air temperature within an interior volume at any given time. Similarly, internal heat gains, including occupancy, were represented as single point source loads, nominally determined by a user-assigned schedule as per the engineering manuals of the time (York and Cappiello, 1981). The location of an occupant in any simulated volume would be either fully non-spatial or located in an assumed fixed position of the floor space.

While the processing capabilities of computers today are worlds apart from the computers used in the early days of BPS, the legacy of this simplified approach to representing occupants and building geometry lives on. The same numerical methods DOE-2 used to represent occupants in its original source code remains engrained in the engineering of established, present-day BPS tools, such as EnergyPlus (the direct successor to DOE-2) (Crawley *et al*., 2001). Hence, it is common that users of BPS tools today specify similar time-based schedules and densities to represent building occupancy as would have been done by their predecessors 40 years ago. The term *diversit*y is often used to describe these schedules, in recognition that peaks are unlikely to occur simultaneously (e.g., an office might only have 80% of occupants at a given time, compared to maximum or nominal capacity). To reinforce the simplicity of common occupant modeling practice, Table 6.1 provides a summary of common methods to model different aspects of occupants based on the results of O'Brien *et al*. (2016) and O'Brien *et al*. (2020).

An example of a common modeling approach for occupancy is shown in Figure 6.2, where the occupancy density and schedule for numerous countries' energy code specifications are compared. These graphs show that

Table 6.1 Summary of commonly considered occupant-related domains and the corresponding modeling methods

Domain	*Common modeling approaches/assumptions*
Occupancy (presence)	Daily diversity schedules (hourly resolution) with a corresponding density (e.g., m^2 per occupant), usually specified for different building or space types
Plug-in equipment and appliances	Daily diversity schedules with a corresponding power density (e.g., watts per m^2)
Operable windows	Windows are closed
Lighting	Daily diversity schedules or daylight-controlled (otherwise turned on with occupancy) with a corresponding lighting power density (e.g., watts per m^2)
Window blinds/shades	Always open/non-existent (considered furnishing) or closed during glare events (e.g., above 1,000 lux, as per IES LM 83 [IESNA, 2012])
Water appliances (e.g., showers, toilets, sinks)	Hot water volume or energy per day per person or per floor area (e.g., L/person/day)
Thermostats	Daily setpoint schedules with the possibility to turn off systems or use a temperature setback for unoccupied and/or overnight periods
Clothing level	Seasonal schedule (e.g., 0.5 clo in summer and 1.0 clo in winter [ASHRAE, 2020])

typical occupancy modeling approaches are remarkable similar across different regions. They also show the inconsistency among different countries, suggesting a need for a global effort to standardize the way occupants are considered in building simulation.

A common question about occupant modeling approaches is where values and rules originated. Unfortunately, to date, data to support the development of occupant-related schedules has been obtained in a relatively dated and ad hoc way (e.g., "engineering judgment") (Abushakra *et al*., 2004; Deru *et al*., 2011; Duarte *et al*., 2013). O'Brien *et al*. (2020) reported that several building codes' occupancy density values have roots in non-energy applications, such as fire codes, which may be intentionally conservative. In the case of fire codes, for example, the relative risk of human safety is considered over the accuracy of energy estimates.

Aside from the challenge and importance of selecting appropriate schedule values to represent occupants, the schedule-based approach has fundamental problems. While these traditional occupant modeling methods are straightforward (e.g., mathematically simple), consistent (i.e., same results each run), and transparent (to the BPS tool user and stakeholders alike), they also have some drawbacks:

- **They lack recognition of two-way interactions between people and buildings**. The models assume occupants behave the same regardless of building

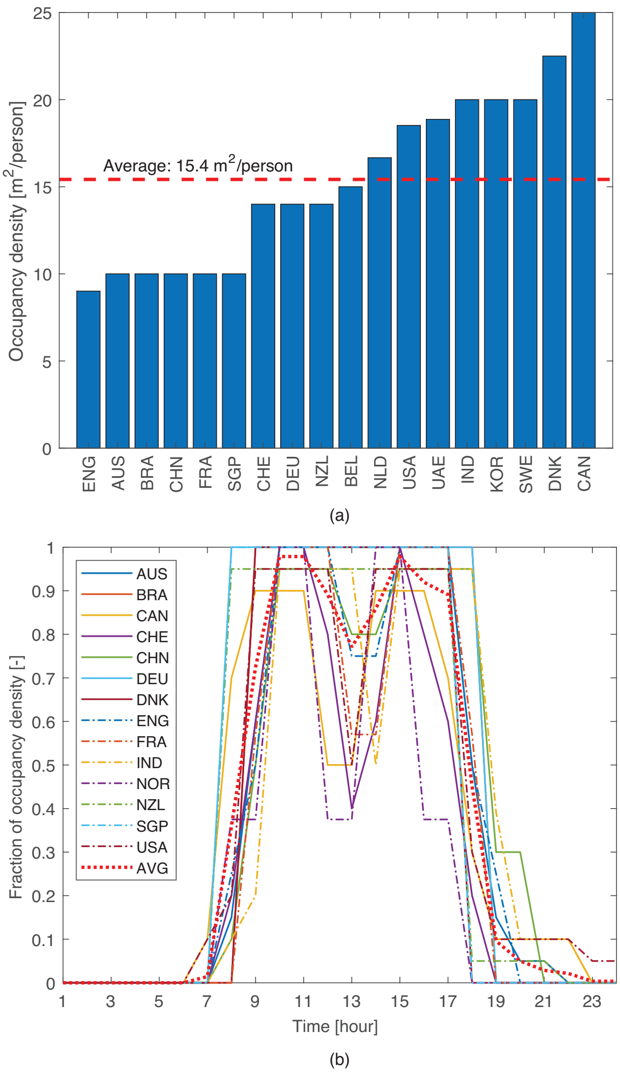

Figure 6.2 Example comparison of occupancy levels and schedules for office buildings in 15 different countries.

design. For example, they assume occupants control lights the same regardless of window geometry.

- **They are deterministic, which means that possible ranges of building performance and occupant behavior are not modeled**. They assume that every occupant behaves the exact same way for a given set of circumstances (e.g., all occupants turn on their light at a specific time of day).

- **They separate each occupant-related domain separately, without considering interdependencies**. For example, schedule-based models tend not to consider the linkages between occupant presence and adaptive actions (e.g., opening windows, turning on fans).
- **They are rather coarse and abstract, thus allowing practitioners to avoid deeply considering occupants**. Superficial occupant modeling does not require design practitioners to think about how building design can affect behavior (e.g., accessibility to and ease of opening windows).

These limitations have major implications for building design practice (see Chapters 8 and 11), and significantly limit the power of simulation-aided building design. Traditional methods are rooted in confirming or estimating building energy performance, rather than exploiting a better understanding of the two-way relationship between buildings and their occupants.

6.3 Advanced Occupant Modeling

In contrast to traditional methods of modeling occupants (see Section 6.2), more advanced occupant models tend to have one or more of the following possible and desirable traits (see Chapters 7 and 8 for additional discussion):

- **Stochastic:** A randomness to consider the reality that occupants' individual decisions are often diverse, unpredictable, and inconsistent. Stochastic modeling is used given that we cannot fully characterize, through any measurement, all the boundary conditions that might lead to a specific action. Moreover, there is unknown diversity among people and how they respond to current conditions, which means there is uncertainty about the specific individual occupants who will occupy a building.
- **Dynamic**: The recognition that conditions (e.g., air temperature) alter the way occupants behave and locate themselves within a space. In this way, the two-way relationship between occupants and buildings is characterized such that building design and operations can affect occupants' decision-making.
- **Data-driven:** The trait that occupant models are generated based on measurements. While most existing occupant modeling approaches are based on some measurements or observations, more advanced occupant models tend to use some form of model fitting (e.g., regression to relate behavior to one or more other variables).
- **Agent-based:** The acknowledgement that occupants interact with buildings and/or each other through a series of decisions that are likely a result of one or more conditions (e.g., IEQ, presence or behavior of others). While any of the methods described in Section 6.3 could be considered agent-based, the term is normally reserved for particularly sophisticated models (as discussed in Section 6.3.3).

In the following section, we provide an overview and mathematical details of some of the most common advanced occupant modeling approaches that include some or all of the above traits. We aim to provide an overview, coupled with key technical and mathematical details, and references where readers can seek greater detail. At the end of the section, we highlight two more advanced occupant modeling methods: agent-based modeling and personas.

6.3.1 Deterministic Models

Deterministic or non-probabilistic models are based on fixed values (e.g., an average and constant value for the internal gains in residential buildings) or schedules that are derived from assumptions or empirical observations, such as those described in Section 6.2. As argued in that section, such models offer ease of application, transparency, and reproducibility. However, they are independent of design and operations, and they typically do not capture uncertainty.

We should note that schedules and other non-probabilistic models could be made stochastic, though this is rare in practice. For example, schedules or densities could be stochastic (e.g., shape parameters randomly chosen from distribution) and data-driven (O'Brien *et al.*, 2019). They could also be customized based on a particular building design (Ouf *et al.*, 2019), or several clusters of occupant types with stochastic weightings could be used.

6.3.2 Stochastic Models

Probabilistic, or stochastic, models make use of stochastic processes to reproduce occupancy and a variety of behaviors, resulting in a probabilistic distribution of predicted results, from the timestep up to annual results. Several stochastic models have been used to reproduce human actions within buildings; in this chapter, we focus on four such models (described in each of the next four sections): binomial models, Markov chains models, hidden Markov chain models, and mixed effect models. Table 6.2 provides a general summary of the purpose and potential application of each model type. We describe each of the models in more depth in the sections that follow, with a focus on the models' application to the field of occupant modeling. For a more extensive explanation of the mathematics behind the different models, we recommend referring to more detailed sources (D'Oca *et al.*, 2019; Mahdavi *et al.*, 2017).

6.3.2.1 Binomial Model

A well-established statistical model used to both analyze and model binary dependent variables is the binomial model, often referred to as logistic regression[1] (Hastie and Tibshirani, 2017) when using the logit function as a link function. It can be used to model, for example, the state of a window

Table 6.2 Summary of four common occupant modeling approaches

Model type	*Typical purpose*	*Application*
Binomial model	Data analysis (e.g., to understand which factors influence occupants to execute an action) and stochastic modeling (e.g., to simulate human operations in building performance simulation software)	A model for predicting binary outcomes (e.g., yes/no, awake/asleep, open/closed, opening action/closing action)
Markov chains	Stochastic modeling with time dependencies (e.g., to model an event that is more likely to happen at a particular time of day, or a particular day of week)	A model for predicting outcomes with *n* states, where *n* can be an integer and represent—for example, specific locations in a building, occupant presence (e.g., present and awake, present and asleep, absent), or position of a window (open, half opened, closed), e.g., at different times of the day
Hidden Markov chain	Data analysis and stochastic modeling	A model for predicting outcomes with *n* unmeasured states, where the states are not measured but are deduced by related information (e.g., the presence of an occupant in a specific room, while only their activity is known)
Mixed effect	Data analysis and stochastic modeling	A model for predicting binary outcomes (see *binomial model regression*, above)

(e.g., closed or open) or the change of state of a window (e.g., from closed to open and vice versa) (Andersen *et al*., 2013; Calì *et al*., 2016).

Binomial models can be used for both analysis and predictive modeling purposes. For the former, it can be used, for instance, to understand the drivers (i.e., leading causes for change, e.g., Fabi *et al*., 2012) leading occupants to take an action. The results can provide researchers with background about how occupants make decisions depending on the indoor environment, weather, time of the day or day of the week, and/or any other measured entity. An illustrative example of binomial models is in Calì (2016), who applied binomial modeling with multiple explanatory variables to 300 monitored windows to generate 300 different models. For each window, the author determined which of the measured explanatory variables had a major influence on the probability of a change of window state and which did not. The variables were then classified depending on the number of times they appeared in the 300 models, where the more frequent the variable, the more important it was considered.

For modeling purposes, the binomial model can be used to dynamically model occupants (e.g., presence in a room, opening/closing a window) within a building simulation model. The model can be called at each timestep or at some selection of timesteps (e.g., only if an occupant is present in a room) and it reacts to the actual room conditions.

The binomial model using the logit function as a link function is based on the logistic function as expressed in Equation (6.1). $p(x)$ expresses the probability function for a certain event to happen (e.g., a window state changes) depending on an explanatory variable x, and, by definition, $p(x) \in [0,1]$, $\forall\ x$. Equation (6.1) can be rewritten as in Equation (6.2).

$$p = \frac{1}{1 + e^{-(\alpha + \beta x)}} \tag{6.1}$$

$$ln\left(\frac{p}{1-p}\right) = \alpha + \beta x \tag{6.2}$$

where α is the intercept, β is a coefficient, and x is the explanatory variable. Equation (6.1) describes the probability of a certain event (e.g., opening a window, switching off the heating system) depending on one explanatory variable (e.g., the outdoor temperature) and is therefore used for simple linear regression analysis. For regression analysis with n explanatory variables, the probability function p can be expressed as in Equation (6.3).

$$\ln\left(\frac{p}{1-p}\right) = \alpha + \beta_0 x_0 + \beta_1 x_1 + \ldots + \beta_n x_n \tag{6.3}$$

Andersen *et al.* (2013) suggested the inclusion of interaction terms in the probability function for some circumstances. That is, the probability of an action might depend on x_i at one level of x_j as compared to another level of x_j. For example, the probability of opening (or closing) a window, the coefficient β_i of the x_i explanatory variable at a certain period, e.g., in the morning, might differ from the coefficient β_i at a different period, e.g., at night. Also, there might be cases where an increase in the room air temperature might result in an increase in the probability of opening a window in the morning, and in a decrease in the probability of opening a window in the evening. Equation (6.4) can be used to include interaction terms (γ). It is good practice to use only interaction terms between continuous and categorical variables—time of day can be represented, for instance, in categorical variables such as morning, afternoon, evening, and night.

$$\ln\left(\frac{p}{1-p}\right) = \alpha + \beta_0 x_0 + \beta_1 x_1 + \ldots + \gamma_{1,2} x_1 x_2 + \ldots + \gamma_{1,n} x_1 x_n + \ldots + \gamma_{n-1,n} x_{n-1} x_n \tag{6.4}$$

As mentioned above, binomial models can be used to understand and model a state or a change of state. For occupants' use of building systems, binomial models can be used to model the action rather than the state (e.g., light switching action rather than on/off state). As noted by Fabi *et al.* (2012), the status of the window itself influences the indoor environment (hence the explanatory variables used for the modeling) and therefore affects the model.

Calì (2016) provides an example of the application of a binomial model for modeling occupant interactions with operable windows in a residential building. Figure 6.3 shows sample plots of the analysis, with the probability of the opening action of a specific window from a specific living room of a specific apartment, which was found to vary by time of the day, the indoor CO_2 concentration, and the indoor air temperature. The results suggest that window opening probability increases with indoor temperature and CO_2 concentration. Also, occupants are much more likely to open the window during the day than at night.

6.3.2.2 Markov Chain Models

This section describes discrete-time Markov chain models of the first order—henceforth, simply Markov chain models. Markov chain models are useful to model processes with two or more states, such as the position of a window (e.g., closed, open, half open) or the state of a fan (e.g., on or off, low flow, medium flow, high flow). When the state that has to be modeled is measured, a discrete-time Markov chain of the first order can be used (e.g., Calì *et al.*, 2018; Haldi and Robinson, 2009; McKenna *et al.*, 2015; Page *et al.*, 2008). Alternatively, when the state that has to be modeled is measured indirectly (e.g., the position of a window is inferred by the CO_2 concentration in the room, or the presence of occupants with one specific room is inferred based on a time use survey indicating only the activity of the occupants), hidden Markov models (see next section) can be used.

The paragraphs that follow include a brief description of the principles of Markov chain, inverse function sampling, and the Markov chain Monte Carlo technique. A deeper illustration of the Markov chain technique can be found in Feller (1968). The Markov chain Monte Carlo method is well described in (Gilks *et al.*, 1995).

To begin, a Markov chain is a random process that, within a state space, undergoes a transition from one state to another. The Markov property, which characterizes the Markov chain (illustrated in Equation (6.5)) states that the probability distribution of the next state (X_{n2}) depends on the current state (X_{n1}) and not on the events that preceded it. This property is also known as the memory-less property since the Markov process does not keep previous states in memory.

$$P\{X_{n1}|X_{n2}, X_{n3}, X_{n4}, X_{n5}, \ldots\} = P\{X_{n1}|X_{n2}\} \tag{6.5}$$

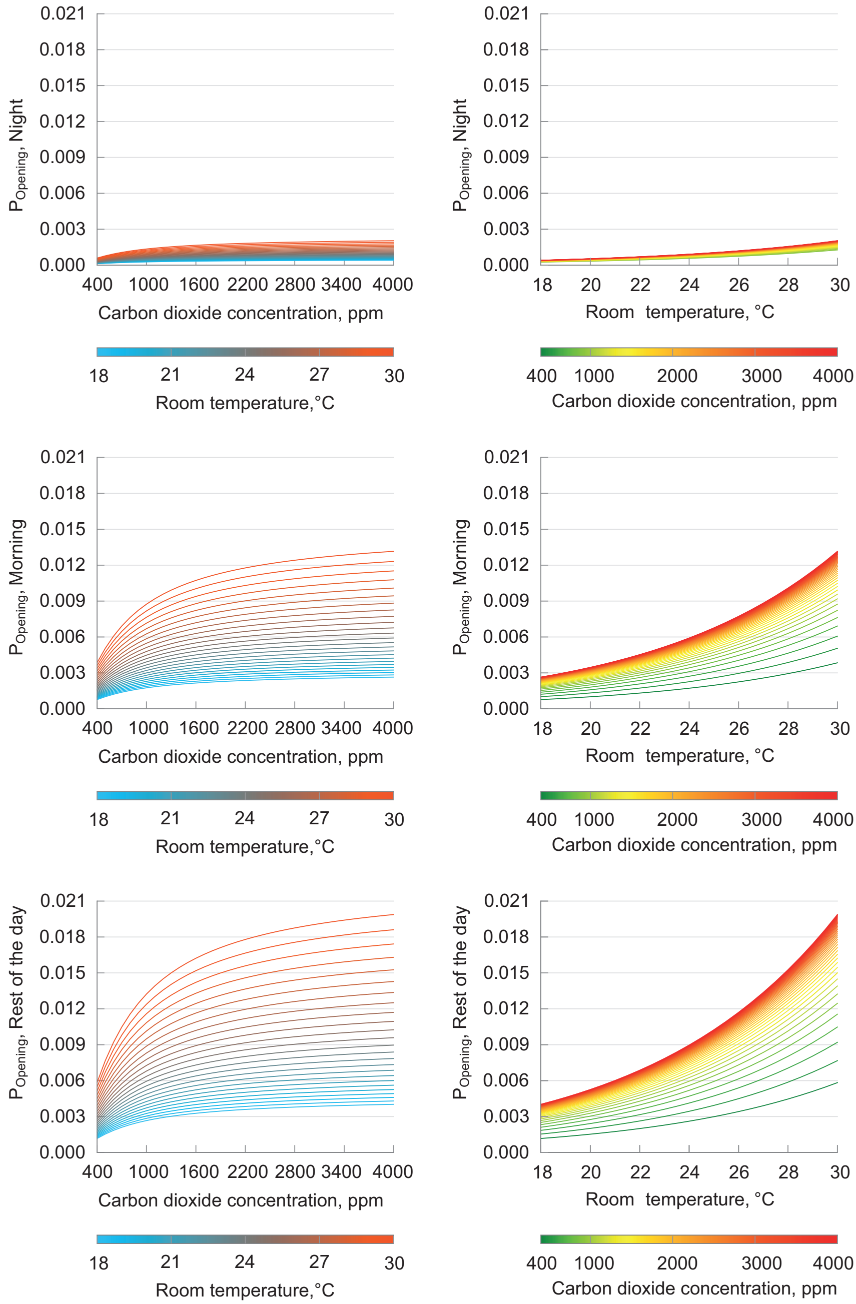

Figure 6.3 Probability of opening action of a window in a living room of a specific apartment (Calì, 2016) at three different times of day, within the next minute.

State changes in the process are called transitions; the probability of such transitions is stored in transition probability matrices (TPMs). When the transition probability does not depend on the time—and hence, the transition probability does not vary with the time—the Markov chain is called time-stationary or time-homogeneous, and this property is expressed in Equation (6.6).

$$P\{X_{n1}|X_{n2}\} = P\{X_{n2}|X_{n3}\} \tag{6.6}$$

Examples of time-homogeneous Markov chains include the "random walk" or the number of successes on bets by flipping a coin. As mentioned above, the presence or absence of occupant(s) as well as the state of a window can be modeled through Markov chains. However, in those cases, the probability of a change of the state (e.g., window opened/closed, occupant present/absent) varies over the time: in such cases, the Markov chain is time-inhomogeneous or simply inhomogeneous.

Equation (6.7) and Figure 6.4 show a two-state TPM for the status of a window at a given point in time: the state "0" indicates a closed window, while the state "1" indicates an open window; $S_{n,00}$ indicates the probability that a closed window (first 0) stays closed (second 0); $S_{n,01}$ indicates the probability that a closed window (0) will be opened (1); $S_{n,10}$ indicates the probability that an open window (1) will be closed (0); $S_{n,11}$ indicates the probability that an open window (first 1) stays open (second 1). For this particular example, at the given time n, there is a probability of $S_{n,00} = 0.95$ that the window remains closed if it was already closed at the preceding time n-1; in the case of the window being open at time $n - 1$, the probability that the window remains open is $S_{n,11} = 0.75$. The numbers in red in the figure represent the probability of a state change: $S_{n,01} = 0.05$ for a change from closed to open and $S_{n,10} = 0.25$ for a change from open to closed. The two dimensions illustrated in the example are related to the change of status for the time interval [n - 1, n].

Occupant behavior depends on time; for instance, occupants are more likely to sleep at night, windows are more likely to get opened in the morning,

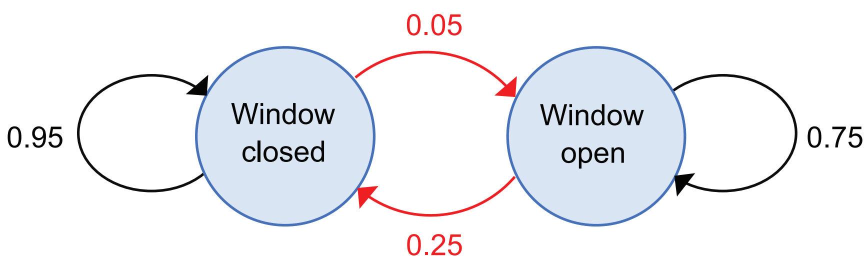

Figure 6.4 Two-state transition graphic for the TPM, at a given time instance, as shown in Equation (6.3) (Calì, 2016).

$$P_{n=1440} = \begin{pmatrix} s_{n,00} & s_{n,01} \\ s_{n,10} & s_{n,11} \end{pmatrix} = \begin{pmatrix} 0.98 & 0.02 \\ 0.11 & 0.89 \end{pmatrix} \quad n = 1440$$

$$\vdots$$

$$P_{n=n} = \begin{pmatrix} s_{n,00} & s_{n,01} \\ s_{n,10} & s_{n,11} \end{pmatrix} = \begin{pmatrix} 0.94 & 0.06 \\ 0.2 & 0.8 \end{pmatrix}$$

$$\vdots$$

$$P_{n=1} = \begin{pmatrix} s_{n,00} & s_{n,01} \\ s_{n,10} & s_{n,11} \end{pmatrix} = \begin{pmatrix} 0.93 & 0.07 \\ 0.21 & 0.79 \end{pmatrix} \quad n = 1$$

Figure 6.5 TPM over n states (Calì, 2016).

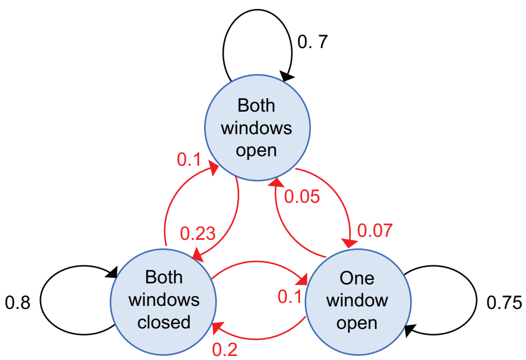

Figure 6.6 Three-state transition graphic for the TPM, at a given time instance, as given in Equation (6.4) (Calì, 2016).

and so on. Thus, the TPM needs a third dimension that allows probabilities of state change that vary over time. For a two-state process with transitions changing each minute during an entire day, the TPM shape will be 2 × 2 × 1,440 (Figure 6.5).

$$P_n = \begin{pmatrix} s_{n,\,00} & s_{n,\,01} \\ s_{n,\,10} & s_{n,\,11} \end{pmatrix} \tag{6.7}$$

There are cases where two states are not enough to model occupant behavior, such as when the goal is to model the presence of a number N of occupants or the opening and closing of a window with two movable panes. In the latter case, for instance, if a distinction between the panels is not necessary—for example because the two panels are the same size—a three-state Markov chain can be used. In the case of the panels being different sizes, a four-state Markov chain is necessary. Equation (6.8) and Figure 6.6 demonstrate an example of a three-state TPM for the status of a double-paneled window (with panels of equal size) at time n. In this example, "0" indicates a completely closed window, “1” indicates that one panel of the window is open, and “2” indicates that both panels are open.

As for the two-state TPM, for the three-state TPM, the sum of the values of each row at each time is equal to 1. In this way, the Markov chain does not stop within the simulation process.

$$P_n = \begin{pmatrix} s_{n,\,00} & s_{n,\,01} & s_{n,\,02} \\ s_{n,\,10} & s_{n,\,11} & s_{n,\,12} \\ s_{n,\,20} & s_{n,\,21} & s_{n,\,22} \end{pmatrix} \tag{6.8}$$

The generation of the TPMs can be done separately for each window or together for all windows based on the observed status changes of those windows at each measured time interval.

6.3.2.3 Hidden Markov Chain Models

Unlike the Markov chain model, the hidden Markov chain model (HMM) consists of two components: an unobserved Markov chain $\{X_t\}$ and an observed sequence $\{Y_t\}$. Y_t only depends on the current state X_t, and not on its own history $\boldsymbol{Y}^{(t-1)}$, as expressed in Equation (6.9).

$$P\left(Y_t|X_t,\, \boldsymbol{Y}^{(t-1)},\, \boldsymbol{X}^{(t-1)}\right) = P\left(Y_t|X_t\right) \tag{6.9}$$

The distribution of $Y_t|X_t$ is called response distribution. In an HMM, the parameters are given by the set $\{\{A,\, B,\, \pi\}\}$, where A corresponds to the TPM, B corresponds to the response distribution, and π corresponds to the distribution of the unobserved state of X_0 in the initial timestep. For the estimations of the parameters, the Baum-Welch algorithm can be used (Rabiner, 1989; Zucchini and MacDonald, 2009). The Baum-Welch algorithm is based on the maximum likelihood estimation principle. When dealing within the context of HMM, the most likely sequence of unobserved states for a given sequence of observations can be of interest. This sequence, called global decoding, can be efficiently calculated through the Viterbi algorithm. An example of an application of a hidden Markov chain for the generation of occupants’ presence profiles within buildings based on a time-use survey is provided by Wolf *et al.* (2019).

6.3.2.4 Mixed Effect Models

We previously described a generalized linear model (GLM; see Footnote 1), specifically a binomial model with the logit function as a link function, that describes the probability of an action for a specific window, in a specific living room, and in a specific apartment. Thus, if the goal is to use the GLM for simulating the performance of a building with a number X of apartments, each with a number Y of rooms, we will need to have $X{\cdot}Y$ models for the opening action and $X{\cdot}Y$ models for the closing action: one model for each window. Hence, within the simulation of the performance of a building, it will be difficult to choose among one of the many models, for each room and each apartment. Ideally, there would be a unique model able to address behavioral diversity.

A solution can be represented by the addition of a further predictor of random nature, x_k (McCulloch *et al.*, 2003; Pinheiro and Bates, 2006), following the approach proposed by Haldi (2013) and resulting in a generalized linear mixed model (GLMM), as demonstrated in Haldi *et al.* (2016) and O'Brien *et al.* (2017). An example of a mixed model is the mixed-effects logistic model defined as in Equation (6.10), where there is a fixed effect (like in Equation (6.3)) and a mixed effect. An application of this model to the case of window action, applied to residential and non-residential buildings from Germany, Denmark, and the United Kingdom is illustrated in Haldi *et al.* (2016).

$$\operatorname{logit}(p) = \beta_0 + b_0 + \sum_{k=1,\ldots,n} \left(\beta_k x_k + b_k x_k\right) \tag{6.10}$$

In conclusion, binomial models are GLMs that can be used to analyze and predict the probability of specific binary events, such as opening or closing a window or switching on or off a device. Markov models are particularly useful to model the probability of an action that is observable, the state of a window, if this state has been observed, and that varies with time. Hidden Markov models are useful to analyze and model the probability of an action that has not been observed (e.g., the state of the window) based on an observable variable (e.g., the carbon dioxide concentration in the room). Finally, generalized linear mixed models can be used to model the probability of a particular event, adding a mixed effect to represent the differences among the population (e.g., different apartments, different occupants).

6.3.2.5 Selection of Explanatory Variables

When addressing a modeling case with different potential explanatory variables, it is important to decide which explanatory variables (e.g., outdoor and indoor temperature and humidity, indoor carbon dioxide concentration) are relevant to evaluate and select the most appropriate model. Schweiker and Shukuya (2009) suggest using "forward" and "backward" selection of the variables for the regression models and scoring the models using the Akaike information criterion (AIC). This process allows the selection of a "best model" containing only the most important explanatory variables (i.e., variables that have a consistent impact on the probability function). Besides the

AIC, other criteria can be used, such as the Bayesian information criteria (BIC) (Schwarz, 1978).

The process for the selection of the best model can be executed by using the step function within the glm function in the statistical language R, with *n* explanatory variables. This process is described as follows:

1. Each coefficient of each variable is fitted by the regression model in a single variable model, and the related AIC is computed for each fit;
2. The variable with the lowest AIC is selected, and the model is fitted *n*-1 times with the selected variable and each of the *n*-1 remaining variables;
3. The model based on two variables with the lowest AIC is selected. Then, the AIC of this model is compared to the AIC of the best single-variable model (the single-variable model with the lowest AIC). Then:
   a. If the new model (two-variables model) had a consistently lower AIC, the process can go to step 4;
   b. Otherwise, the single-variable model is selected;
4. The previously excluded *n*—2 variables are then used to fit the model together with the two variables of the "two variables model" with the lowest AIC from step 3, in a "three variables model" (this is the so-called "forward selection"). Hence, from the three variables model, three two-variables models, obtained by dropping each of the variables recursively, are fitted (this is the so-called "backward selection"). Then:
   a. In the case that none of the three-variable or "new generated" two-variable models has a consistently lower AIC than the two-variables model with the lowest AIC from step 3, the model with the lowest AIC from step 3 is the final model,
   b. Otherwise, the process goes as in step 4, adding one more variable recursively.

6.3.2.6 Inverse Transform Sampling

The generated TPMs can be used to generate occupants' profiles within a simulation process. Within this scope, the so-called inverse transform sampling (ITS) or "inverse function method" is utilized to sample random numbers in Page *et al.* (2008). Through this method, sample numbers can be randomly generated from any probability distribution given its cumulative distribution function (cdf). For the case of windows or occupancy, a uniform distribution can be used. The first step of the ITS is related to the generation of a random number from a uniform distribution, between zero and one. Thus, the generated random number p is compared to the cdf in order to define the next state of the Markov chain. Using the window case as an example, if the generated value p is smaller than the probability of a state change of the window $P_{n+1,\ XX}$, at the given time, the window remains in the same state; otherwise, the window changes its state. Figure 6.7 is a flowchart

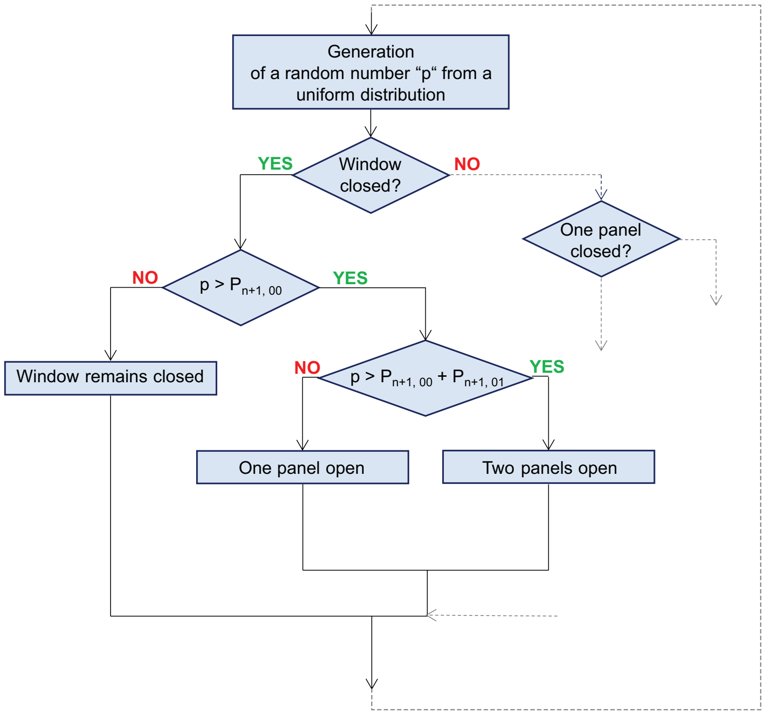

Figure 6.7 Inverse transform sampling Markov chain flow chart example (Calì, 2016).

of the simulation process (for the time instance "$n + 1$") of a double-paneled window, where both panels are closed at time n (Calì, 2016).

6.3.2.7 Evaluation and Validation of Occupant Models

Previously, we described a procedure to develop a model with an optimal selection of explanatory variables. Yet, the generated model needs to be evaluated and validated. It is generally understood that an exact prediction of building performance or each occupant-related event is impractical and unrealistic (e.g., timing of window opening actions). However, modelers can still strive for models that yield reasonable estimates and direct designers to near-optimal designs. This section briefly discusses occupant model evaluation and validation. Interested readers are encouraged to consult additional resources (Langevin *et al*., 2015; Mahdavi and Tahmasebi, 2017; Tahmasebi and Mahdavi; 2016).

The first quality of interest is that the model can reproduce the occupant actions or states for the building from which the data was first collected. In other words, the validation process relates to the model at hand and how well it predicts the behavior of a particular process (e.g., the act of opening a window) in a particular space or building. This validation should not be confused with a generalization of the model to other situations. For instance, this validation does not indicate the applicability of this model to other contexts (e.g., buildings, climates, occupant types). The second quality of interest is that the occupant model can predict occupant actions or states in other contexts. This has been proven to be more difficult to achieve since behavior can be sensitive to building technologies, local customs, and climates (Schweiker *et al.*, 2012).

One technique for validating of a model for a single space or building is the k-fold cross-validation. To apply it, each data sample (i.e., the set of data of the observed phenomenon that is being modeled and the potential explanatory variables) is partitioned into k-ordered subsamples. If k=10, for example, nine subsamples are used for training a model following the method described above and one subsample is used to test the model. The test of the model is done by using the measured input variables of the 10th subsample (i.e., the subsample that was not used for the training) as input to the model, thus comparing the model output with the actual, monitored change of window position. This operation is executed ten times in total; the ten combinations of nine out of ten subsamples are used to train the model, while the last subsamples is used each time as a validation subsample. The process is summarized in Figure 6.8 and described further in Calì (2016).

When creating a model, a validation process of the model should be undertaken to select the best possible model and ensure that the selected model is correctly representing the behaviors it is intended to portray. In the case of a model with a binary outcome (e.g., the change of state of a window from closed to open or from open to closed) to infer the "state change probability" (i.e., probability of opening or closing actions), the data sample should be partitioned into two subsamples:

1 Subsample A "window closed": This subsample is used to infer the probability that a window will change its status to open.
2 Subsample B "window open": This subsample is used to infer the probability that a window will change its status to closed.

The complete modeling process to achieve, as an example, a model describing the opening and closing operation of windows, is described in Figure 6.8.

To evaluate the applicability of an occupant model to another building, the model can be simulated in another context (e.g., climate, building design) to assess whether it accurately predicts occupant behavior in that building. The results may be compared on numerous metrics, such as a fraction of time when the state is correct and the number of actions per year (Mahdavi and Tahmasebi, 2017).

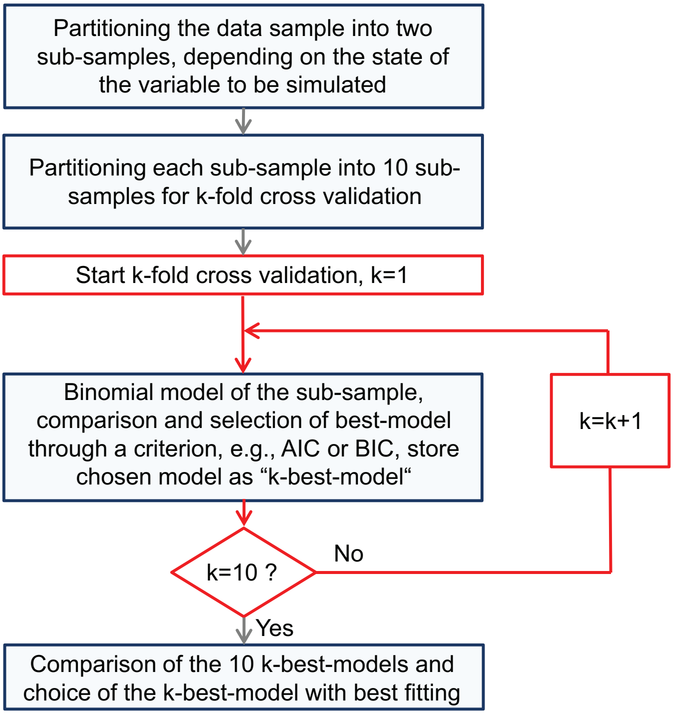

Figure 6.8 Flowchart of the process to generate and obtain the best fitting binomial model (Calì, 2016).

6.3.3 Agent-based Models

Building upon the previous sections and modeling techniques, agent-based modeling (ABM) is a technique capable of representing autonomous agents, their interactions with each other and their environment, and the resulting impact on the system as a whole (Gilbert, 2019). Agents (e.g., building occupants, households, cars) are assigned attributes that govern their interaction with each other and their environment (e.g., building space or geographical area). Each agent can evaluate the environment and the state of other agents and decide whether to take action (or not) based on a set of rules. The global behavior of the system then emerges from the micro-actions and interactions of these agents. The unique ability to simulate decision-making at the individual agent level enables ABM to simulate real-world systems with complex, nonlinear, and dynamic properties (Bonabeau, 2002).

Figure 6.9 shows the main steps to build an agent-based model, based on the work of Salgado and Gilbert (2013) and Sayama (2015). The core of the figure describes the ABM implementation stage following the specification

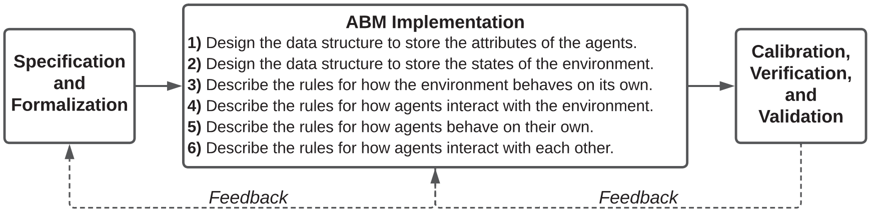

Figure 6.9 Main steps to building an agent-based model, adapted from Salgado and Gilbert (2013) and Sayama (2015).

and formalization of the problem to solve. The main implementation steps include the design of data structures for agents and the environment followed by steps to describe the behavioral rules and interactions (between agents and with the environment). Once the model is executed, calibration, verification, and validation efforts are performed with the model being revised as needed.

In terms of the programming environment, agent-based models can be implemented using general programming languages (e.g., Python, Java, C++) or software packages and toolkits created to help in the development and visualization of simulations (e.g., RePast, NetLogo, Anylogic).

ABM shares attributes with other modeling techniques, such as the probabilistic methods described in Section 6.3. For instance, it is common to define probabilistic rules that guide agents' actions based on information collected from their environment or other agents. As an example, when entering a shared office with uncomfortable thermal conditions, a "person" agent might interact with other nearby persons in the area (based on probabilistic rules) and adjust thermostat settings based on the group's preferences. Similarly, probabilistic rules could be used to model agents adapting their preferences (and behavioral rules) following interactions with other agents. In general, Bonabeau (2002) recommends the use of ABM when the real-world system to model has one or more of the following characteristics:

1. When interactions between agents exist and are of a complex, nonlinear, or discontinuous nature (e.g., the behavior of an agent potentially being influenced by that of another agent)
2. When the topology of interactions is heterogeneous (e.g., in social networks)
3. When space is an essential element of the problem with dynamic positions of agents (e.g., agents moving and interacting within an environment)
4. When the population of agents is heterogeneous (e.g., agents with different characteristics and adaptive behaviors)
5. When agents show complex behaviors (e.g., with learning and adaptation features)

ABM has been applied in numerous fields, including epidemiology (Tracy *et al.*, 2018), population dynamics (Pablo-Martí *et al.*, 2015), economics (Tesfatsion, 2002), transportation (Bernhardt, 2007), electricity grids (Ringler *et al.*, 2016), among others. In the past decade, ABM applications have been extended to cover the building science domain, particularly occupant behavior applications. Berger and Mahdavi (2020) reviewed scholarly articles that applied ABM to simulate building occupants for energy and indoor-environmental performance analysis. Papadopoulos and Azar (2016) presented an ABM framework that captured different and changing energy use characteristics of agents while accounting for their level of control over building systems. Their ABM framework also featured surrogate models of building systems to translate the agents' characteristics to building energy performance estimates. Lee and Malkawi (2014) proposed an ABM approach to mimic the behavior of real-world occupants of commercial buildings in response to environmental stimuli. After evaluating their indoor conditions, agents could increase their comfort levels by adjusting their clothing and activity levels or controlling building systems, such as windows, blinds, fans, and space heaters. Other applications of ABM include occupants' water consumption patterns (Linkola *et al.*, 2013), occupants' movements and shared activities (Schaumann *et al.*, 2017), and HVAC control optimization (Gopika, 2015; Sangi *et al.*, 2017).

Despite the many advancements in applying ABM to understanding and improving building performance, several limitations exist that motivate future work on the topic. First, most ABM applications in building studies are focused on understanding and improving building operation. Little research extends the scope of analysis to include occupant-centric design practice and applications. This gap is not limited to ABM studies but extends to occupant-centric building design research in general (Azar *et al.*, 2020). Second, current ABM studies often fail to provide information on the implementation of their models, particularly on the level of detail and resolution at which they modeled occupants' behaviors. More clarity and consistency are needed to determine the level of complexity needed to achieve the models' specific objectives. Finally, as highlighted by Berger and Mahdavi (2020), ABM applications are rarely based on robust and validated human behavior theories, which are needed to increase the levels of confidence in the developed models and their solutions. Future ABM studies should consider stronger theoretical underpinnings for agent rules and behaviors, in parallel to extensive observation studies for validation purposes.

6.3.4 Personas

The final method for modeling occupants that is considered in this chapter is personas. Personas are archetypal characters that are representative of the expected occupants. While the above mathematical formalisms in Section 6.3 are relatively abstract, the use of personas offers a promising approach to modeling occupant behavior and beliefs and group behavior in a more tangible and understandable by a wide range of stakeholders. Personas are fictional, but representative, characters capturing occupant characteristics, behavior, and goals (Cooper, 1999) for user-centered design. Personas can help designers and simulation users anchor their work in a user's needs (Takai and Ishii, 2010). Like the other modeling approaches described earlier in this chapter, personas can be either data-driven or developed by the designers' judgment. Personas are somewhat analogous to clusters in machine learning, in that representative agents are extracted from a population.

Personas can be fictional (Blythe and Wright, 2006), goal-oriented (Cooper *et al.*, 2014), role-based (Pruitt and Adlin, 2010), or engaging (Nielsen, 2013). Fictional personas may be imaginative or empirical. Goal-oriented personas focus on specific workflows, needs, motivations, and attitudes of the persona to accomplish their goals (e.g., save energy or improve thermal comfort) (Cooper *et al.*, 2014). Role-based personas assume the role the users play in their context and environment (Pruitt and Adlin, 2010). For example, in a large building context, personas may be developed for occupants and building energy managers. Engaging personas consider characters and stories to "produce involvement and insight" (Nielsen, 2013).

6.3.4.1 Past Use of Personas in Building Design and Simulation

While personas are widely employed in fields like human–computer interaction (HCI) and human-centered product design to anchor design in human needs from the beginning to the end of product development, their use in building design to represent different types of occupants is a relatively unexploited opportunity. Only recently, have examples of the application of personas in building design and operation emerged. For example, personas have been used to design spaces for people with dementia (McCracken *et al.*, 2019) and as a lens through which to evaluate the retrofit of buildings according to different behaviors, motivations, and attitudes (Haines and Mitchell, 2014). Bennetts *et al.* (2020) used a persona-based approach to create thermal guidelines for older people in Australia using hierarchical cluster analysis (HCA) on data collected from the participants (ideas, beliefs, knowledge, etc.). Unlike traditional comfort standards, the comfort guidelines were developed for six different thermal personalities (Bennetts *et al.*, 2020).

To date, personas have not been implemented as standard features in mainstream building simulation tools. Goldstein *et al.* (2010) used the schedule-calibrated and weighting coefficients method to generate personas for office buildings. The model considered office parameters such as arrival, departure, desk meetings, team meetings, and onsite and offsite breaks. This method helped create diverse occupant profiles for office buildings, but they did not consider other parameters like comfort and energy-related parameters.

6.3.4.2 Developing Personas for Building Simulation

Personas can be designed for particular building contexts (e.g., police stations, schools), comfort issues, and user types (e.g., older people, children). The data for these personas can be derived from literature, surveys, participatory workshops, among others. For new buildings, information can be estimated by looking to a similar project type, obtaining details from the client, or considering extreme conditions. Here, we focus on data-driven personas. We note that care must be taken to avoid unconscious bias/ discrimination when creating personas, as the associated implications may influence design and neglect certain populations of occupants (e.g., persons with disabilities).

For data-driven personas, the richer the data collection and analysis (e.g., mixed methods, methods that capture user-system context), the more useful the persona will be for designers and simulation users. An example of a data-driven persona is provided by Agee *et al.* (2021). Agee *et al.* (2021) collected both quantitative and qualitative data from 20 multifamily housing developments (representing 239 units) in Virginia, USA. Data were collected and analyzed in four steps, as summarized in Chapter 4, to create the persona in Figure 6.10.

Sadie
Senior Persona

Physical Needs: *safety, easy-to-access and understand spaces and interfaces, level floor surfaces and transitions to avoid tripping hazards*
Physiological Needs: *her comfort is critical, she keeps thermostat between* 72 and 75°F (22 and 24°C), she is keenly aware of drafts/air movement
Psychological Needs: *safety, connection with community and family, continuing to stay active and involved in her family and community*
Attitude: *uses only what she needs, prefers traditional communication (e.g., talking face to face, writing letters), conserves energy to avoid wasting money, feels agnostic toward technology*

Figure 6.10 Data-driven persona representing a senior occupant.
Source: Agee *et al.* (2021).

Sadie
Senior Persona

Sadie is a 78-year-old retiree and widow. She lives by alone, but keeps a full schedule of commitments (e.g., with her church group, visiting with her grandkids, reading, and watching TV). She enjoys learning and keeping an active mind with a daily crossword puzzle and reading her Bible. She spends most of her day at home in her apartment. She lives alone, so feeling safe is important to her sense of well-being. She is cold-natured, and a cozy housing unit is one reason she is more satisfied with her current unit compared to her previous unit. She likes the heat pump in her apartment but is sensitive to direct air blowing on her. She sets her thermostat between 72 and 75°F (22 and 24°C). She uses 88 kWh/m^2/yr of energy. She has an Energy Star-rated dishwasher but cleans her daily dishes by hand. Sadie feels the old ways of life are better. She doesn't like new technology and prefers the old ways of communicating. For example, she writes letters to her friends instead of email. She remembers when times were hard and you didn't waste anything. She is intentional about conserving energy and money (e.g., turning off the TV, lights, and coffee). She lives on a fixed income and cannot afford to be wasteful.	***Behavior:*** *turns off lights and plug loads when not in the room, cleans dishes by hand, takes short to medium length showers, uses space heater to adapt indoor environment*

In closing, personas are a powerful tool to map observed or imagined occupant characteristics onto one or more representations of occupants. While they have not been extensively used in building design, we recommend their future research and implementation because of their desirable traits (e.g., relatability and tangibility to all stakeholders, complexity, and richness in characteristics). Ultimately, for personas to be incorporated into BPS tools, their characteristics must be mapped to simulation inputs and models. There is strong potential for personas to be developed in conjunction with agent-based models and the advanced occupant models described in Section 6.3. For example, a persona could be developed based on a large number of single-behavior models; the model parameter could be varied depending on the persona characteristics (e.g., very reactive to low illuminance levels).

6.4 Implementation of Occupant Models in Simulation Tools

Thus far, in this chapter, we have described a variety of occupant modeling approaches. In this next section, we provide an overview and analysis of common methods of implementing occupant models in simulation tools. Since the remainder of this book focuses on building simulation, this preliminary overview of implementation is an essential step before occupant modeling can be discussed in terms of supporting the design process (Chapters 7 and 8). This section describes current available methods to implement occupant models in BPS tools as well as offering discussion on limitations and future research and development needs.

6.4.1 Occupant-Centric Simulation Tools and Approaches

While modeling occupants using schedules in BPS tools is commonplace (see Section 6.2), more advanced models require more sophisticated means for implementation. In general, BPS tools with a graphical front-end interface are more restrictive, while research-grade tools with open-source capabilities have greater flexibility to implement advanced occupant models. Most BPS tools provide at least one of the following approaches to model occupants (Hong *et al.*, 2018). We hereby divide implementation methods into two categories: those which are integrated into BPS tools and those which generate inputs in advance of integrating them into BPS (i.e., offline and stand-alone). BPS-integrated methods include:

- **Schedules** (deterministic)—These built-in or user-customized schedules generally represent occupant-related states as repeating time-varying parameters (e.g., occupancy profiles, lighting/equipment loads, temperature setpoint). The level of schedule resolution varies among BPS tools, with some allowing sub-hourly resolution and others being restricted to hourly. Ideally, tools represent these schedules graphically to identify errors quickly.
- **Rules** (deterministic/stochastic)—This method enables simulation users to use built-in rules or specify a set of rules for different building systems such as lighting, windows, and shading devices. For example, rules can be set to turn off lights when daylight illuminance reaches a certain threshold to simulation typical occupant behavior. The threshold can be probabilistic to simulate the variability of occupants' manual interactions with lights. While some BPS tools do not use such rules at all, others allow custom rules to be defined, such as User Function in DOE-2 and EMS in EnergyPlus (Gunay *et al.*, 2016). Moreover, models from external standardized languages can be integrated into the simulation tool, such as TRNSYS or IDA-ICE. Although user-customized controls allow more flexibility for the users to incorporate bespoke models, their implementation and debugging require a strong knowledge of the occupant models and programming.
- **User-defined source code** (deterministic/stochastic)—Some occupant models involve more than simple rules and may necessitate that the

source code modifications. However, advanced user knowledge is required for this approach.

- **Co-simulation** (deterministic/stochastic)—Occupant models can also be implemented in BPS via co-simulation that allows the dynamic exchange of information between BPS tools. For example, the occupant behavior Functional Mockup Unit (obFMU) is an example of a co-simulation method that supports reading the data in a standardized XML format through a new schema, titled 'occupant behavior XML' (obXML). The initial repository of obXML contains 52 models (Belafi *et al.*, 2019). A more advanced and flexible interface is Building Control Virtual Test Bed (BCVTB) which is based on a stand-alone interface (*Ptolemy II*) to host certain programs (Wetter, 2011). Similar to user-customized source code, advanced knowledge is required. Moreover, co-simulation can significantly increase computation time.

6.4.1.1 Stand-Alone (Offline) Methods Include:

- **Occupancy simulator** (stochastic)—This method is a web-based platform to provide hourly or sub-hourly occupant presence and movements based on stochastic models for an individual occupant in the form of CSV files (Chen *et al.*, 2017), which can then be used as an input for BPS tools such as EnergyPlus. Although this approach considers the diversity and stochasticity of occupancy, it is limited to occupancy schedules without considering two-way interactions between occupants and the environment. Moreover, this approach typically neglects interdependencies between different aspects of occupant behavior.
- **Offline techniques** (deterministic/stochastic)—An alternative method is to conduct sequential simulations to integrate occupant interactions in a building. Programming languages such as Python or R have the capability of high-level programming functions using a wide range of libraries and packages. The two main approaches are: (1) a pre-processing stage where occupant-centric control metrics are derived as inputs for further evaluation using BPS tools (Hobson *et al.*, 2021); or (2) a post-processing stage when a set of design alternatives are initially simulated as datasets to program occupant-centric control functions (Ouf *et al.*, 2019). Both techniques can deliver deterministic (e.g., rule-based) or probabilistic (e.g., supervised learning) controls to model occupant behavior. Further discussion on ways to simulate occupants to inform design is provided in Chapter 8.

While stand-alone offline methods potentially offer greater transparency and versatility, they do not offer dynamic interaction with the simulation. As such, interactions between triggers (e.g., IEQ) and occupant actions are not captured; thus, offline methods are more suitable for non-adaptive occupant features such as occupancy and office equipment use.

6.4.2 Current Limitations and Recommendations

Each of the major stages of building performance simulation—inputs, simulation, and outputs—have limitations with respect to occupant modeling that should be addressed in the future.

- **Inputs—**Generally speaking, increasing the number and complexity of occupant-related inputs will increase the level of occupant modeling detail in building simulation. These inputs can include occupant demographics and diversity, details on energy-related occupant behavior, and relationships between occupants (and the impact of these relationships on behavior). Current methods to specify occupant behavior are often abstract and implicit (e.g., refer to traditional occupant modeling approaches as in Section 6.2). Moreover, most tools treat occupants in much the way building systems are specified rather than as active participants in building performance. For example, in EnergyPlus, occupants' actions regarding blinds control are categorized as *window properties* rather than *people* objects and there are missing quantitative metrics to control certain functions such as shading systems through vertical eye illuminance (Tabadkani *et al.*, 2020). We recommend that occupant-related inputs are reframed and increased in detail to parallel recent research developments (e.g., more advanced models). Additionally, given the significant uncertainty during the design stage about the occupants that will occupy a space, features to allow ranges of occupant traits is a beneficial feature (Ouf *et al.*, 2019).
- **Simulation—**Most common BPS tools have very limited capabilities regarding occupant modeling (i.e., similar to those described in Section 6.2, rather than Section 6.3). Thus, for the reasons argued in Section 6.3, we strongly recommend an increase in the number and capability of occupant models in research-grade and mainstream BPS tools. Common BPS tools can process only a single simulation at a time without defining a correlation between occupants' behavioral aspects (e.g., occupancy profile and light switching) (Ouf *et al.*, 2018). However, more complex occupant models often necessitate multiple simulation runs, e.g., to quantify uncertainty and stochastic model distributions. Thus, we recommend new BPS tool features to automate batch simulations. While co-simulation has shown significant flexibility for implementing and simulating occupant models, it is not compatible with many BPS tools and requires advanced modeling knowledge, which hinders industry adoption. To overcome existing limitations of common BPS tools in terms of linking different occupant-related variables together (e.g., occupant density and lighting/equipment loads), parametric design tools such as open-source Ladybug Tools can be used to allow the definition of correlations among inputs algorithmically. Parametric-based interfaces enable simulating a large number of iterations automatically to efficiently quantify the impact of different occupants or occupant models.

- **Outputs—**Because BPS is rooted in annual energy use predictions, BPS tool outputs tend to focus on building performance rather than occupants (e.g., discomfort hours of the building rather than discomfort hours of occupants). BPS tools should be more informative and use an occupant-centric approach such that results are output and presented them from an occupant experience perspective. Moreover, many occupant-related simulation outputs are not available for reporting (e.g., number of light switching actions, view to outdoors). Future BPS tool should have features that support the output and visualization of occupant uncertainty (and other sources of uncertainty), such as probability distributions resulting from stochastic occupant models. Further discussion on simulation outputs and communicating results is presented in Section 6.5.

This section briefly summarized existing methods through which occupant models can be incorporated and implemented into BPS tools. It also provided recommendations on BPS tool inputs, simulation, and outputs to support occupant modeling. The next section explores the improvement of transparency of occupant modeling for practitioners and other users.

6.5 Communication and Practical Application

As the complexity of modeling approaches and their underlying statistical methods has increased, so has the number of variables taken into account when creating occupant behavioral models. Several researchers have attempted to classify the growing number of data sources and modeling approaches. For example, Mahdavi and Taheri (2017) presented an ontology for the classification of building performance data (e.g., air temperature, energy use), and others have discussed ways to select the most appropriate model for a specific simulation task (Gaetani *et al.*, 2016; Mahdavi and Tahmasebi, 2017; Tahmasebi and Mahdavi, 2016) (see also Chapter 7). As discussed in Section 6.2, most current approaches focus on schedules, which are relatable and simple to interpret for practitioners. In contrast, more advanced modeling approaches presented in scientific literature (and Section 6.3) are not suitable to communicate model results such as schedules or, for stochastic models, the variance in behavioral patterns.

Accordingly, there is a need to communicate occupant model properties and results in a comprehensible way, especially for those who apply these models, such as building engineers, without expertise in statistics. This argument is emphasized by O'Brien *et al.* (2016) who conclude, based on data from a survey among practitioners, that time and understanding are major obstacles of using more advanced occupant modeling. Thus, increasing the comprehensibility of occupant models may be a prerequisite for their widespread application in building performance simulation for design and operation of buildings. To date, there are only a few attempts to communicate to simulation users the impact of occupant modeling choices (e.g., Chen *et al.*, 2017; Gunay *et al.*, 2016; Ouf *et al.*, 2019; Schweiker *et al.*, 2019).

Discussing all potential methods to communicate occupant models is beyond the scope of this chapter and still debated among researchers. The most important aspects of a model's behavior to be communicated depend on the characteristics of interest and whether the practitioner is, for example, an engineer applying a model in communication with the researcher who developed the model, or an architect or investor communicating with the simulation engineer. Basic characteristics need to be communicated to enable (1) the comparison between different models and (2) judgment of the suitability of a model, including the number and type of input and output variables, potential hidden values, the basis of model (e.g., type of data collection, type of building and occupants monitored, region, climate), and the validation status of the model together if available with validation results.

To ease the understanding of a models' behavior, a breakdown of potentially complex model behaviors into transferable and communicable parameters is desirable. Such parameters could be descriptive values, such as the predicted mean duration of the behavior, the number of actions, the sensitivity of model to variance in input parameters, or the effect of the predicted behavioral patterns on other outcome parameters (Gunay *et al.*, 2016; Schweiker *et al.*, 2019). For example, Gunay *et al.* (2016) presented a method to compare a variety of occupant behavior models in terms of behavioral characteristics as well as energy use variations. Other ways to present the behavior of complex models is the generation of exemplary schedules resulting from their application. Such an approach is presented by Ouf *et al.* (2019) for stochastic models of lighting usage and by (Schweiker *et al.*, 2019) for window opening models' behaviors. The latter presented a method to compare model behaviors parametrically for combinations of different climates and building properties (see also Figure 6.11).

6.6 The Future of Occupant Modeling and Simulation

Major challenges and opportunities exist regarding occupant modeling, in the context of the methods proposed in Section 6.3. Care must be taken to balance accuracy gained by the relatively advanced statistical modeling of that section (relative to the knowledge of most BPS practitioners) with the opaqueness and obscurity that results (e.g., see Section 6.5). Ultimately, a simple model that is fit-for-purpose (see Chapter 7) is better than an inappropriately advanced model. With the Internet of Things (IoT), Internet-connected building automation systems, and other smart building technologies, the availability of occupant-related data is improving and becoming less costly to collect. We expect this to greatly enhance the ability to develop robust, data-driven occupant profiles for a variety of domains, building types, climates, etc. However, centrally managed and coordinated efforts, such as ASHRAE's Global Occupant Behavior Database (Dong *et al.*, 2021) are still

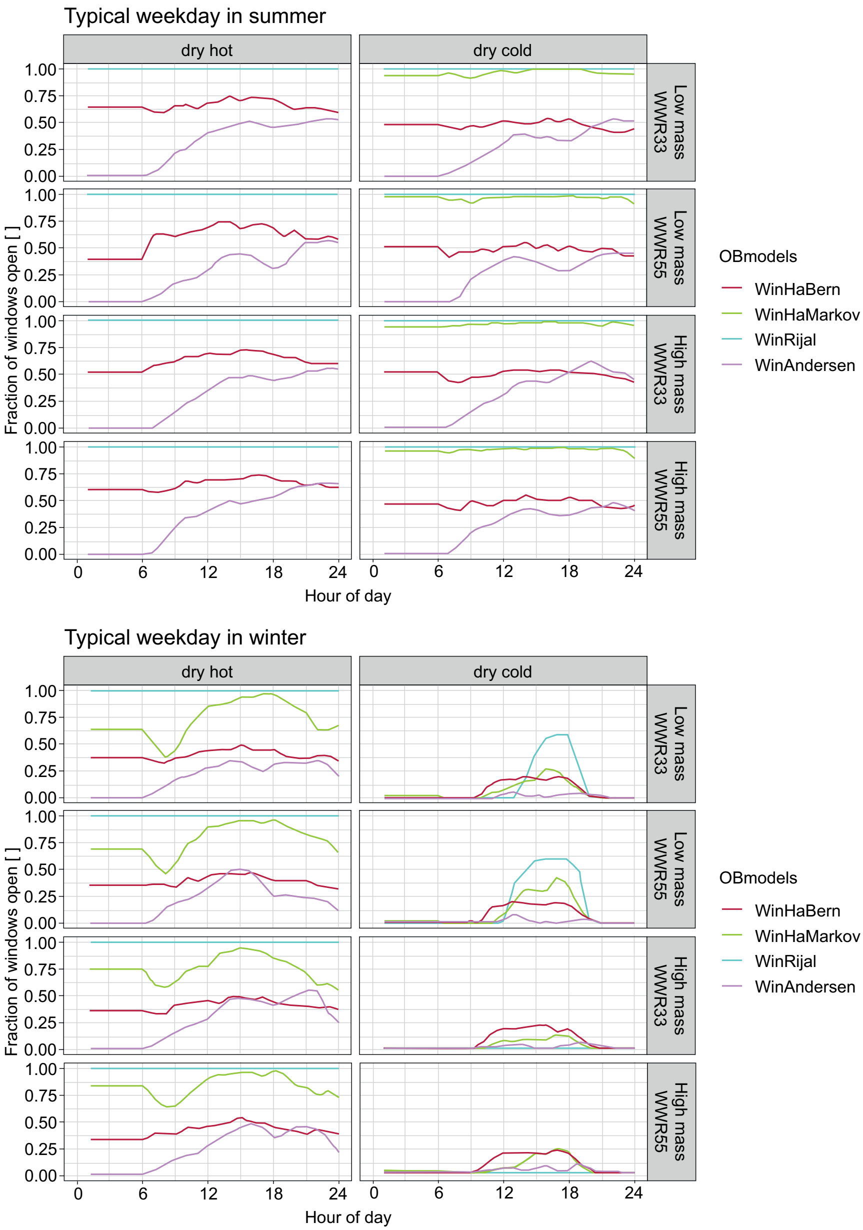

Figure 6.11 Example visualization of the behavior of four different window opening behavior models for simulated summer and winter day depending on weather data file and building characteristics.

required to maintain model quality, reliability, and consistency. Commensurate with any centrally-managed occupant data and model repository should also be rigorous verification of the generalizability of occupant models, much like the early work of (Schweiker *et al*., 2012).

We must also recognize that while the buildings industry will witness the emergence and widespread use of new occupant modeling techniques in the future, as summarized above and throughout this chapter, it is likely that the industry will simultaneously experience a widening or shifting of the disciplines that undertake the activity in practice.

Despite the many advances that have been made with respect to methods for generating occupancy schedules, the inherent simplification of occupancy presence in existing BPS tools remains a research gap facing the future of occupancy modeling as a whole. Rooms within buildings, like an open-office space, are subject to thermal and occupant asymmetries. An occupant sitting near a window will face unique thermal conditions, and may respond uniquely to environmental control decisions compared to an occupant located in a different position of the same modeled room (Brager *et al*., 2004). Different HVAC concepts can also produce unique asymmetrical thermal environments, where the specific position and location of an occupant in an HVAC-conditioned space, including the extent to which one's own limbs are exposed in that space, produces a unique regime of thermal sensation across the occupant's body (De Dear, 2011). These asymmetries are known to increase uncertainty in predicting occupant control decisions and predicting building energy demand using existing BPS tools (Halawa *et al*., 2014).

Accurately simulating spatial asymmetries between occupants and the built environment involves overcoming at least two challenges: (1) predicting the specific location and orientation of an occupant with respect to 3D space and time; and (2) directly modeling the asymmetrical relationship between the occupant and the indoor environment. Established advances in coupling BPS with computational fluid dynamics (CFD) have long-since illustrated how the latter challenge can be overcome (Zhai *et al*., 2002). The first challenge persists, however, albeit with a number of emerging solutions in the research pipeline. Most interesting is that these solutions are originating from fields that have historically lagged behind BPS, namely, architecture and computer-aided design.

Whereas decades ago, only a few architects were using computers for design, let alone simulation, the division of computer and simulation literacy between engineers and architects has narrowed considerably. Simulation and software programming has not only been introduced to architects, it is also fast becoming a standard skillset in the field (Riekstins, 2018). Credit for this goes particularly to Grasshopper 3D, a visual programming language that was created in 2007 by Rutten and McNeel (2007) to enable parametric, programmable computer-aided design. Grasshopper is effectively a functional mock-up environment which connects a programmable computer-aided design process with a growing suite of third-party simulation tools and other plug-ins written natively for the Grasshopper environment. Like the coupling of BPS with CFD modeling, Grasshopper provides the opportunity to couple BPS with highly spatial, parametric design algorithms that can include the modeling of occupant movement, behavior, and thermal sensation in fully-resolved 3D spaces.

Several recent examples of Grasshopper-based occupancy modeling are relevant to acknowledge. Aviv *et al.* (2022) used Grasshopper to develop a raytracing-based radiant heat transfer model to resolve the radiant asymmetries between occupants and the built environment. PedSim Pro, a pedestrian movement simulation tool developed for Grasshopper, was used by Pan *et al.* (2021) to generate time- and space-varying building occupancy profiles. Yi (2020) achieved a similar outcome by using Grasshopper to couple a BPS model with a hybrid agent-based model of occupant movement and behavior. As more Grasshopper-based BPS tools emerge and become popular, such as ClimateStudio (Solemma Inc., 2022), we can expect the field of highly-spatial occupancy modeling to grow more capable, and commonplace, in the years to come. We can also expect to see more and more architects leading this charge in future practice.

6.7 Closing Remarks

In this chapter, we provided an overview of occupant modeling from traditional and current practices to advanced occupant modeling. We explained why we should model occupants and that representing occupants using fixed schedules has some major limitations in simulation-aided building design.

We also covered the major traits of occupant models (stochastic, dynamic, data-driven, and agent-based) and their implications for simulation and building design. It explained how the different model types can be developed from various sources of occupant data. Next, we provided an overview of methods to implement occupant models into building performance simulation tools, ranging from schedules to co-simulation. We concluded the chapter with a discussion on how occupant models and their characteristics can be better communicated to users as well as in future work.

While in this chapter we discussed model selection in the context of accuracy and strengths and weaknesses, the next two chapters delve into details on selecting the most appropriate occupant models for a given purpose and then methods to use occupant models to support building design.

Note

1 Both logistic and probit regression are generalized linear models (GLM). In such models, instead of using the outcome Y, a link function is used, which is a function of the mean of Y. The difference between logit and probit models is in the link function: logit models make use of an inverse normal function, and the probit model makes use of a logit link function.

References

Abushakra B, Haberl JS, and Claridge DE (2004), Overview of existing literature on diversity factors and schedules for energy and cooling load calculations, *ASHRAE Transactions*, 110(1) 164–176.

Agee P, Gao X, Paige F, McCoy A, and Kleiner B (2021), A human-centred approach to smart housing, *Building Research & Information*, 49(1):84–99.

Andersen R, Fabi V, Toftum J, Corgnati SP, and Olesen BW (2013), Window opening behaviour modelled from measurements in Danish dwellings, *Building and Environment*, 69:101–113.

Aviv D, Gros J, Alsaad H, Teitelbaum E, Voelker C, Pantelic J, and Meggers F (2022), A data-driven ray tracing simulation for mean radiant temperature and spatial variations in the indoor radiant field with experimental validation, *Energy and Buildings*, 254:111585.

Azar E, O'Brien W, Carlucci S, Hong T, Sonta A, Kim J, Andargie MS, Abuimara T, El Asmar M, and Jain R (2020), Simulation-aided occupant-centric building design: a critical review of tools, methods, and applications, *Energy and Buildings*:110292:1–21.

Belafi ZD, Hong T, and Reith A (2019), A library of building occupant behaviour models represented in a standardised schema, *Energy Efficiency*, 12(3):637–651.

Bennetts H, Martins LA, van Hoof J, and Soebarto V (2020), Thermal personalities of older people in South Australia: a personas-based approach to develop thermal comfort guidelines, *International Journal of Environmental Research and Public Health*, 17(22):8402.

Berger C, and Mahdavi A (2020), Review of current trends in agent-based modeling of building occupants for energy and indoor-environmental performance analysis, *Building and Environment*, 173:106726.

Bernhardt K (2007), Agent-based modeling in transportation, *Artificial Intelligence in Transportation*, Vol. 113 72–80.

Blythe MA and Wright PC (2006), Pastiche scenarios: fiction as a resource for user centred design, *Interacting with Computers*, 18(5):1139–1164.

Bonabeau E (2002), Agent-based modeling: methods and techniques for simulating human systems, *Proceedings of the National Academy of Sciences*, 99(suppl 3):7280–7287.

Brager G, Paliaga G, and De Dear R (2004), Operable windows, personal control and occupant comfort, *ASHRAE Transactions*, 110(2).

Calì D (2016), *Occupants' Behavior and Its Impact upon the Energy Performance of Buildings*, Universitätsbibliothek der RWTH Aachen.

Calì D, Andersen RK, Müller D, and Olesen BW (2016), Analysis of occupants' behavior related to the use of windows in German households, *Building and Environment*, 103:54–69.

Calì D, Wesseling MT, and Müller D (2018), WinProGen: a Markov-chain-based stochastic window status profile generator for the simulation of realistic energy performance in buildings, *Building and Environment*, 136:240–258.

Carlucci S, De Simone M, Firth SK, Kjærgaard MB, Markovic R, Rahaman MS, Annaqeeb MK, Biandrate S, Das A, Dziedzic JW, Fajilla G, Favero M, Ferrando M, Hahn J, Han M, Peng Y, Salim F, Schlüter A, and van Treeck C (2020), Modeling occupant behavior in buildings, *Building and Environment*, 174:106768, URL https://doi.org/10.1016/j.buildenv.2020.106768.

Carpino C, Mora D, Arcuri N, and De Simone M (2017), Behavioral variables and occupancy patterns in the design and modeling of Nearly Zero Energy Buildings. Paper read at Building Simulation.

Chen Y, Liang X, Hong T, and Luo X (2017), Simulation and visualization of energy-related occupant behavior in office buildings. Paper read at Building Simulation.

Clarke J (2001), *Energy Simulation in Building Design*, Butterworth-Heinemann, Oxford, 2nd edition.

Cooper A (1999), *The Inmates Are Running the Asylum: Why High-Tech Products Drive Us*. Pearson Higher Education, Indianapolis, IN

Cooper A, Reimann R, Cronin D, and Noessel C (2014), *About Face: The Essentials of Interaction Design*, John Wiley & Sons, Indianapolis, IN.

Crawley DB, Lawrie LK, Winkelmann FC, Buhl WF, Huang YJ, Pedersen CO, Strand RK, Liesen RJ, Fisher DE, Witte MJ, and Glazer J (2001), EnergyPlus: creating a new-generation building energy simulation program, *Energy and Buildings*, 33(4):319–331, URL https://doi.org/10.1016/S0378-7788(00)00114-6.

D'Oca S, Gunay HB, Gilani S, and O'Brien W (2019), Critical review and illustrative examples of office occupant modelling formalisms, *Building Services Engineering Research and Technology*, 40(6), URL https://doi.org/10.1177/0143624419827468.

De Dear R (2011), Revisiting an old hypothesis of human thermal perception: alliesthesia, *Building Research & Information*, 39(2):108–117.

Deng Z and Chen Q (2021), Reinforcement learning of occupant behavior model for cross-building transfer learning to various HVAC control systems, *Energy and Buildings*, 238:110860.

Deru M, Field K, Studer D, Benne K, Griffith B, Torcellini P, Liu B, Halverson M, Winiarski D, and Rosenberg M (2011), *Commercial Reference Building Models of the National Building Stock*, US Department of Energy, Golden, Colorado.

Diamond SC and Hunn BD (1981), *Comparison of DOE-2 Computer Program Simulations to Metered Data for Seven Commercial Buildings*, Los Alamos National Lab (LANL), Los Alamos, NM.

Dong B, Hong T, and O'Neill Z (2021), *ASHRAE Global Occupant Behavior Database*, ASHRAE, Atlanta, Georgia.

Dong B, Li Z, and McFadden G (2015), An investigation on energy-related occupancy behavior for low-income residential buildings, *Science and Technology for the Built Environment*, 21(6):892–901, URL https://doi.org/10.1080/23744731.2015.1040321.

Duarte C, Van Den Wymelenberg K, and Rieger C (2013), Revealing occupancy patterns in an office building through the use of occupancy sensor data, *Energy and Buildings*, 67:587–595, URL https://doi.org/10.1016/j.enbuild.2013.08.062.

Fabi V, Andersen RV, Corgnati S, and Olesen BW (2012), Occupants' window opening behaviour: a literature review of factors influencing occupant behaviour and models, *Building and Environment*, 58(December):188–198.

Feller W (1968), *An Introduction to Probability Theory and Its Applications*, Wiley.

Gaetani I, Hoes P-J, and Hensen JLM (2016), Occupant behavior in building energy simulation: towards a fit-for-purpose modeling strategy, *Energy and Buildings*, 121:188–204, URL https://doi.org/10.1016/j.enbuild.2016.03.038.

Gilani S, O'Brien W, Gunay HB, and Carrizo JS (2016), Use of dynamic occupant behavior models in the building design and code compliance processes, *Energy and Buildings*, 117(1):260–271, URL https://doi.org/10.1016/j.enbuild.2015.10.044.

Gilbert N (2019), *Agent-Based Models*, Sage Publications, London, UK.

Gilks WR, Richardson S, and Spiegelhalter D (1995), *Markov Chain Monte Carlo in Practice*, CRC Press Boca Raton, FL.

Goldstein R, Tessier A, and Khan A (2010), Customizing the behavior of interacting occupants using personas, in *Conference: IBPSA-USA SimBuild* New York, NY, August 11–13.

Gopika S (2015), Agent based HVAC optimization model for building energy efficiency. Paper read at 2015 *IEEE 9th International Conference on Intelligent Systems and Control (ISCO)* Coimbatore, India, January 9–10.

Goubran S, Cucuzzella C, and Ouf MM (2021), Eyes on the goal! Exploring interactive artistic real-time energy interfaces for target-specific actions in the built environment, *Sustainability*, 13 (4):1996.

Guerra Santin O, Itard L, and Visscher H (2009), The effect of occupancy and building characteristics on energy use for space and water heating in Dutch residential stock, *Energy and Buildings*, 41(11):1223–1232.

Gunay HB, O'Brien W, and Beausoleil-Morrison I (2016), Implementation and comparison of existing occupant behaviour models in EnergyPlus, *Journal of Building Performance Simulation*, 9(6):567–588.

Haines V and Mitchell V (2014), A persona-based approach to domestic energy retrofit, *Building Research & Information*, 42(4):462–476.

Halawa E, van Hoof J, and Soebarto V (2014), The impacts of the thermal radiation field on thermal comfort, energy consumption and control—a critical overview, *Renewable and Sustainable Energy Reviews*, 37:907–918.

Haldi F (2013), A probabilistic model to predict building occupants'diversity towards their interactions with the building envelope, in *Building Simulation*, Chambery, France, August 26–29.

Haldi F, Calì D, Andersen RK, Wesseling M, and Müller D (2016), Modelling diversity in building occupant behaviour: a novel statistical approach, *Journal of Building Performance Simulation*:1–18, URL https://doi.org/10.1080/19401493.2016.1269245.

Haldi F and Robinson D (2009), Interactions with window openings by office occupants, *Building and Environment*, 44(12):2378–2395, URL https://doi.org/10.1016/j.buildenv.2009.03.025.

Hastie TJ and Tibshirani RJ (2017), *Generalized Additive Models*. Routledge New York, NY.

Hobson BW, Huchuk B, Gunay HB, and O'Brien W (2021), A workflow for evaluating occupant-centric controls using building simulation, *Journal of Building Performance Simulation*:1–19.

Hoes P, Hensen JLM, Loomans MGLC, de Vries B, and Bourgeois D (2009), User behavior in whole building simulation, *Energy and Buildings*, 41(3):295–302.

Hong T, Chen Y, Belafi Z, and D'Oca S (2018), Occupant behavior models: a critical review of implementation and representation approaches in building performance simulation programs. Paper read at Building Simulation.

Hong T, Yan D, D'Oca S, and Chen C-F (2017), Ten questions concerning occupant behavior in buildings: the big picture, *Building and Environment*, 114:518–530.

IESNA. 2012. LM-83-12 IES Spatial Daylight Autonomy (sDA) and Annual Sunlight Exposure (ASE). In *IESNA Lighting Measurement*. New York.

Iwashita G, and Akasaka H (1997), The effects of human behavior on natural ventilation rate and indoor air environment in summer—a field study in southern Japan, *Energy and Buildings*, 25(3):195–205.

Jakubiec JA and Reinhart CF (2012), The 'adaptive zone'–a concept for assessing discomfort glare throughout daylit spaces, *Lighting Research and Technology*, 44(2):149–170.

Langevin J, Wen J, and Gurian PL (2015), Simulating the human-building interaction: development and validation of an agent-based model of office occupant behaviors, *Building and Environment*, 88:27–45, URL https://doi.org/10.1016/j.buildenv.2014.11.037.

Lee YS and Malkawi AM (2014), Simulating multiple occupant behaviors in buildings: an agent-based modeling approach, *Energy and Buildings*, 69:407–416.

Li S, Yang Z, Hu D, Cao L, and He Q (2021), Understanding building-occupant-microbiome interactions toward healthy built environments: a review, *Frontiers of Environmental Science & Engineering*, 15(4):1–18.

Linkola L, Andrews CJ, and Schuetze T (2013), An agent based model of household water use, *Water*, 5(3):1082–1100.

Liu Y, Zhang Z, and Mao Z (2020), Analysis of influencing factors in pre-evacuation time using interpretive structural modeling, *Safety Science*, 128:104785.

Lizhong Y, Weifeng F, and Weicheng F (2003), Modeling occupant evacuation using cellular automata—effect of human behavior and building characteristics on evacuation, *Journal of Fire Sciences*, 21(3):227–240.

Luna-Navarro A, Loonen R, Juaristi M, Monge-Barrio A, Attia S, and Overend M (2020), Occupant-facade interaction: a review and classification scheme, *Building and Environment*, 177:106880, URL https://doi.org/10.1016/j.buildenv.2020.106880.

Mahdavi A, Tahmasebi F, Gunay B, O'Brien W, and D'Oca S (2017), *Occupant Behavior Modeling Approaches and Evaluation*, EBC Birmingham, UK.

Mahdavi A, Berger C, Amin H, Ampatzi E, Andersen RK, Azar E, Barthelmes VM, Favero M, Hahn J, and Khovalyg D (2021), The role of occupants in buildings' energy performance gap: myth or reality?, *Sustainability*, 13(6):3146.

Mahdavi A and Taheri M (2017), An ontology for building monitoring, *Journal of Building Performance Simulation*, 10(5–6):499–508, URL https://doi.org/10.1080/19401493.2016.1243730.

Mahdavi A and Tahmasebi F (2017), On the quality evaluation of behavioural models for building performance applications, *Journal of Building Performance Simulation*, 10(5–6):554–564, URL https://doi.org/10.1080/19401493.2016.1230148.

McCracken I, Harpeand RD, and Ruvo MD (2019), Developing dementia personas for user centered architectural design considerations in non-specialized contexts. Paper read at Dementia Lab Conference.

McCulloch CE, Searle SR, and Neuhaus JM (2003), *Generalized, Linear, and Mixed Models Generalized, Linear, and Mixed Models*, John Wiley & Sons, Hoboken, NJ, URL https://doi.org/10.1198/tech.

McKenna E, Krawczynski M, and Thomson M (2015), Four-state domestic building occupancy model for energy demand simulations, *Energy and Buildings*, 96:30–39.

Nielsen L (2013), *Personas-User Focused Design*, Springer, London, UK.

Norford LK (1984), *An Analysis of Energy Use in Office Buildings: The Case of ENERPLEX*, Princeton University.

O'Brien W and Gunay HB (2014), The contextual factors contributing to occupants' adaptive comfort behaviors in offices – a review and proposed modeling framework, *Building and Environment*, 77:77–87.

O'Brien W, Tahmasebi F, Andersen RK, Azar E, Barthelmes V, Belafi ZD, Berger C, Chen D, De Simone M, and d'Oca S (2020), An international review of occupant-related aspects of building energy codes and standards, *Building and Environment*:106906.

O'Brien W, Abdelalim A, and Gunay HB (2019), Development of an office tenant electricity use model and its application for right-sizing HVAC equipment, *Journal of Building Performance Simulation*, 12(1):37–55, URL https://doi.org/10.1080/19401493.2018.1463394.

O'Brien W, Gaetani I, Carlucci S, Hoes PJ, and Hensen JL (2016), International survey on current occupant modelling approaches in building performance simulation, *Journal of Building Performance Simulation*:1–19, URL https://doi.org/10.1080/19401493.2016.1243731.

O'Brien W and Gunay HB (2015), Mitigating office performance uncertainty of occupant use of window blinds and lighting using robust design, *Building Simulation*, 8(6):621–636.

O'Brien W, Gunay HB, Tahmasebi F, and Mahdavi A (2017), A preliminary study of representing the inter-occupant diversity in occupant modelling, *Journal of Building Performance Simulation*, 10(5–6):509–526, URL https://doi.org/10.1080/19401493.2016.1261943.

O'Brien W, Wagner A, Schweiker M, Mahdavi A, Day J, Kjærgaard MB, Carlucci S, Dong B, Tahmasebi F, Yan D, and Hong T (2020), Introducing IEA EBC Annex 79: key challenges and opportunities in the field of occupant-centric building design and operation, *Building and Environment*:106738.

Ouf M, O'Brien W, and Gunay B (2018), Improving occupant-related features in building performance simulation tools, *Building Simulation*, URL https://doi.org/10.1007/s12273-018-0443-y.

Ouf M, O'Brien W, and Gunay B (2019), A method to generate design-sensitive occupant-related schedules for building performance simulations, *Science and Technology for the Built Environment*, 25(2):221–232.

Pablo-Martí F, Santos JL, and Kaszowska J (2015), An agent-based model of population dynamics for the European regions, *Emergence: Complexity & Organization*, 17(2).

Page J, Robinson D, Morel N, and Scartezzini J-L (2008), A generalised stochastic model for the simulation of occupant presence, *Energy and Buildings*, 40(2):83–98.

Pan J, Cho TY, and Bardhan R (2021), Redesigning the working space for social distancing: modelling the movement in an open-plan office, in *CIBSE Technical Symposium. Conference: Virtual Event*, 13–14 July.

Papadopoulos S and Azar E (2016), Integrating building performance simulation in agent-based modeling using regression surrogate models: a novel human-in-the-loop energy modeling approach, *Energy and Buildings*, 128:214–223.

Pinheiro J and Bates D (2006), *Mixed-Effects Models in S and S-PLUS*, Springer Science & Business Media, New York.

Porter E, Hamdar SH, and Daamen W (2018), Pedestrian dynamics at transit stations: an integrated pedestrian flow modeling approach, *Transportmetrica A: Transport Science*, 14(5–6):468–483.

Pruitt J and Adlin T (2010), *The Persona Lifecycle: Keeping People in Mind throughout Product Design*, Elsevier, San Francisco, CA.

Rabiner LR (1989), A tutorial on hidden Markov models and selected applications in speech recognition, *Proceedings of the IEEE*, 77(2):257–286.

Riekstins A (2018), Teaching parametricism as a standard skill for architecture, *Journal of Architecture and Urbanism*, 42(1):34–39.

Ringler P, Keles D, and Fichtner W (2016), Agent-based modelling and simulation of smart electricity grids and markets–a literature review, *Renewable and Sustainable Energy Reviews*, 57:205–215.

Rutten D and McNeel R (2007), *Grasshopper3D*, Robert McNeel & Associates, Seattle, WA.

Sailor DJ, Anand J, and Kalkstein L (2021), Potential overall heat exposure reduction associated with implementation of heat mitigation strategies in Los Angeles, *International Journal of Biometeorology*, 65(3):407–418.

Salgado M and Gilbert N (2013), Emergence and communication in computational sociology, *Journal for the Theory of Social Behaviour*, 43(1):87–110.

Sangi R, Bünning F, Fütterer J, and Müller D (2017), A platform for the agent-based control of HVAC systems. Paper read at *Proceedings of the 12th International Modelica Conference*, Prague, Czech Republic, May 15–17, 2017.

Sayama H (2015), *Introduction to the Modeling and Analysis of Complex Systems*, Open SUNY Textbooks, Geneseo, NY.

Schaumann D, Breslav S, Goldstein R, Khan A, and Kalay YE (2017), Simulating use scenarios in hospitals using multi-agent narratives, *Journal of Building Performance Simulation*, 10(5–6):636–652.

Schwarz G (1978), Estimating the dimension of a model, *The Annals of Statistics* 6(2): 461–464.

Schweiker M, Carlucci S, Andersen RK, Dong B, and O'Brien W (2018), *Exploring Occupant Behavior in Buildings*, chapter Occupancy and occupants' actions, edited by A. Wagner, W. O'Brien, and B.Dong, Springer, Cham, pp. 7–38.

Schweiker M, Haldi F, Shukuya M, and Robinson D (2012), Verification of stochastic models of window opening behaviour for residential buildings, *Journal of Building Performance Simulation*, 5(1):55–74.

Schweiker M, O'Brien W, and Gunay B (2019), Characterization of occupant behaviour models for simulation engineers and architects, in *Conference: Building Simulation*, Rome, Italy.

Schweiker M and Shukuya M (2009), Comparison of theoretical and statistical models of air-conditioning-unit usage behaviour in a residential setting under Japanese climatic conditions, *Building and Environment*, 44(10):2137–2149.

Solemma Inc. (2022), *Climate Studio*, Cambridge, MA, URL https://www.solemma.com/climatestudio.

Stopps H and Touchie MF (2021), Residential smart thermostat use: an exploration of thermostat programming, environmental attitudes, and the influence of smart controls on energy savings, *Energy and Buildings*, 238:110834.

Tabadkani A, Tsangrassoulis A, Roetzel A, and Li HX (2020), Innovative control approaches to assess energy implications of adaptive facades based on simulation using EnergyPlus, *Solar Energy*, 206:256–268.

Tahmasebi F and Mahdavi A (2016), An inquiry into the reliability of window operation models in building performance simulation, *Building and Environment*, 105:343–357, URL https://doi.org/10.1016/j.buildenv.2016.06.013.

Takai S and Ishii K (2010), A use of subjective clustering to support affinity diagram results in customer needs analysis, *Concurrent Engineering*, 18(2):101–109.

Tang T-Q, Wu Y-H, Huang H-J, and Caccetta L (2012), An aircraft boarding model accounting for passengers' individual properties, *Transportation Research Part C: Emerging Technologies*, 22:1–16.

Tesfatsion L (2002), Agent-based computational economics: growing economies from the bottom up, *Artificial Life*, 8(1):55–82.

Tracy M, Cerdá M, and Keyes KM (2018), Agent-based modeling in public health: current applications and future directions, *Annual Review of Public Health*, 39:77–94.

Vine EL, Craig PP, Cramer JC, Dietz TM, Hackett BM, Kowalczyk DJ, and Levine MD (1982), The applicability of energy models to occupied houses: summer electric use in Davis, *Energy*, 7(11):909–925.

Wetter M (2011), Co-simulation of building energy and control systems with the Building controls virtual test bed, *Journal of Building Performance Simulation*, 4(3):185–203.

Wolf S, Calì D, Krogstie J, and Madsen H (2019), Carbon dioxide-based occupancy estimation using stochastic differential equations, *Applied Energy*, 236:32–41.

Yan B, Hao F, and Meng X (2021), When artificial intelligence meets building energy efficiency: a review focusing on zero energy building, *Artificial Intelligence Review*, 54(3):2193–2220.

Yan D, O'Brien W, Hong T, Feng X, Gunay HB, Tahmasebi F, and Mahdavi A (2015), Occupant behavior modeling for building performance simulation: current state and future challenges, *Energy and Buildings*, 107:264–278, URL https://doi.org/10.1016/j.enbuild.2015.08.032.

Yang Y, Deng J, Xie CC, and Jiang YT (2014), Design and implementation of fire safety evacuation simulation software based on cellular automata model, *Procedia Engineering*, 71:364–371.

Yi H (2020), 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*, 12(16):6672.

York DA and Cappiello CC (1981), *DOE-2 Engineers Manual (Version 2.1A)*, Lawrence Berkeley Lab, Berkeley, CA; Los Alamos National Lab, Santa FE, NM.

Zhai Z, Chen Q, Haves P, and Klems JH (2002), On approaches to couple energy simulation and computational fluid dynamics programs, *Building and Environment*, 37(8–9):857–864.

Zucchini W and MacDonald IL (2009), *Hidden Markov Models for Time Series: An Introduction Using R*, Chapman and Hall/CRC, Boca Raton, FL.