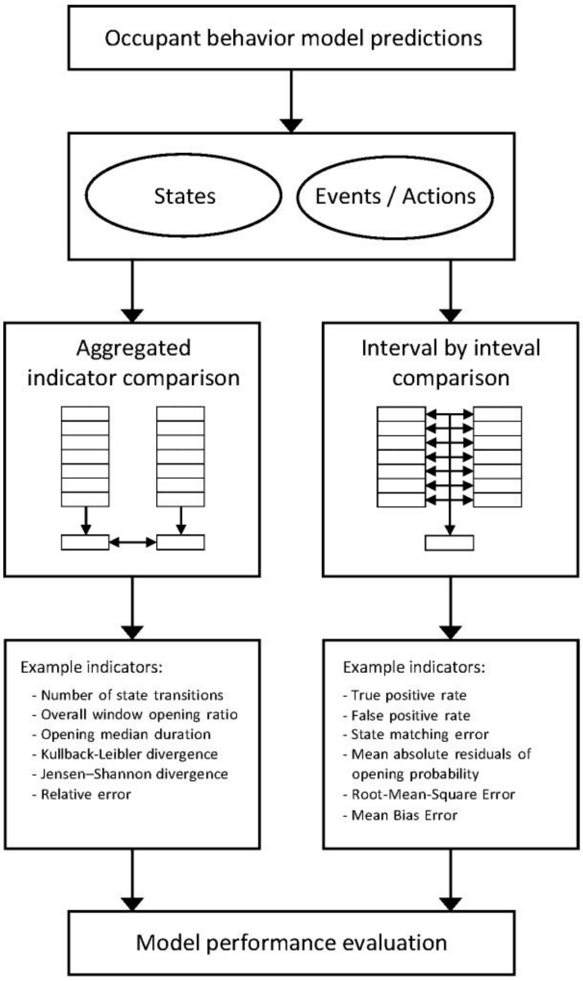
Annex 66: Definition and Simulation of Occupant Behavior in Buildings

Technical Report:

Occupant Behavior Modeling Approaches and Evaluation

November 2017



EBC is a programme of the International Energy Agency (IEA)

Published by Tsinghua University, Beijing 100084, China and LBNL, 1 Cyclotron Road, Berkeley, CA 94720, USA

ISBN 978-0-9996964-1-5

Participating countries in EBC: Australia, Austria, Belgium, Canada, P.R. China, Czech Republic, Denmark, France, Germany, Ireland, Italy, Japan, Republic of Korea, the Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom and the United States of America.

Additional copies of this report may be obtained from:

EBC Bookshop
C/o AECOM Ltd
The Colmore Building
Colmore Circus Queensway
Birmingham B4 6AT
United Kingdom
Web: www.iea-ebc.org
Email: essu@iea-ebc.org

Authors:	Ardeshir Mahdavi, TU Wien
	Farhang Tahmasebi, TU Wien
	Burak Gunay, Carleton University
	William O’Brien, Carleton University
	Simona D’Oca, LBNL
Reviewers:	Paul Ruyssevelt, UCL
	Michael Donn, Victoria University of Wellington
Edited by:	Tianzhen Hong, LBNL

Table of content

Preface

The International Energy Agency

The International Energy Agency (IEA) was established in 1974 within the framework of the Organisation for Economic Co-operation and Development (OECD) to implement an international energy programme. A basic aim of the IEA is to foster international co-operation among the 29 IEA participating countries and to increase energy security through energy research, development and demonstration in the fields of technologies for energy efficiency and renewable energy sources.

The IEA Energy in Buildings and Communities Programme

The IEA co-ordinates international energy research and development (R&D) activities through a comprehensive portfolio of Technology Collaboration Programmes. The mission of the Energy in Buildings and Communities (EBC) Programme is to develop and facilitate the integration of technologies and processes for energy efficiency and conservation into healthy, low emission, and sustainable buildings and communities, through innovation and research. (Until March 2013, the IEA-EBC Programme was known as the Energy in Buildings and Community Systems Programme, ECBCS.)

The research and development strategies of the IEA-EBC Programme are derived from research drivers, national programmes within IEA countries, and the IEA Future Buildings Forum Think Tank Workshops. The research and development (R&D) strategies of IEA-EBC aim to exploit technological opportunities to save energy in the buildings sector, and to remove technical obstacles to market penetration of new energy efficient technologies. The R&D strategies apply to residential, commercial, office buildings and community systems, and will impact the building industry in five focus areas for R&D activities:

- Integrated planning and building design
- Building energy systems
- Building envelope
- Community scale methods
- Real building energy use

The Executive Committee

Overall control of the IEA-EBC Programme is maintained by an Executive Committee, which not only monitors existing projects, but also identifies new strategic areas in which collaborative efforts may be beneficial. As the Programme is based on a contract with the IEA, the projects are legally established as Annexes to the IEA-EBC Implementing Agreement. At the present time, the following projects have been initiated by the IEA-EBC Executive Committee, with completed projects identified by (*):

Annex 1: Load Energy Determination of Buildings (*)
Annex 2: Ekistics and Advanced Community Energy Systems (*)
Annex 3: Energy Conservation in Residential Buildings (*)
Annex 4: Glasgow Commercial Building Monitoring (*)
Annex 5: Air Infiltration and Ventilation Centre
Annex 6: Energy Systems and Design of Communities (*)
Annex 7: Local Government Energy Planning (*)
Annex 8: Inhabitants Behaviour with Regard to Ventilation (*)
Annex 9: Minimum Ventilation Rates (*)
Annex 10: Building HVAC System Simulation (*)
Annex 11: Energy Auditing (*)
Annex 12: Windows and Fenestration (*)
Annex 13: Energy Management in Hospitals (*)
Annex 14: Condensation and Energy (*)
Annex 15: Energy Efficiency in Schools (*)
Annex 16: BEMS 1- User Interfaces and System Integration (*)
Annex 17: BEMS 2- Evaluation and Emulation Techniques (*)
Annex 18: Demand Controlled Ventilation Systems (*)
Annex 19: Low Slope Roof Systems (*)
Annex 20: Air Flow Patterns within Buildings (*)
Annex 21: Thermal Modelling (*)
Annex 22: Energy Efficient Communities (*)
Annex 23: Multi Zone Air Flow Modelling (COMIS) (*)
Annex 24: Heat, Air and Moisture Transfer in Envelopes (*)
Annex 25: Real time HVAC Simulation (*)

Annex 26: Energy Efficient Ventilation of Large Enclosures (*)
Annex 27: Evaluation and Demonstration of Domestic Ventilation Systems (*)
Annex 28: Low Energy Cooling Systems (*)
Annex 29: Daylight in Buildings (*)
Annex 30: Bringing Simulation to Application (*)
Annex 31: Energy-Related Environmental Impact of Buildings (*)
Annex 32: Integral Building Envelope Performance Assessment (*)
Annex 33: Advanced Local Energy Planning (*)
Annex 34: Computer-Aided Evaluation of HVAC System Performance (*)
Annex 35: Design of Energy Efficient Hybrid Ventilation (HYBVENT) (*)
Annex 36: Retrofitting of Educational Buildings (*)
Annex 37: Low Exergy Systems for Heating and Cooling of Buildings (LowEx) (*)
Annex 38: Solar Sustainable Housing (*)
Annex 39: High Performance Insulation Systems (*)
Annex 40: Building Commissioning to Improve Energy Performance (*)
Annex 41: Whole Building Heat, Air and Moisture Response (MOIST-ENG) (*)
Annex 42: The Simulation of Building-Integrated Fuel Cell and Other Cogeneration Systems (FC+COGEN-SIM) (*)
Annex 43: Testing and Validation of Building Energy Simulation Tools (*)
Annex 44: Integrating Environmentally Responsive Elements in Buildings (*)
Annex 45: Energy Efficient Electric Lighting for Buildings (*)
Annex 46: Holistic Assessment Tool-kit on Energy Efficient Retrofit Measures for Government Buildings (EnERGo) (*)
Annex 47: Cost-Effective Commissioning for Existing and Low Energy Buildings (*)
Annex 48: Heat Pumping and Reversible Air Conditioning (*)
Annex 49: Low Exergy Systems for High Performance Buildings and Communities (*)
Annex 50: Prefabricated Systems for Low Energy Renovation of Residential Buildings (*)
Annex 51: Energy Efficient Communities (*)
Annex 52: Towards Net Zero Energy Solar Buildings (*)
Annex 53: Total Energy Use in Buildings: Analysis & Evaluation Methods (*)
Annex 54: Integration of Micro-Generation & Related Energy Technologies in Buildings (*)
Annex 55: Reliability of Energy Efficient Building Retrofitting - Probability Assessment of Performance & Cost (RAP-RETRO) (*)
Annex 56: Cost Effective Energy & CO2 Emissions Optimization in Building Renovation (*)
Annex 57: Evaluation of Embodied Energy & CO2 Equivalent Emissions for Building Construction (*)
Annex 58: Reliable Building Energy Performance Characterisation Based on Full Scale Dynamic Measurements (*)
Annex 59: High Temperature Cooling & Low Temperature Heating in Buildings (*)
Annex 60: New Generation Computational Tools for Building & Community Energy Systems (*)
Annex 61: Business and Technical Concepts for Deep Energy Retrofit of Public Buildings (*)
Annex 62: Ventilative Cooling
Annex 63: Implementation of Energy Strategies in Communities
Annex 64: LowEx Communities - Optimised Performance of Energy Supply Systems with Exergy Principles
Annex 65: Long-Term Performance of Super-Insulating Materials in Building Components and Systems
Annex 66: Definition and Simulation of Occupant Behavior in Buildings
Annex 67: Energy Flexible Buildings
Annex 68: Indoor Air Quality Design and Control in Low Energy Residential Buildings
Annex 69: Strategy and Practice of Adaptive Thermal Comfort in Low Energy Buildings
Annex 70: Energy Epidemiology: Analysis of Real Building Energy Use at Scale
Annex 71: Building Energy Performance Assessment Based on In-situ Measurements

Working Group - Energy Efficiency in Educational Buildings (*)
Working Group - Indicators of Energy Efficiency in Cold Climate Buildings (*)
Working Group - Annex 36 Extension: The Energy Concept Adviser (*)
Working Group - Survey on HVAC Energy Calculation Methodologies for Non-residential Buildings

Introduction to Annex 66

Energy-related occupant behavior in buildings is a key issue for building design optimization, energy diagnosis, performance evaluation, and building energy simulation. Actions such as adjusting the thermostat for comfort, switching lights, opening/closing windows, pulling up/down window blinds, and moving between spaces, can have a significant impact on the real energy use and indoor environmental quality in buildings. Having a deeper understanding of occupant behavior, and quantifying their impact on the use of building technologies and building performance with modeling and simulation tools is crucial to the design and operation of low energy buildings where human-building interactions are the key. However, the influence of occupant behavior is under-recognized or over-simplified in the design, construction, operation, and retrofit of buildings.

Occupant behavior is complex and requires a multi-disciplinary approach if it is ever to be fully understood (Figure 1). On one hand, occupant behavior is influenced by external factors such as culture, economy and climate, as well as internal factors such as individual comfort preference, physiology, and psychology; On the other hand, occupant behavior drives occupants' interactions with building systems which strongly influence the building operations and thus energy use/cost and indoor comfort, which in-turn influences occupant behavior thus forming a closed loop.

There are over 20 groups all over the world studying occupant behavior individually. However, existing studies on occupant behavior, mainly from the perspective of sociology, lack in-depth quantitative analysis. Furthermore, the occupant behavior models developed by different researchers are often inconsistent, with a lack of consensus in common language, in good experimental design and in modeling methodologies. Therefore, there is a strong need for researchers to work together on a consistent and standard framework of occupant behavior definition and simulation methodology.

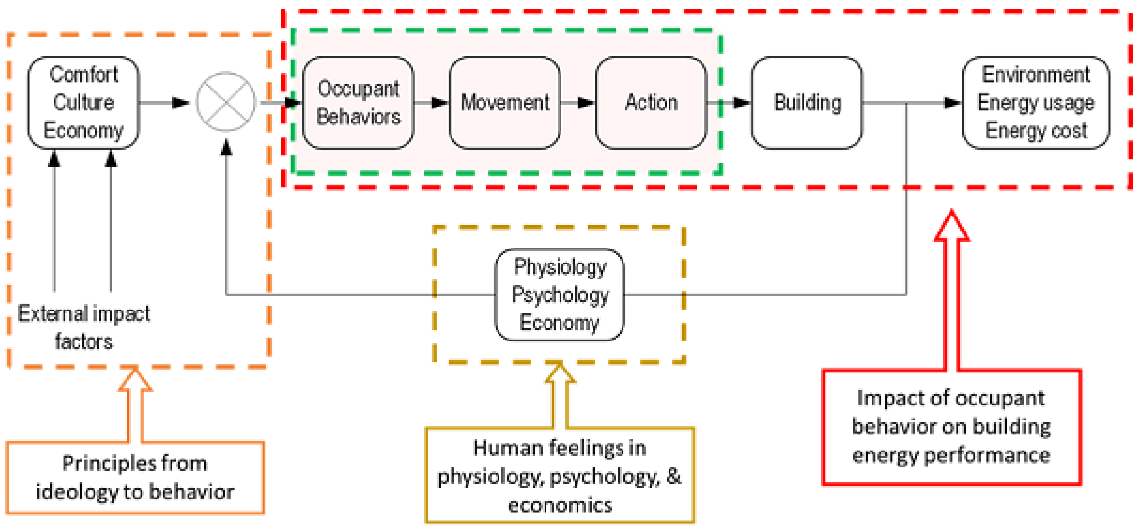


Figure 1: Relationship between occupants and buildings

The Annex 66 project was approved unanimously at the 74th Executive Committee Meeting of the International Energy Agency's Energy in Buildings and Communities Programme, held on 14th November 2013 in Dublin, Ireland. Operating Agents are Dr. Da Yan of Tsinghua University and Dr. Tianzhen Hong of Lawrence Berkeley National Laboratory. The Annex aims to (1) set up a standard occupant behavior definition platform, (2) establish a quantitative simulation methodology to model occupant behavior in buildings, and (3) understand the influence of occupant behavior on building energy use and the indoor environment. The project has five subtasks:

Subtask A - Occupant movement and presence models. Simulating occupant movement and presence is fundamental to occupant behavior research. The main objective of the subtask is to provide a standard definition and simulation methodology to represent how an occupant presents in his/her office and moves between spaces.

Subtask B - Occupant action models in residential buildings. Occupant action behavior in residential buildings affects building performance significantly. This subtask aims to provide a standard description for occupant action behavior simulation, systematic measurement approach, and modeling and validation methodology for residential buildings.

Subtask C - Occupant action models in commercial buildings. Some specific challenges of occupant behavior modeling exist in commercial buildings, where occupant behavior is of high spatial and functionality diversity. This subtask aims to provide a standard description for occupant action behavior simulation, systematic measurement approach, and modeling and validation methodology for commercial buildings.

Subtask D – Development of new occupant behavior definition and modeling tools, and integrating them with current building performance simulation (BPS) programs. This subtask will enable applications by researchers, practitioners, and policy makers and promote third-party software development and integration. A framework for an XML schema and a software module of occupant behavior models are the main outcomes.

Subtask E - Applications in building design and operations. This subtask will provide case studies to demonstrate applications of the new occupant behavior modeling tools. The occupant behavior modeling tools can be used by building designers, energy saving evaluators, building operators, and energy policy makers. Case studies will verify the applicability of the developed modeling tools by comparing the measured and simulated results.

17 countries and 123 participants from universities, research institutes, software companies, design consultant companies, operation managers, and system control companies participated in this Annex. All parties expressed an interest in developing a robust understanding of energy-related occupant behavior in buildings, via international collaboration on developing research methodologies and simulation tools that can bridge the gap between occupant behavior and the built environment. The Preparation Phase started in November 2013 and continued through November 2014. The Working Phase started in December 2014 and lasted for two and a half years. The Reporting Phase took place from July 2017 to May 2018.

Summary

In line with the activities within IEA-EBC Annex 66 Subtask C, this report includes two main sections, namely occupant behavior modeling approaches, and evaluation of occupant behavior models.

The first section presents a critical review of the occupant modeling approaches. Included are, in particular, methods used to model occupants' presence in offices as well as their use of electric lighting, blinds, windows, plug-in equipment, and thermostats. In addition, illustrative examples for each model are presented with reference to two independent datasets obtained from two different office buildings located in Ottawa, Canada and Hartberg, Austria. The models are divided into adaptive and non-adaptive domains. The former category refers to behaviors that are related to occupants' attempts to improve their comfort by adapting the building or themselves (e.g., turning on a window or changing clothing level). Whereas non-adaptive domains include those which do not directly improve comfort but are part of the occupants' objectives (e.g., presence and use of plug-in office equipment). These categories are further divided into the common model formalisms from the literature, including: schedules, Bernoulli models, discrete-time Markov models, discrete-event Markov models, regression-based models, and survival models. The strengths and weaknesses of the models are identified through the illustrative examples. Schedules can be used if significant knowledge is known about the building, but they do not allow uncertainty or dynamic occupant behavior to be characterized. Bernoulli models introduce stochasticity but are generally not suitable for predicting events (e.g., the number of light switch-on events). Markov models are powerful for predicting the likelihood of an action or change in state and the discrete-event type provides flexibility with regards to the models not being timestep-specific. Survival models predict the duration that a system will remain in a certain state (e.g., probability that a light will be turned off as a function of occupant absence) and should be primarily used for non-adaptive domains. While the examples used are from offices, the principles can also be applied to other building types.

The second section of the report addresses first a number of key conceptual requirements for the improvement of the quality of model validation practices in behavioral modelling. Both general model evaluation requirements as well as specific circumstances pertaining to models of building inhabitants are discussed. To explain these requirements in a concrete fashion, the section also includes an illustrative case study of a typical occupant behavior model evaluation process. As a paradigmatic model evaluation instance, this case study, which concerns occupants' operation of windows, provides a twofold opportunity to discuss: i) the need for clear documentation of uncertainties associated with existing behavioral models in different deployment scenarios, ii) the development of more generally applicable occupancy-related models across different contexts and different building types. It is thus expected that the present treatment can encourage developers and users of occupant behavior models toward a more systematic and critical stance.

1. Occupant behavior modeling approaches

1.1. Introduction

In many office buildings, zone level building components and systems (e.g., window blinds, electric lighting, operable windows, and thermostats) and heat gains associated with power-consuming devices and occupants are controlled or affected by occupants. The way these building components and systems are used accounts for substantial uncertainty over a building's energy use and occupants' comfort (Norford et al. 1994; Haldi and Robinson 2011). Therefore, without realistically representing occupants' interactions with the building control systems and components in building performance simulation (BPS), it is less likely that meaningful performance predictions and appropriate design decisions can be made.

Occupant interacting components in BPS tools are typically represented in terms of static schedules and power or occupant densities (Fadzli et al. 2013), meaning that these values do not change from design to design nor do they vary from individual to individual (Hoes et al. 2009). This implies that occupants are passive recipients of the indoor climates chosen for them; while in reality there is a dynamic interaction between a building and its occupants, in a way to restore comfort conditions. A substantial body of research has reviewed thermal comfort in buildings (Djongyang et al. 2010; Taleghani et al. 2013), specifically focusing on historic (Martínez-Molina et al. 2016), health care (Khodakarami and Nasrollahi 2012), and educational (Zomorodian et al. 2016) buildings.

Both passive human comfort solutions (Omer 2008) and advanced optimized control systems engineering for energy and comfort management in a building environment are widely available (Dounis and Caraiscos 2009; Shaikh et al 2014; Carlucci et al. 2015). More recently, the impacts of personalized indoor environmental condition controls on thermal comfort and energy performance have been discussed (Vesely and Zeiler 2014), especially in high performing buildings (Wang et al. 2016).

Occupants can adapt the indoor climate (e.g., by interacting with their lights, blinds, windows, thermostats) (Nicol and Humphreys 2004; Borgeson and Brager 2008; Roetzel et al. 2010; Fabi et al. 2013; Keyvanfar et al. 2014; Konstantoglou and Tsangrassoulis 2016)), and can adapt to the indoor climate by changing their clothing or activity levels (Newsham 1997; Nicol and Humphreys 2002; Morgan and de Dear 2003; Schiavon and Lee 2013). These behaviors are classified as adaptive behaviors (Gunay et al. 2013), as their primary intent is to restore comfort (thermal, visual, acoustic comfort, and indoor air quality).

On the other hand, there are non-adaptive behaviors such as plug-in equipment use and light switch-off behaviors immediately before departure from a space (Gunay et al. 2013). These behaviors are not undertaken to mitigate discomfort, but they still play a major role in a building’s energy performance. The non-adaptive behaviors are mainly driven by contextual factors (non-physical factors affecting occupants’ behaviors, habits, attitudes (Sadeghi et al. 2016)) rather than physical discomfort (O'Brien and Gunay 2014). For example, office occupants’ computer (Menezes et al. 2014; Gunay et al. 2016) and light switch off (Pigg et al. 1996) behaviors at departure exhibit a close relationship with the duration of absence following the departure.

Evaluation of building energy systems (Harish and Kumar 2016) and comfort reliability of building designs can be performed by stochastic simulation (Sulaiman and Olsina 2014). Occupant behavior and presence models mimic the interactions of occupants with zone level building components (e.g., window, blinds) or systems (e.g., lights), and with themselves (e.g., clothing insulation) (Clarke et al. 2006). They are statistical models developed upon long-term observational studies. A large community of researchers has been examining methods to model energy-related occupant behavior and to incorporate these models into the building performance simulation based design process. A number of occupant behavior data acquisition technologies, modeling methodologies and simulation coupling mechanisms for building energy efficiency have been surveyed (Jia et al. 2017). In a recent article, Gaetani et al. (2016) listed over 500 research papers on topics related to occupant behavior modeling in buildings. Consequently, the modeling methods remained fragmented amongst a large number of articles. This represents a major obstacle to those who intend to join in this research endeavor (i.e., to model energy-related human behavior in buildings).

1.2. Objective and case studies

The objective of this deliverable is to provide a critical review of the existing occupant modeling methodologies. Moreover, illustrative examples are heavily relied upon to demonstrate concepts, in such a way as to be a pedagogical resource. To this end, a comprehensive survey of the state-of-the-art literature was conducted, and examples for each modeling approach were provided upon two datasets gathered from two office buildings in Ottawa, Canada, and Hartberg, Austria.

Representing the randomness in occupants’ presence and behavior patterns entails mimicking not only the day-to-day variations of a group of occupants’ overall occupancy and behaviors but also the habitual and behavioral differences amongst these occupants (Haldi 2013; Mahdavi and Tahmasebi 2015). However, this report focuses only on the methodologies to represent the former; the latter – studying the diversity amongst different occupants – is not within the scope of this report.

The exemplars of occupant models presented in this report focus on office building behaviors. However, while the occupant modeling approaches reviewed are applicable to residential buildings too, occupant behavior patterns in residential buildings differ from those in commercial buildings. This report presents illustrative examples to demonstrate each occupant model type from the literature. To this end, occupant behavior (for lighting, blinds, and plug-in equipment use) and presence data were gathered from two office buildings. One of them is located in Ottawa, Canada and the other is in Hartberg, Austria. The photos of the buildings are shown in Figure 2.

Table 1 presents an overview of the characteristics of the data used in the illustrative examples. The exterior windows of the monitored offices are Northeast-facing in the Hartberg Building and West-facing in the Ottawa Building. The window-to-wall and window-to-floor ratios are 32% and 34% in the monitored offices in the Ottawa Building, and they are 24% and 18% in the Hartberg Building, respectively. The visible light transmission coefficient is 70% in the Ottawa Building offices, whereas it is 75% in the Hartberg office building.

The occupants of the Ottawa Building were full-time faculty members in a university, and they were full-time municipal employees in a government building in the Hartberg Building. Further details about the data can be found elsewhere (for the Ottawa Building in (Gunay et al. 2016; Gunay et al. 2016) and for the Hartberg Building in (Mahdavi et al. 2008)).

Table 1: Overview of the datasets from the two case studies

	Number of offices studied		*Lighting*	*Blinds*	*Plug loads*	**Indoor illuminance*	*Solar irradiance*
Hartberg Building	6	Monitoring period	Nov 2005 - Aug 2006	—	—	Nov 2005 – Aug 2006	
		Sampling frequency	Event-based	—	—	5 min	
Ottawa Building	10	Monitoring period	Jan 2012 – Apr 2016	Feb 2014 – Nov 2016	Nov 2014 – Mar 2016	Mar 2015 – Apr 2016	Oct 2013 – Mar 2016
		Sampling frequency	Event-based	30 min	60 min	15 min	

* The indoor illuminance was measured on the ceiling and at the workplane in the Ottawa and Hartberg buildings, respectively.

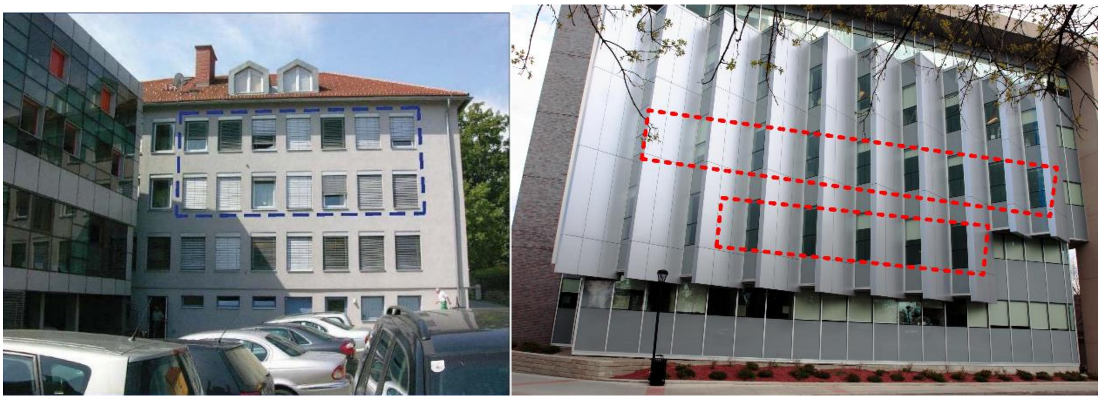


Figure 2: The buildings from which the datasets were collected (left: the Hartberg Building, right: the Ottawa Building). The dotted lines enclose the windows of the rooms studied in this report.

1.3. Modeling adaptive behaviors

In the reviewed literature, four different adaptive behavior model forms were found: (1) schedules, (2) Bernoulli models, (3) discrete-time Markov models, and (4) discrete-event Markov models. The formalisms classify whether the models predict the occupants' adaptive actions or the state of the building components with which occupants interact. Hong et al. (2015) defined the first two model forms as implicit, and the last two as explicit. Implicit models predict the states of the building components with which occupants frequently interact, whereas the explicit models directly predict occupants' interactions with these building components. Each of the modeling approaches is explored using the above case study buildings' data, after which their strengths and weaknesses are analyzed.

1.3.1. Building schedules

The traditional way of modeling adaptive behaviors is building schedules – e.g., presenting the ratio of the lights on or the mean blind occlusion rate averaged over a week or a month (Schweiker et al. 2012). Figure 3 presents the mean weekday lighting schedule for the two datasets and the lighting schedules used in the United States Department of Energy archetype office buildings (ASHRAE 2013). The data points in the plots represent the mean value by the time of day across many weekdays. As illustrated in Figure 3, the approach provides information that is easy to interpret and does not require data from indoor environmental quality sensors. This model form is established based on the assumption that the time of the week or the month of the year alone is adequate to make predictions for occupant behavior. This steady-periodicity assumption arises from the fact that indoor and outdoor environmental factors that influence adaptive behaviors tend to recur in daily or seasonal cycles. However, it is worth noting that schedules can be built from hourly (or subhourly) measured observations. For example, in an annual whole-building energy simulation, a simulationist can use 8760 hours' worth of plug-in equipment load data as a schedule, when such data are available. The steady-periodicity of schedules does not apply to these cases.

However, when a simulationist or a building operator wants to determine the outcomes of a design or a control strategy, the indoor climatic conditions that affect the occupants' behavior will inevitably change. For example, changing the glazing material and geometry, shading material and controls, and lighting fixture and controls will change indoor environmental conditions, thereby playing a role over occupants' use of lighting. Because schedules do not incorporate indoor environmental proxies (e.g., workplane illuminance) to explain occupant's behavior, these models may fail to replicate adaptive actions effectively (Hoes et al. 2009). In addition, as they are deterministic, they fail to represent the inherent randomness in occupants' behaviors.

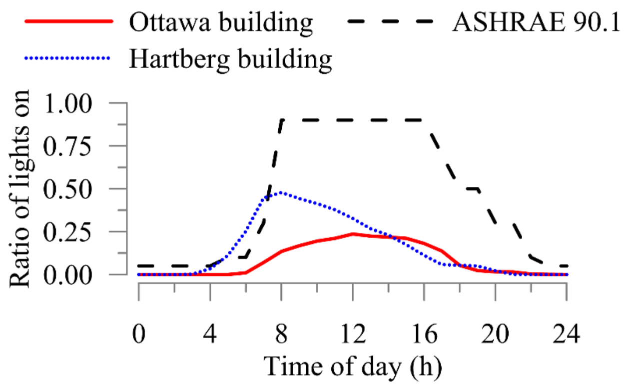


Figure 3: Lighting use schedule for weekdays in the two office buildings and the ASHRAE Standard 90.1 (ASHRAE 2013).

1.3.2. Bernoulli models

Another method employed in adaptive behavior modeling is Bernoulli random processes (Haldi and Robinson 2008; Herkel et al. 2008). Bernoulli models predict the likelihood of finding a building component with which occupants interact at a given state (i.e., a window open or closed, lights on or off). Each scatter point in Figure 4 presents the ratio of the occupied duration when the lights were on to the occupied duration at varying solar irradiance levels. For example, the probability of finding the lights on when the incident solar irradiance on the façade is less than 50 W/m^2 is 0.72 in the Ottawa building, while it is 0.74 when the horizontal solar irradiance is less than 50 W/m^2 in the Hartberg building.

This model form does not provide any information about the occupants' adaptive comfort (Gunay et al. 2015), and thus, it is appropriate to use when the objective of occupant modeling is to better represent a building's energy use – not the indoor comfort conditions. Occupants' adaptive actions are predictable with discomfort proxies (e.g., workplane illuminance is a predictor for insufficient daylight levels). On the contrary, the reversals of an adaptive action (i.e., blinds opening or light switch off) can happen long after the discomfort conditions disappear (Rubinstein et al. 1989; Foster and Oreszczyn 2001; Reinhart 2004; Sutter et al. 2006; Rijal et al. 2008). As a result, the environmental predictors often cannot explain a significant variation in the occupant controlled building components. For instance, as presented in Figure 4, even when the solar irradiance reaches to its upper limits, a considerable portion of the lights remained on in both buildings – meaning that users in these perimeter spaces do not actively adjust their blinds to exploit daylighting potential to replace electric lighting. In line with this, the blind occlusion rate exhibits an insignificant variation as a function of the incident solar irradiance on the façade of the Ottawa building – as shown in Figure 5.

Although Bernoulli occupant models have been developed with both indoor and outdoor explanatory variables in the literature (Nicol and Humphreys 2004; Haldi and Robinson 2008); Gunay et al. (2016) proposed that Bernoulli occupant models are more appropriate to be used with outdoor variables. This is because the adaptive behaviors trigger changes in the indoor environment, which contains both the explanatory and the response variables. For example, the ratio of lights on when the workplane illuminance is less than 500 lux would be zero, if the electric lighting can provide at least 500 lux at the workplane when it is switched on. Therefore, a Bernoulli model for lighting use with a predictor of indoor illuminance does not indicate the real lighting use patterns at an indoor illuminance lower than 500 lux.

The advantages of developing models with outdoor variables instead of indoor variables are the reduction in the cost of sensors and data collection, and the reduced risk of gathering biased information due to the Hawthorne effect (Humphreys and Nicol 1998; Mahdavi 2011). The major weakness in using environmental conditions as the explanatory variables with the Bernoulli models is that they cannot be used in other buildings because they neglect the influence of differences in buildings' geometry and material properties.

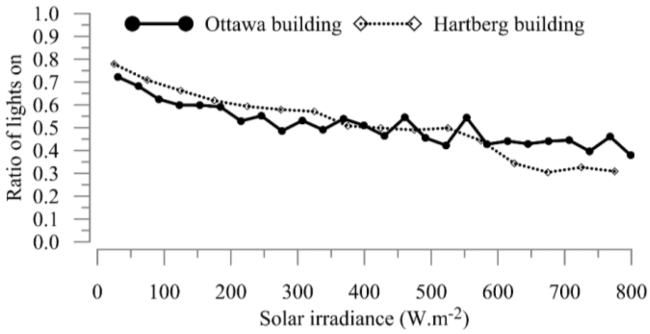


Figure 4: Bernoulli models predicting the fraction of the occupied period with lights on as a function of the solar irradiance in the Ottawa and the Hartberg building. Solar irradiance values represent the incident solar irradiance on the façade for the Ottawa building and the horizontal solar irradiance for the Hartberg building

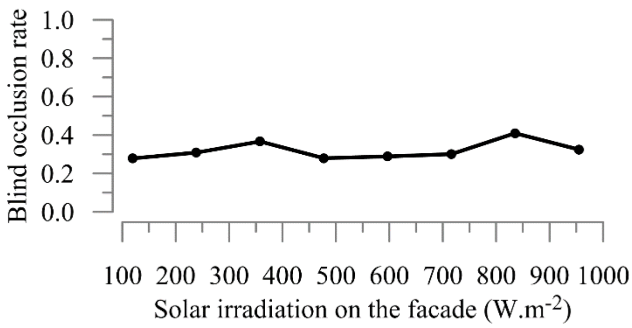


Figure 5: A Bernoulli model predicting the blinds occlusion rate as a function of the solar irradiance in the Ottawa building.

1.3.3. Discrete-time Markov models

The third method used in modeling adaptive behaviors is the discrete-time Markov chains (Fritsch et al. 1990; Lindelöf and Morel 2006; Rijal et al. 2008; Haldi and Robinson 2009; Chen et al. 2017). The discrete-time Markov models predict the likelihood of undertaking an adaptive behavior in the next timestep. They can be developed by both indoor and outdoor environmental variables because they are derived from the conditions just before occupants undertake the action. The Markov models treat adaptive actions and their reversals independently and have been suggested to predict behavior patterns more realistically. However, a common issue regarding the discrete-time Markov models is their dependency on fixed time steps (Gunay et al. 2014). They only provide the likelihood of an occupant action in the next time step. The fixed timestep concept implies that the frequency of an occupant's instances of decision-making remains constant; it is logical that these cases increase in frequency during periods in which environmental conditions rapidly change (e.g., at arrival).

Examples of the discrete-time Markov models in the literature include Haldi and Robinson (2009)'s model for window opening/closing behaviors during intermediate occupancy; Haldi and Robinson (2010)'s model for blind closing/opening behaviors during intermediate occupancy; and Reinhart (2004)'s model for light switch on behavior during intermediate occupancy.

Figure 6 presents two discrete-time Markov models predicting the likelihood of a light switch-on action in the next 15 min for the Ottawa building and the Hartberg building. The scatter points represent the ratio of occupied timesteps with a light switch-on action to the total number of occupied time steps at a particular indoor illuminance level. To calculate the discrete likelihood values, the total number of occupied timesteps when the lights were off was grouped from all occupants at each bin (25 lux for the Hartberg and 10 lux for the Ottawa building). Some of these timesteps were followed by a light switch-on action, while some were not. The ratio of those timesteps that led to a light switch-on action to the total occupied timesteps with lights off provides the likelihood of observing the light switch-on action in the next timestep. As shown in Figure 6, this ratio is significantly higher when the indoor illuminance levels are less than 50 lux.

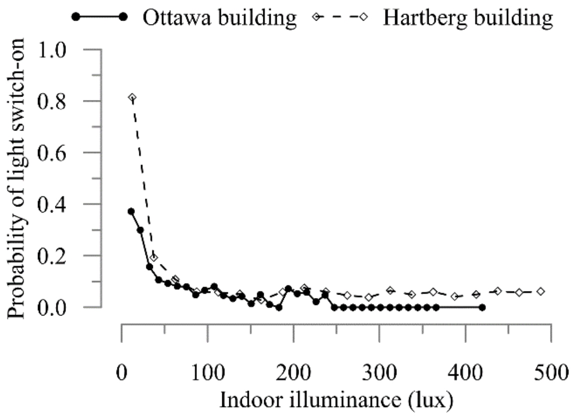


Figure 6: Discrete-time Markov models predicting the likelihood of a light switch-on action in the next 15 min as a function of the indoor illuminance. In the Ottawa building, the indoor illuminance measurements were taken on the ceiling and, they were taken at the workplane in the Hartberg building.

1.3.4. Discrete-event Markov models

Discrete-event Markov models (fourth method) link an occupant action model to an external event (Reinhart 2004; Herkel et al. 2008; Rijal et al. 2008; Yun and Steemers 2008). For example, in Reinhart (2004)'s light switch model, simulated occupants are modeled to turn on their lights more likely at arrivals (event). In Rijal et al. (2008)'s window operation model, occupants were modeled to consider window opening and closing upon a change in the predicted mean vote (event) (ASHRAE 2004). In a similar fashion, Gunay et al. (2015) treated discrete events for the light switch-on behavior as a change larger than 100 lux in the workplane illuminance levels.

When relevant events triggering the behavioral adaptation of the occupants can be identified, the models' predictive accuracy are shown to improve in contrast to discrete-time Markov models (Gunay et al. 2015). However, the discrete-event Markov modeling approach is challenged by finding an appropriate event triggering the occupant's action, to replace the timestep concept. Another limitation of this method is that its predictive performance relies on the accuracy of the external events' predictions. For example, the predictive performance of the discrete-event Markov light switch model for arrival is subject to our ability to detect the intermediate arrival and departure events accurately.

Figure 7 presents two discrete-event Markov models predicting the likelihood of a light switch-on action at arrival (including the first and intermediate arrivals) for the Ottawa and Hartberg buildings. The ratio of arrivals that result in a light switch-on action to the total arrival timesteps with lights off provides the likelihood of observing the light switch-on action in the timestep right after the arrival.

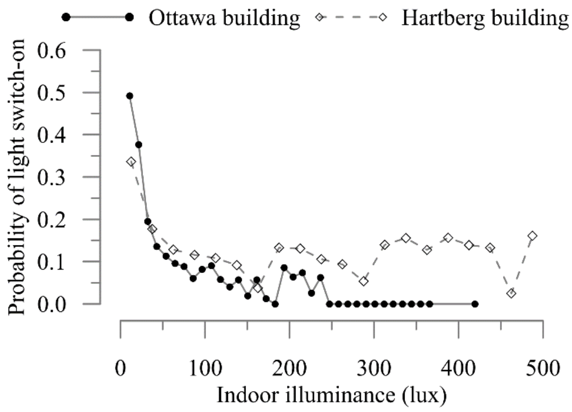


Figure 7: Discrete-event Markov model predicting the likelihood of a light switch-on action at arrival as a function of the indoor illuminance.In the Ottawa building, the indoor illuminance measurements were taken on the ceiling and, they were taken at the workplane in the Hartberg building.

1.3.5. Regression methods for adaptive behavior models

The discrete likelihood weights in adaptive behavior models (e.g., see Figures 4 to 7) are often fitted as regression models to represent the information of the model with a number of parameter coefficients and to regularize them as continuous distributions.

In the reviewed literature on adaptive behavior modeling, two regression methods were found: (1) linear regression (e.g., linear or polynomial regression) (Warren and Parkins 1984; Inoue et al. 1988; Foster and Oreszczyn 2001; Inkarojrit and Paliaga 2004) and (2) generalized linear regression (e.g., logistic, probit regression) (Nicol 2001; Clarke et al. 2006; Rijal et al. 2007;

Haldi and Robinson 2008; Inkarojrit 2008; Rijal et al. 2008; Haldi and Robinson 2009; Haldi and Robinson 2010; Haldi and Robinson 2011; Zhang and Barrett 2011; Zhang and Barrett 2012).

The shortcoming of linear regression is that it is not appropriate for probabilistic models where the response variables are bound between 0 and 1. Thus, the generalized linear regression has become the de-facto standard in adaptive behavior modeling (Haldi and Robinson 2011). It employs a nonlinear link function (e.g., probit or logit) to map the explanatory variables (e.g., indoor temperature) onto bounded response variables (e.g., the probability of observing a thermostat override). By employing the maximum likelihood method, one can develop the generalized linear models. Statistical packages for established programming environments provide built-in functions to develop generalized linear models (e.g., statsmodels in Python, glmfit or fitglm in Matlab, glm in R-programming).

Figure 8 presents a logistic regression fit for the discrete-time Markov light switch-on model for the Ottawa building. The areas of the bubble plots in Figure 8 indicate the observed occupancy duration when lights are off at each ceiling illuminance level. Note that the occupied durations are not homogeneously distributed at each illuminance level. Thus, an important consideration for building generalized linear models is to ensure that the representative number of observations is acquired from a wide-range of predictor conditions (e.g., monitoring light switch behavior from 0 to 1000 lux on the workplane, monitoring thermostat use behavior from 18 to 27°C).

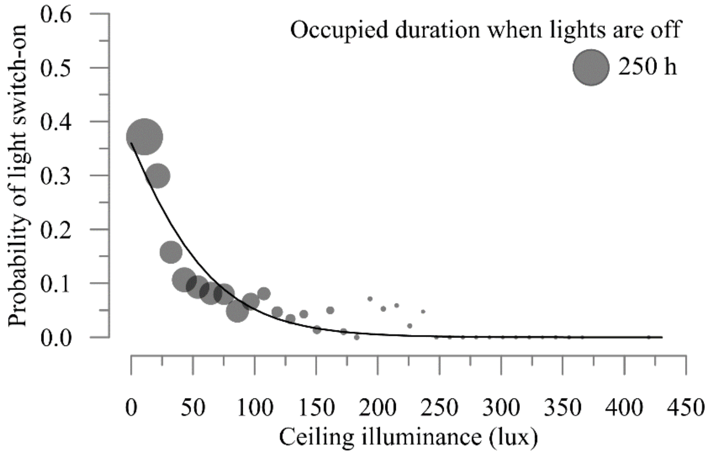


Figure 8: Probability of switching on the lights in the next 15 min (discrete-time Markov) in the Ottawa Building. The univariate logistic regression model is in the following form: $p = 1/\left(1 + e^{-(a+bE_{lux})}\right)$.

The regression models can be univariate, where the model is fitted with respect to a single predictor (e.g., predicting the blind closing action with the workplane illuminance), or multivariate where the model is fitted with respect to two or more predictors (e.g., predicting the blind closing action with the workplane illuminance and the indoor temperature). As mentioned in Haldi and Robinson (2011), increasing the number of predictors will provide diminishing improvements in the predictive accuracy.

For the evaluation of the regression models, when the dataset is large enough to be partitioned into training and validation sets, the cross-validation method should be employed to ensure the models' fitness (Haldi and Robinson 2009; Haldi and Robinson 2010). If a model is not overfitted, the model developed by the training set would be in agreement with the model developed by the data retained for the validation. Alternatively, the relative model quality can be assessed by computing the Akaike or the Bayesian information criteria. For example, Figure 9 contrasts the quality of two univariate logistic regression models (discrete-time and discrete-event Markov models) for the same dataset from the Ottawa building. By examining the Akaike information criterion values (smaller values are favorable), the discrete-time model appears to be a relatively better model for the dataset. Another metric for the assessment of the regression models is R-squared. Note that for binomial data, ordinary R-squared should not be used. If needed, the modelers should use pseudo-R-squared values to assess the fitness of the model. The readers can refer to (McCullagh and Nelder 1989) for further information on generalized linear model development, selection, and validation procedures.

Discrete-time logistic fit properties

	Coefficients	Standard Error	p-value		
				AIC	216
a	-0.573	0.063	0.000	**Null deviance**	731
b	-0.023	0.001	0.000	**Residual deviance**	123

Discrete-event logistic fit properties

	Coefficients	Standard Error	p-value		
				AIC	242
a	-0.171	0.067	0.011	**Null deviance**	824
b	-0.026	0.001	0.000	**Residual deviance**	150

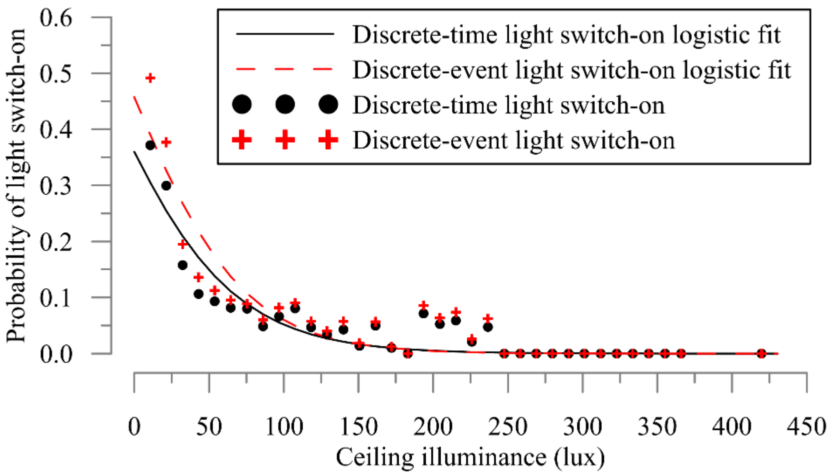


Figure 9: Probability of switching on the lights in the next 15 min (discrete-time Markov) and at arrival (discrete-event Markov) in the Ottawa Building. The univariate logistic regression model is in the following form: $p = 1/\left(1 + e^{-(a+bE_{lux})}\right)$. The properties of the regression parameters are annotated in the figure.

1.4. Modeling non-adaptive behaviors

Non-adaptive behaviors such as plug-in appliance use, light switch off at the time of departure, and blind opening are driven primarily by factors other than physical discomfort. Non-adaptive behaviors are considered those which are not undertaken to improve comfort; but instead are motivated by things such as saving energy, improving views to outside, or engaging in a task. In the reviewed literature, three modeling methods were identified for non-adaptive behavior modeling: (1) building schedules (e.g., Masoso and Grobler 2010; Menezes et al. 2012), (2) using the occupancy schedules (Mahdavi and Pröglhöf 2009), and (3) building survival models (Haldi 2010; Parys et al. 2011).

1.4.1. Schedules

Similar to the adaptive behaviors, the traditional way of modeling non-adaptive occupant behaviors is building weekly schedules. For example, Figure 10a presents the mean weekday plug-in appliance load intensity in the Ottawa building. Similar to the lighting schedules shown in Figure 2, the data points in the plots shown in Figure 10a represent the mean value by the time of day across many weekdays. This method might be appropriate for modeling the non-adaptive occupant behaviors if they were developed from a similar building archetype (Deru et al. 2011). Note that the Ottawa building had substantially different plug-in equipment schedules from the ASHRAE Standard 90.1 (ASHRAE 2013) reference office building, which is frequently used for simulating that type of building (see Figure 10a).

1.4.2. Using occupancy schedules as a predictor

Using the occupancy schedules is another model form found in the reviewed literature for the non-adaptive behavior modeling (Mahdavi and Pröglhöf 2009). The low-occupancy Ottawa office building and the ASHRAE Standard 90.1 (2013) reference office building appeared to have similar plug-in appliance load intensities when they were normalized with the occupancy rate (see Figure 10b). Recently, Mahdavi et al. (2016) developed a new model predicting the plug-in equipment usage by looking at the mean occupancy rate. This model can be considered as another example of this model form.

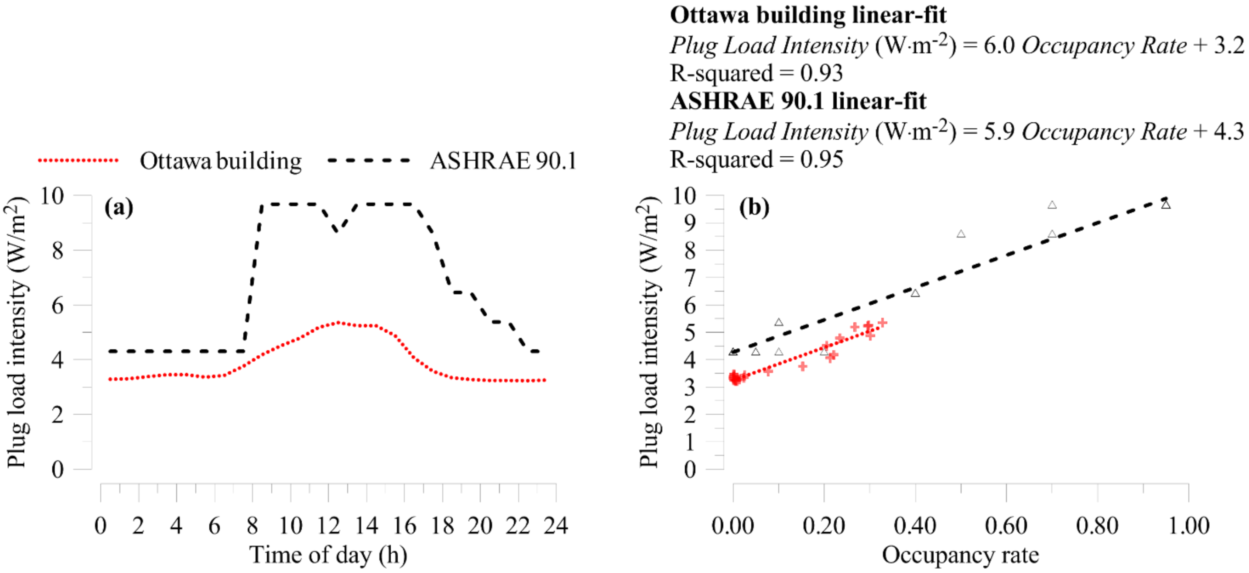
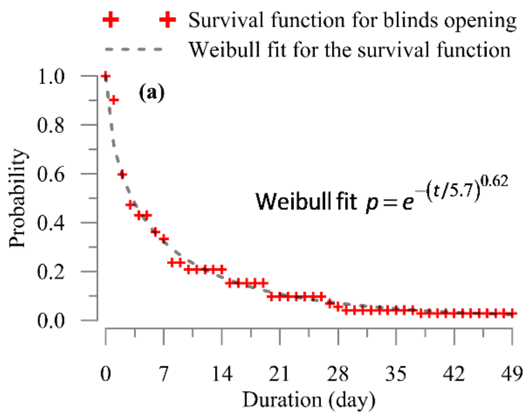


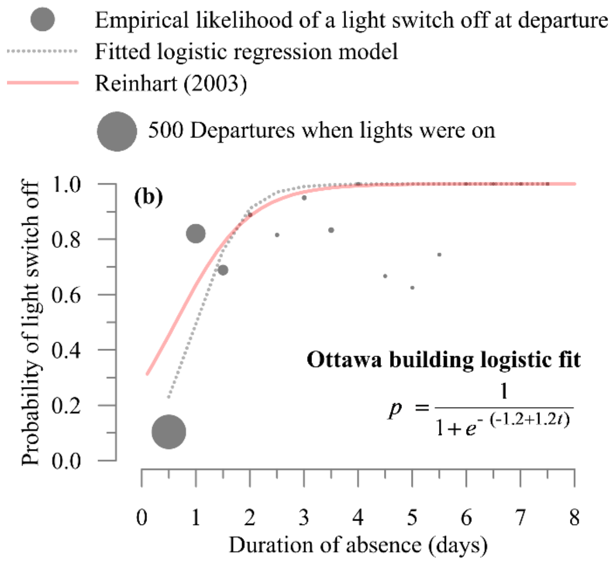
Figure 10: The average plug-in appliance load intensity on a weekday in the Ottawa building as (a) a schedule and (b) its relationship with the mean occupancy rate.

1.4.3. Survival models

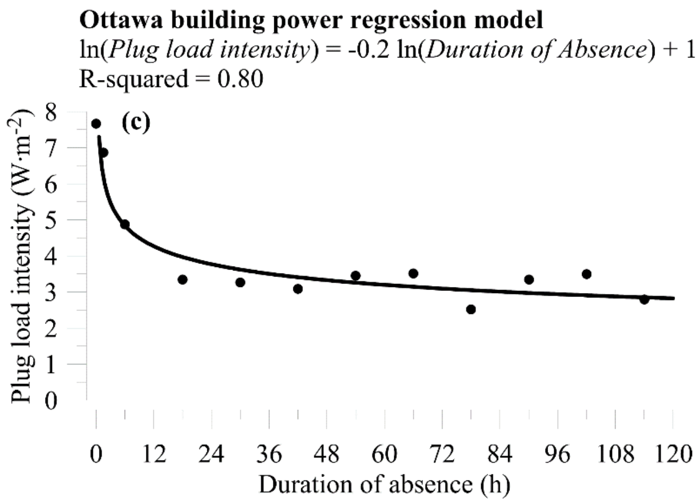
The third method used in modeling non-adaptive behaviors is the survival models. The survival models found in the reviewed literature predict the lifetime of an occupant action or the state of a building component with which occupants interact (Haldi 2010; Parys et al. 2011). For example, the likelihood of a light switch off action at departure was modeled to increase as a function of the duration of absence following the departure (Boyce 1980; Pigg et al. 1996; Mahdavi and Pröglhöf 2009). In a similar fashion, the plug-in appliance load intensities during vacancy periods were modeled as a function of the duration of the vacancy period (Gunay et al. 2016). The survival models exploit the availability of matching occupancy data to elaborate the relationship between the non-adaptive behaviors and the occupancy/vacancy state, albeit with the added complexity to collect concurrent occupancy data. Figure 11 presents three survival models built upon the data gathered from the Ottawa building. The first one (Figure 11a) presents the survival model for the time between consecutive blinds closing and opening actions. Results indicate that in 30% of the cases it takes more than a week to reopen the blinds once they are closed. This observation is in line with the literature that the users' blind opening behavior is quite infrequent (Gunay et al. 2013). The second example (Figure 11b) presents the ratio of departures with a manual light switch off action to the total number of departures when the lights were on as a function of the duration of absence. The results indicate that in almost 70% of the cases the users left their dimmable and motion detector-automated artificial lighting on (with a 30 min delay) during intermediate breaks. The third example (Figure 11c) presents the plug-in appliance load intensities during vacancy periods as a function of the duration of the absence. The model was established upon the mean plug load values at varying durations of absence. The scatter points represent the mean plug load measured at different periods of occupancy/vacancy – in 12 h bins. Results indicate that the mean plug-in equipment load per occupant was about 8 W/m^2 during occupancy, and it decreased to 3 W/m^2 during absences longer than three days. This can be interpreted as occupants' tendency to turn off their plug-in equipment increases as a function of the duration of their absence. The model mimics this behavior through a regression model in which the mean plug load exponentially reduces as a function of the length of absence. Examples of this model type were used in modeling blinds, plug-in equipment, operable windows, and lighting (Pigg et al. 1996; Reinhart 2004; Haldi and Robinson 2009; Parys et al. 2011; Gunay 2016; Gunay et al. 2016; Gunay et al. 2016). In a few cases, the survival models were also used in modeling adaptive behaviors (e.g., windows use) (Haldi and Robinson 2009; Haldi and Robinson 2010; Haldi and Robinson 2011). However, in these cases, the survival curve was modified as a function of indoor environmental variables. For example, in Haldi and Robinson's (2009) window use model, the lifetime of a window position can be predicted by a survival model which is a function of the indoor temperature. The limitation of this approach is that if the indoor environmental variable used in the model (e.g., indoor temperature) changes after making a prediction (e.g., lifetime of a window's position), the duration predicted by the initial survival model will become unrepresentative.



(a)



(b)



(c)

Figure 11: Different survival models built upon the data gathered from the Ottawa building: (a) time between consecutive blinds closing and opening actions, (b) likelihood of a light switch off at departure as a function of the duration of absence, and (c) plug-in appliance load intensity during vacancy as a function of the length of the absence period.

1.5. Modeling presence

As occupants' behaviors are conditional upon their presence, it is essential to understand and characterize the randomness inherent in occupants' presence patterns to represent their behaviors realistically. In modeling presence in buildings, three different methods have been typically used: (1) schedules (Chang and Hong 2013; Duarte et al. 2013), (2) discrete-time Markov models (Parys et al. 2010; Wang et al. 2011; Andersen et al. 2014), and (3) survival models (Wang et al. 2005).

1.5.1. Occupancy schedules

The most common occupancy modeling method is building weekly occupancy schedules – presenting the ratio of presence as a function of the time of day and the day of the week (Gunay et al. 2015; Mahdavi and Tahmasebi 2016). Figure 12 presents the weekday occupancy schedule in the two office buildings of this report and the ASHRAE Standard 90.1 (2013). Results indicate that the occupancy in the Hartberg building peaks in the morning, whereas the occupancy in the Ottawa building peaks in the afternoon. The occupancy rates in the Ottawa building – an academic office building used by professors – was noticeably lower than the Hartberg building – a government building used by municipal employees. However, the occupancy rates in both buildings were substantially lower than those of ASHRAE Standard 90.1 (2013). The advantage of this model form is that it is easy to interpret by building operators and control technicians. Moreover, it is suitable for large building scales (i.e. office floors, schools, entire large buildings). Building specific occupancy schedules provide valuable insights that can help operators choose operating schedules. Simulation experts can incorporate them quickly into building models to represent occupancy. Recently, Mahdavi and Tahmasebi (2015) introduced a method to generate occupancy time-series data (i.e., sequential presence and absence information) from an occupancy schedule.

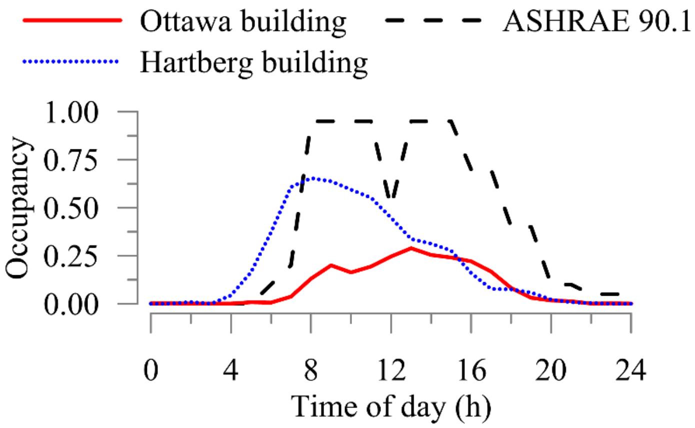


Figure 12: Occupancy schedule for weekdays in the two office buildings and the ASHRAE Standard 90.1 (ASHRAE 2013).

1.5.2. Discrete-time Markov models

The second method used in occupancy modeling is the Markov chains (Page et al. 2008; Wang et al. 2011). The model predicts the likelihood of an arrival when occupants are absent, and it predicts the probability of a departure when occupants are present. For example, Figure 13 presents discrete-time Markov models predicting the likelihood of observing a first arrival or a last departure in the next hour based on the two example datasets (Harberg and Ottawa buildings). The models were built by computing the ratio of the number of first arrivals (last departures) to the total number of unoccupied duration (occupied duration) at a certain hour of a weekday. The results indicate the occupants in the Hartberg building tend to arrive earlier and leave later than the occupants in the Ottawa building. The occupants' first arrival and last departure distributions exhibit a rather weak bimodality in the Ottawa building; meaning that occupants' first arrivals may take place in the afternoon, or their last departures can occur in the morning. This type of behavior was not observed in the Hartberg building. The strength of this approach – unlike the traditional schedule-based models – lies in the fact that the likelihood of observing an arrival or a departure from the rest of the day can be estimated given current time and the current state of presence. This may help to make midday control decisions such as temperature setbacks when the likelihood of observing an arrival is too small for the rest of the day (Gunay et al. 2015). The Markov occupancy models are also capable of creating realistic occupancy time-series which can be used in BPS models (Chang and Hong 2013). A weakness of the Markov occupancy models is that they treat arrival and departure events independently. In reality, occupants may depart early when they arrive early, or they may depart late when they arrive late (Page 2007).

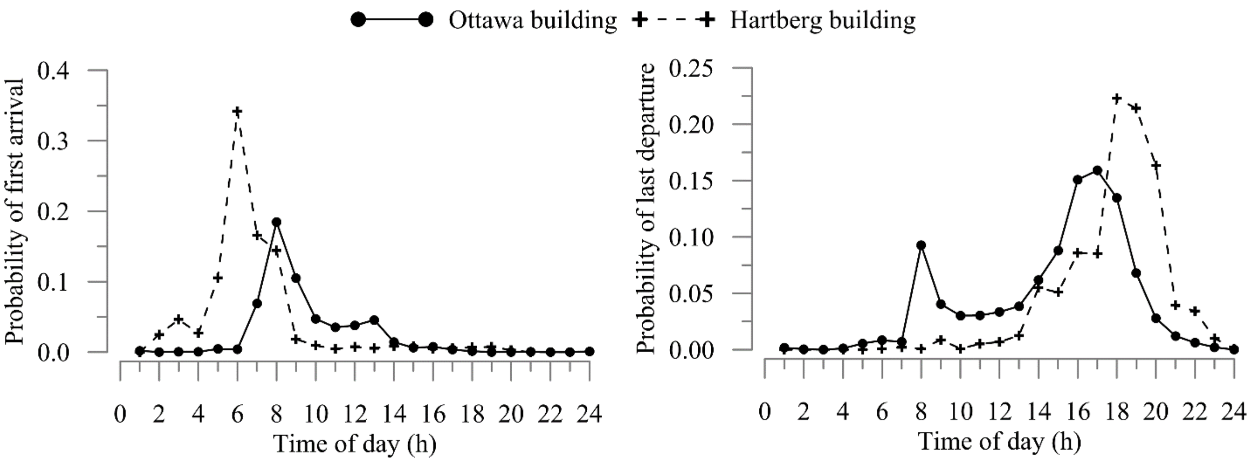


Figure 13: Discrete-time Markov models providing the likelihood of observing a first arrival or a last departure in the next hour on a weekday.

1.5.3. Survival models

Survival occupancy models predict the duration of an intermediate vacancy period following a departure or they can predict the length of an intermediate occupancy period upon an arrival (Wang et al. 2005). Note that the term intermediate vacancy period represents a coffee or a lunch break during a workday. The term intermediate occupancy period represents an occupied period between a consecutive arrival and departure. Figure 14 presents survival models predicting the duration of an uninterrupted intermediate occupancy/vacancy period for the two example datasets (Hartberg and Ottawa building). Results indicate that more than 30% of the intermediate vacancy periods were longer than 1.5 h in the Hartberg building. This interval was about 2.5 h in the Ottawa building. Similarly, 30% of the uninterrupted intermediate occupancy periods were longer than 1 h in the Hartberg building. This was about 2 h in the Ottawa building. Therefore, the occupants in the Ottawa building tend to stay in their offices for longer periods without taking breaks. However, their intermediate breaks tend to persist longer than those in the Hartberg building.

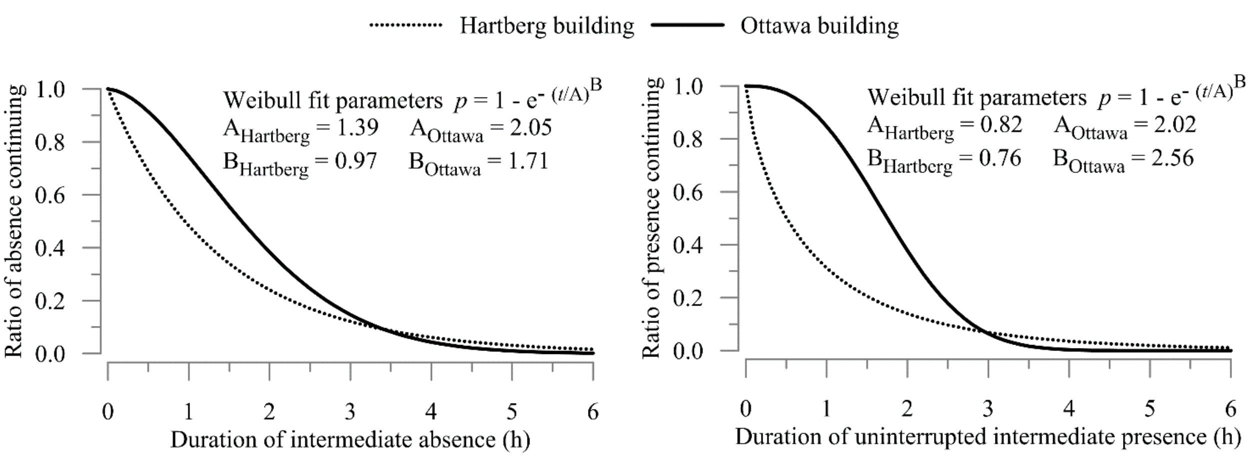


Figure 14: Survival models predicting the duration of an intermediate vacancy or uninterrupted presence period.

1.6. Analysis

The modeling methods discussed in this report have different strengths and weaknesses. Depending on the nature of the use cases involved, they can be used in the BPS-based design process. Accordingly, the following sections provide a summary of the occupant modeling forms with regard to their use cases (Section 1.6.1), as well as their strengths, and weaknesses (Section 1.6.2).

1.6.1. Inappropriate occupant behavior modeling approaches

Some of the inappropriate use cases found in the literature – which the authors suggest should be avoided in future research and development – can be listed as follows:

- *Schedules* and *Bernoulli models* for adaptive behaviors should not be used in comparing design alternatives that affect the distribution of the indoor physical stimuli of the behavior (Hoes et al. 2009). For example, changing the window-to-wall ratio will influence the users' lighting and blind use behaviors (Gilani et al. 2016). Similarly, simple variations in the interior shading systems (e.g., addition of light shelves) can affect users' lighting use behavior significantly (Sanati and Utzinger 2013). However, the schedules and Bernoulli models (developed upon outdoor conditions) will not be able to mimic the changes in users' behaviors as they overlook the link between the indoor climate and the user behavior.

- *Bernoulli models* should not be developed with indoor environmental variables affected by the behavior – e.g., developing Bernoulli lighting use models with workplane illuminance data or developing Bernoulli window use models with indoor temperature data (Haldi and Robinson 2008; Gunay et al. 2015). For example, when the lights are switched on in a typical office environment, the workplane illuminance would not fall below 300-500 lux. As a result, the model predictions for the ratio of lights on become dependent on the lighting state rather than workplane illuminance. Note that this is not an issue for the Markov models as they use the conditions just before the adaptive actions take place.

- *Discrete-time Markov models* predict the likelihood of an occupant action in the next time step (Haldi 2010). The modelers should report these timesteps. On the other hand, *discrete-event Markov models* predict the likelihood of an occupant action at an event instance. The modelers should define these event steps in which the occupant models will be invoked, and then stick to them in the energy simulation phase. Some of the early examples of occupant models found in the reviewed literature (e.g.,Hunt 1979) did not report when these models should be called during simulation (Gunay et al. 2016).

- *Survival models* should not be used for adaptive behavior modeling. For example, when Haldi and Robinson (2009) developed survival models to predict the duration windows remain open, they had to vary the shape of the survival curve for different indoor and outdoor temperatures because the window closing behavior is influenced by the indoor and outdoor temperatures. Given that the indoor and outdoor temperatures can change substantially in time, the opening durations predicted by the survival model can become inappropriate before the predicted opening period elapses. Because non-adaptive behaviors are primarily not driven by indoor climate variables, the survival analysis is more appropriate in modeling non-adaptive behaviors.

1.6.2. Strengths and weaknesses of occupant behavior modeling approaches

Some of the strengths and weaknesses of occupant behavior modeling approaches are listed as follows:

- The main advantage of using *schedules* (i.e., average value model) of an occupant behavior is their ease of development and application to a range of adaptive behaviors and building archetypes. The strength of this model form is that only a single data type is necessary to build it and it is easy to interpret for building operators and simulators. For this reason, schedules have been extensively used in BPS practice and introduced as recommendations in the design standards and codes (e.g., ASHRAE 2013; NRC 2015). Conventionally, the BPS models represent occupants via standard design conditions – occupancy levels, ventilation rates, thermostat setpoints – illustrated using schedules and threshold values, without detailed consideration of occupancy or indoor and outdoor environmental parameters. The model form is established based on the assumption that the time of the week or the month of the year alone is adequate to make predictions for the occupant behavior and presence. This assumption arises from the fact that occupancy and indoor and outdoor environmental factors that influence adaptive behaviors tend to recur in daily or seasonal cycles. However, when a building designer or operator wants to find out the outcomes of a design or a control strategy, the indoor climatic conditions that affect the occupants' behavior will inevitably change for different design alternatives. Because this model type does not incorporate the indoor environmental proxies (e.g., workplane illuminance) to explain occupants' adaptive behaviors, they may fail to characterize them under other building design and control scenarios. Moreover, the same occupant may respond differently, on different occasions, even in response to identical adaptive and non-adaptive stimuli. We may also encounter considerable differences in responses between individuals to even identical stimuli (Fabi et al. 2013). This randomness can have significant implications for building energy demand, leading to inconsistencies between simulated and actual building energy performance (Turner and Frankel 2008).

- *Bernoulli processes* provide some improvements in explaining occupants' adaptive behaviors with environmental explanatory variables. The limitation of Bernoulli processes is that models are developed based on the observations of a building component's state (e.g., window states), not the actual interactions with it (window opening or closing). In other words, these models do not describe the probability of window opening or closing, but the likelihood for a window to be found open, as a function of explanatory variables. Furthermore, Bernoulli processes ignore the particular patterns caused by occupancy events, like arrivals or departures of occupants.

- *Discrete-time Markov models* have become popular among the research community because of the straightforwardness of pairing consecutive probabilities of occupants' actions with some simulated changes in environmental conditions. Markov models fit simulation processes in BPS tools in the format of discrete timesteps. However, in reality, occupants' adaptive action events take place at irregular time intervals. In such view, discrete-time

Markov models may fail to capture behavior patterns. As an example, occupants tend to close their blinds when the conditions are bright; but they tend to leave them closed when the ambient conditions change and the risk of glare diminishes. Discrete-time Markov lighting and blind use models can mimic this better than Bernoulli lighting and blinds models because they predict the light switch or the blind closing/opening actions instead of the light state or the blind occlusion rate. Discrete-time Markov occupancy models can represent the frequency and timing of the arrivals and departures. However, the major drawback of using this form for occupancy modeling is that the consecutive arrival and departure events are treated independently from each other.

- *Discrete-event Markov models* link the calling points of an occupant action model to an external event. This modeling approach may address some of the limitations of the aforementioned methodologies. It appears to overcome some of the barriers in using fixed schedules, Bernoulli models, and discrete-time Markov models. It is, however, challenged by finding an appropriate event definition to replace the timestep concept. Another limitation of this approach is that its predictive performance relies on the accuracy of the external events' predictions.

- *Survival models* predict the duration that an occupant interacting system's state remains unchanged. This characteristic makes the modeling approach suitable for transferability to other building models and archetypes. While discrete-time Markov occupancy models may fail to capture potential dependencies between arrival and departure times – because they treat these two variables independently – survival occupancy models can tackle this limitation, by linking the timing of the arrival and departure events with each other. However, because they are continuous-time random processes, the rounding errors can be significant when used with large time steps in BPS.

1.6.3. Unresolved modeling issues and future requirements

A fundamental question when looking for a model of occupant behavior is the model choice. Starting from a monitoring campaign of occupants' presence and behaviors and predictor variables, a crucial point is defining the best model that can reproduce the relationship between occupants and the monitored variables. The models should strive to achieve a compromise between complexity and usability. An occupant model should characterize the observed occupant behavior patterns by looking at a small set of explanatory variables that are isolated from a large number of indoor and outdoor environmental indicators.

Furthermore, some adaptive actions are undertaken infrequently such as drinking hot or cold beverage, adjusting clothing level, or changing blind position (Haldi and Robinson 2011). Consequently, it becomes very expensive (and time-consuming) to gather an adequate dataset. This burden raises the question whether or not having a dynamic adaptive behavioral model is practical for building energy simulation purposes. Besides, the reference commercial building archetypes commonly used in North America (Deru et al. 2011) and in Europe (Schimschar et al. 2011; Mata et al. 2014) should be refined with the insights acquired in studying occupant behavior and presence. For example, as shown in this study, the occupancy patterns in an academic office building can be vastly different from an office building used by public employees – while the data from these buildings were vastly different from the current modeling practice (ASHRAE 2013). One area of future investigation is to develop integrated occupant models that consider comfort requirements, behavioral actions, psychological, physiological, and sociological factors altogether. These domains are currently implicitly considered in an unintegrated way to some extent. Greater modeling capability and accuracy would be afforded by an interdisciplinary occupant modeling approach going further.

1.7. Closing remarks

In this section of the report, a critical review of the occupant modeling methodologies from the literature was presented. Also, illustrative examples were developed upon two independent datasets from an academic office building in Ottawa, Canada and a government building in Hartberg, Austria. Based on the literature and the analyses of these datasets, the strengths, weaknesses, and use cases of each model form were discussed.

This report categorized the occupant models into three groups: (1) adaptive behavior models, (2) non-adaptive behavior models, and (3) occupancy models. The adaptive behaviors are occupant actions primarily undertaken to restore occupant comfort – e.g., light switch-on, blinds closing, thermostat use, window use, and clothing adjustments. The non-adaptive behaviors are actions mainly driven by contextual factors rather than physical discomfort – e.g., plug-in appliance use, light switch off at departure. The occupancy models predict occupants' presence, arrival/departure patterns, and the duration of vacancy/occupancy periods.

In the reviewed literature, the adaptive behavior models were developed as weekly schedules, Bernoulli models, and discrete-time and discrete-event Markov models. Bernoulli models predict the likelihood of finding a building component with which occupants frequently interact at a given state (e.g., ratio of lights switched on at a certain outdoor illuminance level). Markov models predict the likelihood of an adaptive action as a function of the explanatory variables (e.g., probability of a light switch-on in the next time step for the discrete-time Markov models or in the next event step such as at next arrival for the discrete event Markov models).

Non-adaptive behavior models have been developed as weekly schedules, survival models, or by using the occupancy schedules from a similar building. The survival models for non-adaptive behaviors predict the lifetime of an occupant action or the state of a building component with which occupants interact (e.g., lifetime of a blind position before it is changed). Occupancy models have been developed as weekly schedules, discrete-time Markov models predicting the timing and frequency of the arrivals/departures, and survival models predicting the duration of an uninterrupted occupancy/vacancy period.

2. Evaluation of occupant behavior models

2.1. Background

This section of the report primarily concerns the necessary conditions for the systematic evaluation procedures of models of occupants' presence and actions in buildings. To appreciate the critical importance of this issue, a brief reminder of the role of occupants' models in the larger context of building performance simulation is in order. Building performance simulation models typically require input information on context (climate), building geometry, construction, systems, and internal processes. Whereas the specification methods regarding physical building components and properties (pertaining, for example, to buildings' fabric and construction) in building performance simulation are fairly well established, representations of occupants (presence, movement, behavior, perception, and evaluation) are frequently rudimentary. Simplistic representations of people as passive and static entities have been suggested to diminish the reliability of building performance assessment and building operation planning processes (e.g., D'Oca et al. 2014; Liang et al. 2016). Rather, adequate representations of building inhabitants need to address in more detail not only building inhabitants' passive presence, but also the multi-dimensional scope and the dynamic nature of their actions (e.g., interactions with building's indoor environmental control devices and systems). A further, related phenomenon that needs to be considered in any model development activity is the inhabitants' behavioral diversity (Mahdavi & Tahmasebi 2015; O'Brien et al. 2016).

In the past, representations of buildings' inhabitants in performance simulation models have mostly consisted of fixed schedules (so-called diversity profiles) and rule-based action models. As such, it has been argued that these kinds of representations do not realistically reflect the inherent temporal fluctuations of occupancy-related processes and events (e.g., entering, leaving, and moving in buildings, operation of devices such as windows, blinds, luminaires, manipulation of control set-points, equipment usage). Thus, there has been recently a considerable number of efforts – especially by the professionals in the building performance simulation community – to develop more sophisticated dynamic models of occupant presence and actions in buildings in terms of stochastic algorithms (for example, reviewed by Parys et al. 2011) and agent-based representations (e.g., Langevin et al. 2015; Chen et al. 2016).

A significant number of such efforts have focused on the potential of probabilistic methods and associated formalisms. Thereby, a stated objective has been to replace fix schedules and rule-based actions models in performance simulation with high-resolution probabilistic models. A number of such models have been and are being incorporated in building performance simulation applications. Such efforts are undoubtedly important. However, they have not been immune to a number of misconceptions (Mahdavi 2011, 2015; Mahdavi & Tahmasebi 2016b) regarding model evaluation and application considerations. Models have been at times prematurely promoted as valid and reliable, despite wanting empirical evidence and despite lack of information regarding the down-stream deployment scenarios. The inclusion of sophisticated and realistic behavioral models in building performance assessment applications is of course desirable as such. However, it must be done in a careful and systematic manner, lest confusion and poor decision making result due to uncritical implementation and application of all kinds of insufficiently tested behavioral models.

Given this background, this section is primarily motivated by the lack of general procedures and guidelines for the evaluation of proposed user-related behavioral models. To encourage a deeper discourse in this area, we specifically formulate a number of conditions that are necessary for systematic and dependable enrichment of building performance assessment applications with behavioral representations of buildings' inhabitants. Toward this end, we use mostly assertions and findings formulated in a previous publication (Mahdavi & Tahmasebi 2016b). We discuss both general model evaluation requirements as well as specific circumstances pertaining to models of building inhabitants. The section concludes with a case study (Tahmasebi & Mahdavi 2016) to illustrate exemplary model evaluation processes. Given the rapidly evolving state of art in the area of occupancy-related model development and their integration within the workflows pertaining to the building delivery process, it would probably be premature at this point to formulate an ultimate and definitive guideline to model evaluation. The included case study thus is meant to illustrate potentially paradigmatic model evaluation steps using a comparison of a number of recently proposed behavioral models. Thereby, the main objective is to promote a rigorous process toward quality assurance while considering and integrating behavioral representations in building performance assessment tools and processes.

2.2. General principles concerning model evaluation

A central thrust of scientific activity is the development of models that are used to describe phenomena and predict events. Despite the persistence and historical evolution of model development activity across a variety of scientific disciplines (e.g., Hulley et al. 2013; Oleckno & Anderson 2002), a brief treatment of the question of model validation in the context of the occupancy-related behavioral models would be beneficial. Note that a considerable number of shortcomings in the recent model development and evaluation efforts could be shown to be the consequence of the following three circumstances:

- Firstly, systematic occupancy-related studies in the context of the built environment belong to a relatively young field of inquiry. Note that the strength of research standards in a specific domain typically results from expected utility and a critical mass of projects and researchers in that domain. As compared to many other areas of scientific inquiry (such as medical sciences or information technology), research pertaining to inhabitants' behavior in building is much less developed. A closer instance for comparison purposes would be perhaps research on human comfort in general and thermal comfort in particular. The latter has a longer tradition and is arguably better established. But even in the thermal comfort domain many open research questions and challenges persist (Schweiker & Wagner, 2016; Shipworth et al. 2016).

- Secondly, a persistent problem for both model development and model evaluation lies in the rather limited availability of large-scale observational data. Consequently, the demographic basis of the majority of proposed behavioral models is often very small. The coverage and representativeness of behavioral models of buildings' inhabitants depends on the availability and fidelity of observational data. As such data is still hard to come by, models are often developed and disseminated with insufficient empirical backing. This circumstance has also affected the aforementioned thermal comfort research, albeit to a lesser degree.
- Thirdly, behavioral models require – in principle – the concurrent consideration of multiple parameters of physical, physiological, psychological, and socio-cultural nature. To conduct field or controlled studies addressing this complex pattern of potential causal factors is indeed anything but trivial. The multifariousness of potential influencing and contributing factors to behavior actions creates as it were a kind of background "noise". Against this background, it is often difficult to discern the typically low-strength "signal" of causal factors hypothesized to be behind behavioral manifestations.

Obviously a number of the above-mentioned challenges in behavioral model development and evaluation cannot be met in the short run. Collection of vast amount of reliable observational data in the course of field studies is laborious, time-consuming, and costly. Likewise, conducting experimental behavioral studies is exceedingly difficult and the corresponding results cannot be readily generalized. This, however, does not mean that the invested community cannot improve the related conditions and processes. Toward this end, a critical assessment of the past efforts in model development and application would be essential. Specifically, avoiding certain common misconceptions would help to guide the behavioral modelling discourse in a more solid direction (Mahdavi 2015). Some of the key issues may be formulated as follows:

- Arguments pertaining to certain occupancy-related modeling approaches frequently display a certain confusion of simulation (computational, typically dynamic representation of a system's behavior) with prediction. Long-term exact predictions of buildings' energy and thermal performance are unrealistic, even under the speculative assumption that the internal (occupancy-dependent) processes could be accurately modelled. Alone the long-term unpredictability of external weather conditions falsifies claims of exact predictions. A more reasoned view of performance simulation appears to lie in its utility toward complex system analysis, rather than accurate long-term predictions. As such, it is important to understand that the frequent mismatch between simulation-based predictions and observations of energy use (the so-called performance gap) is not necessarily, or automatically, or exclusively due to behavioral factors. Long-term accurate predictions of building performance indicators are difficult to make due to an extensive list of uncertainties, pertaining not only to internal (occupancy-related) processes and external conditions, but also to assumptions regarding building fabric and building systems.

- In model comparison and evaluation discourse, the term "deterministic", which has a weighty philosophical baggage, is often used in a potentially misleading manner to characterize fixed diversity profiles (e.g., assumed fixed schedules of occupants' presence) and rule-based behavioral models. From this inaccurate terminology is then the inference made that building simulation results would be necessarily more "accurate" if occupancy-related diversity profiles and rule-based assumptions would be simply replaced with more detailed probabilistic ones (e.g., Tahmasebi & Mahdavi 2015, 2016). As such, there is not a conclusive empirical evidence and specific modeling formalisms automatically result in more accurate building performance simulations.
- A class of occupancy-related modeling efforts argue with the notion that "people behave randomly" and hence could be exclusive represented in simulation models via stochastic formalisms. There is as such nothing wrong with constructing black-box models of inhabitants' control actions nor the use of probabilistic methods toward generation of realistic patterns. In fact, many valuable lessons can be learned from careful deployment of probabilistic modeling techniques in representation of inhabitants in building performance simulation. But this does not point to the absence of a motivational (and potentially causally effective) field shaped by physiological, psychological, and social factors. Hence, efforts toward developing grey (or even white) box behavioral models is both warranted and potentially illuminating.
- Any statements about validity of specific behavioral models can be assessed only on the basis of properly and meticulously prepared documents of the model development and evaluation procedures (research design, empirical basis, hypotheses and assumed causal factors, limitations, etc.). This should enable any independent instance to retrace, comprehend, and reappraise such procedures. Moreover, behavioral models should not be claimed to be "validated" based on a limited set of observational data. Specifically, data sets for model development and model evaluation should not be conflated. Paucity of empirical information does not justify testing a model based on the same data set which was used for its development.
- It is important not to extrapolate from a single limited behavioral study to all kinds of populations, building types, locations, and climates. This is especially critical in the case of black-box models, which typically lack explicit causal explanations.
- Similar to other domains where model evaluation is critical, in the behavioral modeling field too we must safe-guard against bias. As such, internal evaluation by model developers does not provide conclusive evidence for a model's general reliability. While not easy to conduct, external evaluation procedures, double blind studies, and round robin tests are undoubtedly in a better position to convincingly support the evaluation of a model's performance.
- It is of great importance to exercise care while incorporating insufficiently documented and rudimentarily tested behavioral models in broadly used simulation applications, lest potential users are misled into assuming such models necessarily capture the "reality" of inhabitants' presence and behavior in buildings.

The last point above points to a critical challenge regarding model evaluation in the building performance evaluation domain. The reliability and appropriateness of a specific behavioral model cannot be discussed in isolation from the specific circumstances of its deployment in the simulation-assisted building performance evaluation workflow. In other words, building simulation can be deployed at very different stages of the building delivery process and for very different purposes. Consequently, it would be misguided to assume that a specific modelling approach or technique can be appropriately applied to all kinds of use cases (see Gaetani et al. 2016, Mahdavi & Tahmasebi 2016). Given the significance of this point, it is treated in more detail in the following section.

2.3. Deployment dependence of model evaluation

Performance simulation models can be generated with different levels of resolution with regard to the representation of the underlying (physical) phenomena, required (input) information, and produced results (output). Generally speaking, the choice of a specific level of resolution in these aspects is not independent of the types of queries, which the simulation model is expected to provide answers for. In this context, an important case in point pertains to possible choices in the type and resolution of representations of people's presence and behavior in building performance simulation models. The relationship between these choices and the purpose of the simulation-assisted analyses is not well understood. This, however, represents a practical problem, as it implies that adopted methods in capturing people's presence and behavior in a simulation process may in fact be inappropriate with regard to specific simulation use scenario at hand. Likewise, it can be argued that the criteria for the evaluation of the representational fidelity of people's presence and behavior in buildings are not independent of the types of the studies undertaken in the course of simulation tool deployment.

There are arguably very few studies that have explicitly addressed the fitness of occupancy-related models with regard to different simulation queries. Gupta and Mahdavi (2004) first proposed – in a different context – a perspective to view and structure the performance queries in terms of a multidimensional query space. The classification of the queries was intended to render them more suitable for analysis, resulting in enhanced responses through selection and execution of appropriate computational tools and techniques. Specific to the deployment of occupancy models, Hoes et al. (2009) used sensitivity analysis to arrive at the minimal required user model resolution with regard to a number of building performance indicators and design parameters. That is, when for example a performance indicator is determined to be more sensitive to the occupancy-related assumptions, the simulation effort should start with a more sophisticated model of occupancy (and if the performance indicator still does not fall within the required target value range, a higher resolution level should be applied). However, the focus of the study is on the design stage and it does not involve empirical data to confirm the conjecture that using more sophisticated models would necessarily provide more accurate results.

Given the multitude of scenarios (i.e., use cases involving different users, different phases of the building delivery process, different queries, etc.) in which building performance simulation can be deployed, a respective well-structured conceptual framework in terms of a multi-dimensional simulation deployment space is of utmost importance. Such a framework is not only a prerequisite for establishing a solid basis for the suitability evaluation of alternative modelling techniques and resolutions with regard to people's presence and behavior in buildings, but also contributes to substantiating the evaluation process of such modelling techniques. Table 2 briefly outlines nine dimensions that may be considered directly relevant for the selection of appropriate occupancy-related models.

Table 2: Dimensions of the proposed simulation deployment space

	Dimension	**Remarks/examples**
i	Phase in the building delivery process	Early design, detail design, HVAC systems design, building operation
ii	Purpose (or nature) of the study	Parametric study of design options, generation of energy compliance documents, HVAC system sizing, HVAC controls
iii	Domain (discipline)	Energy, thermal comfort, lighting, acoustics, fire safety
iv	Building type	Dominant function of the building (residential, commercial, educational, mixed use)
v	Indoor climate control strategy	Passive, hybrid (mixed mode), fully air-conditioned
vi	Physical destination	Building details, whole buildings, campus, district, urban
vii	Zonal destination (resolution)	Whole building, individual floors, orientations, micro-zoning
viii	Performance indicator (results)	Annual heating/cooling demand, peak heating/cooling loads, PMV
ix	Temporal resolution (horizon)	Entire life-cycle, annual, monthly, daily, hourly, sub-hourly

To demonstrate and elaborate on the desirability and usability of such a framework, we tested specific case studies, involving probabilistic and non-probabilistic occupancy models (Mahdavi & Tahmasebi 2016). The findings suggest that we cannot simply declare a priori that a particular modelling technique for generation of occupancy-related input information for performance simulation is superior to others. Rather, we must carefully consider the circumstances pertaining to the nature of application scenario such as time horizon of predictions or granularity of performance indicators. In other words, we have good reasons to suggest that the choice of an appropriate occupancy model and the criteria for evaluating its performance depends on the position of the relevant simulation-based query within the proposed application space.

2.4. An illustrative case study

In the next section, we address some of the aforementioned considerations based on a specific illustrative case study of behavioral models. The material for this case study is taken from a previously published paper of the authors that explored the reliability of various models pertaining to inhabitants' operation of windows for natural ventilation in buildings (Tahmasebi & Mahdavi 2016). In the present context, the results are not so much of interest in the original narrow sense of model comparison. Rather, we use this case study here paradigmatically to elaborate on a number of central model evaluation issues. Note that the case study itself has a number of key limitations (small set of reference empirical data from only one location, small number of models considered, etc.). We could of course argue, Popperian style, that strictly speaking, models cannot be "verified", even with large amount of affirmative evidence. A single counter-example, on the other hand, suffices to "falsify" a model. This is, however, not the point we are making here. In the domain under discussion (assessment of inhabitants' behavioral models), it would be perhaps unwise to set unrealistically high standards regarding models' predictive performance. Consequently, the treatment of this case study's material does not attempt here to definitively evaluate the selected models. For such an objective, neither the original empirical basis upon which those models were developed, nor the empirical basis we used to examine their performance are large enough. Consequently, the case study has a different purpose: The structure and embedded procedure of this external evaluation exercise provides a useful context to specifically address a number of the aforementioned model evaluation challenges.

2.5. Case study: external evaluation of window operation models

2.5.1. Introductory remarks

As already mentioned, the following treatment of external model evaluation issues uses material from a case study from one of our previous publications (Tahmasebi & Mahdavi 2016). Specific details concerning the model comparison process related to this case study may be found in the aforementioned reference. Our focus in the present context and the respective use of the case study is, however, the critical discussion of a number of typical challenges in behavioral model evaluation. Toward this end, we first provide a description of the evaluation case study, followed by an extended discussion of respective results and their general implications.

2.5.2. Selected window operation models for the external evaluation study

As a case in point, the following external evaluation study specifically addresses the performance of window operation models. We studied three existing stochastic and three simple non-stochastic models. The stochastic models (referred here as A, B, and C) are derived based on occupant behavior at office buildings and are widely referenced in the building performance simulation community. They are all Markov chain based logistic regression models that estimate the probability of window opening and closing actions based on the previous window state and a number of occupancy-related and environmental independent variables.

The non-stochastic models (referred as D, E, and F) are defined based on simple rules according to the common practice in use of building performance simulation tools without integration of stochastic models – models D and F are, for example, integrated in EnergyPlus.

In our study, we also included additional variations of models A and C (denoted as A* and C*), as the original models did not capture a key behavioral feature in the building under study where the inhabitants are requested not to leave the windows open when they leave the office due to storm damage risk. In addition, we considered two benchmark pseudo-models (denoted as G and H), whose purpose is to put the performance of the selected models into perspective. For the sake of clarity, a brief description of the aforementioned models is provided below:

- Model A, developed by Rijal et al. (2007), estimates the probability of opening and closing windows based on outdoor and operative temperature, when operative temperature is outside a dead-band (Comfort temperature ± 2°C). This model is derived based on data obtained from 15 office buildings in UK between March 1996 and September 1997.
- Model A*, a variation of Model A, always returns a closing action upon each occupant's last departure.
- Model B, developed by Yun and Steemers (2008), is derived based on summer data (from 13 June to 15 September 2006) obtained from a naturally ventilated office building in UK without night time ventilation. It estimates the probability of opening windows upon first arrival and the probability of window opening and closing actions within intermediate occupancy interval (i.e., after first arrival and before last departure) based on indoor temperature.
- Model C, developed by Haldi and Robinson (2009), estimates the probability of opening and closing actions at arrival times (first and intermediate ones), intermediate occupancy intervals, and the departure times (intermediate and last ones) based on a number of occupancy-related and environmental independent variables (see Tahmasebi & Mahdavi 2016, for the list of independent variables, and the original and adjusted estimates of the coefficients used in this study). This model has been developed based on data obtained from 14 south-facing cellular offices in a building located in the suburb of Lausanne, Switzerland for a period covering December 19th, 2001 to November 15th, 2008.
- Model C*, a variation of Model C, always returns a closing action upon each occupant's last departure.
- Model D, a non-stochastic model, operates as follows: windows are opened if indoor temperature is greater than outdoor temperature and indoor temperature is greater than 26°C. Otherwise the windows are closed.

- Model E, a non-stochastic model, can be specified as follows: Windows are opened if indoor temperature is higher than outdoor temperature and also higher than 26°C. Windows are closed if the indoor temperature is less than 22°C.
- Model F, a non-stochastic model, operates as follows: windows are opened if the operative temperature is greater than the comfort temperature calculated from the EN15251 adaptive comfort model. Following the definition of comfort temperature for free-running period in EN15251, the windows can be opened only if weighted running average of the previous 7 daily average outdoor air temperatures is above 10°C and below 30°C.
- Model G, a benchmark pseudo-model, "predicts" windows are always open.
- Model H, a benchmark pseudo-model, "predicts" windows are always closed.

In case of the stochastic window operation models, to conduct the evaluation in a comprehensive manner, we used both original and adjusted coefficients of the logit functions. Whereas the original coefficients are published by model developers, the adjusted coefficients are obtained from re-fitting the models to a separate set of data obtained from the building under study in the calibration period. We specify the models with original coefficients with a subscript "O" and the ones with calibrated coefficients with a subscript "C". As mentioned before, the latter option (adjusting model coefficients based on observations in actual buildings) has no relevance to model deployment scenarios pertaining to building design support, but may be of some interest in operation scenarios of existing buildings.

The above described process of model selection and specification of the external evaluation study already highlights some of the typical challenges in the external validation studies of behavioral models. Aside from not having gone through a prior external validation study, most published models are limited even in the scope of the underlying internal validation: The published models are often derived based on limited data – typically from a single building – rendering those as non-representative in statistical terms (population, climate, building typology, etc.). Moreover, even for this limited base, models' documentations typically leave many questions open or include questionable assumptions (for instance, the assumption that inhabitants' degree of freedom in operating windows is independent of facility management issues in a typical office building). Likewise, hidden assumptions pertaining, for example, to the assumed one-to-one relationship between an inhabitant and a window, make it difficult for the user to judge if and to which extent socially relevant interaction patterns between inhabitants and the related implications for the window operation are captured in the model.

2.5.3. Empirical data for model calibration and evaluation

An office area at TU Wien (Vienna, Austria) including an open space with multiple work-stations and a single-occupancy closed office acted as the data source for external model assessment. We specifically focused on seven workstations, at which each occupant has access to one manually operable casement window. The occupants' presence, state of windows and a number of indoor environment variables (including air temperature, humidity, and CO_2 concentration) are monitored on a continuous basis. Outdoor environmental parameters (including air temperature and precipitation) are also continuously monitored via the building's weather station. For this study, we used 15-minute interval data from a calendar year (referred to as calibration period) to calibrate the coefficients of stochastic window operation models. As such, this option is only of interest, if the model deployment scenario involves already existing buildings (e.g., model use for optimization of building operation). A separate set of data obtained from another calendar year (referred to as validation period) was used to evaluate the predictive performance of the models.

Note that, in this paradigmatic scenario, efforts were made to satisfy a number of generic model evaluation requirements formulated in the first section of this paper. These included, for example, collection of long-term high-resolution data, a rather rigorous data quality check, and obviously separate data sets for calibration of model coefficients and model comparison. However, a central problem remains: Data available for model evaluation was in this case only from one building and for a relatively small number of inhabitants. This circumstance may remain, at least for some time, unavoidable (large repositories of observational data from different locations and building types are, while highly desirable, not available). This underlines the importance of candid and detailed model documentation, as alluded to in the introduction of the paper.

2.5.4. Calibrated simulation model of the office area

The previous studies on evaluation of stochastic window operation models (Schweiker et al. 2012, Fabi et al. 2015) did not address models' feedback. This circumstance represents a special problem in behavioral model validation, as the impact of behavioral models' output (for instance window states) on the models' input (for instance indoor temperature) is ignored. It is of course logically impossible to obtain empirical data matching every possible sequence of actions predicted by behavioral models. Hence, we need to emulate building's response to behavioral impulses virtually, i.e., via calibrated simulation. Therefore, we suggest the use of a calibrated simulation model as a platform for evaluation of behavioral models whose output (e.g., window states) influences models' input (e.g., indoor temperature). This necessitates a model that can reliably represent the building's behavior.

For the purposes of this case study, we first subjected the building model to an optimization based calibration to adjust the fixed parameters governing the multi-zone air flow simulations (for the details of the calibration procedure, see Tahmasebi & Mahdavi 2012). Secondly, we incorporated the monitored data pertaining to occupancy, plug loads, use of lights, and operation of heating system into the calibrated building model as a set of full-year data streams in terms of 15-minute intervals. This data set was obtained in the validation period. The resulting model, when fed with actual window operation data as the benchmark model, predicts the hourly indoor temperatures in validation year with a Normalized Mean Bias Error of 2.8% and a

Coefficient of Variation of Root-Mean-Square Error of 4.8%. The low values of these indicators (as compared with the criteria introduced in ASHRAE Guideline 14-2014) show the relatively high accuracy of model.

The described building simulation model served as a platform, into which the selected window operation models were integrated, such that in each variation of the building model, the occupants' interactions with windows are represented using one of the selected window models. For each occupant in the building, individual occupancy data and zone-level indoor environmental factors are provided for the window operation model. That is, at each simulation time-step, the window model is executed separately for each occupant. We also built a benchmark model, which contained the actual operation of windows based on the monitored data obtained in the validation period.

As using calibrated building performance simulation for evaluation of occupant behavior models necessitates the deployment of real-year – preferably on-site – weather data, the building model was exposed to the outdoor environmental conditions in the validation period. This was accomplished by generating a weather data file from the on-site weather station measurements. The measured dataset included outdoor air temperature, air humidity, atmospheric pressure, global horizontal radiation, diffuse radiation, wind speed, and wind direction.

2.5.5. Evaluation scenarios for window operation predictions

We evaluated the performance of window operation models to predict inhabitants' interactions with windows for a one-year-long validation period, whereby the models are fed with monitored occupancy-related and outdoor environmental data from the same period according to their independent variables. The required indoor environmental factors, however, are provided from the calibrated building simulation output to include the models' feedback. That is, the calibrated building performance model simulates the impact of window operation models' output on indoor environmental input.

2.5.6. Evaluation statistics

One of the fundamental challenges of evaluation procedures pertaining to behavioral models of building inhabitants pertains to the paucity of systematically classified model performance metrics. The pertinent professional community has arguably not converged toward a systematic and expressive set of statistics for behavioral models' predictive performance. Some of the responsible factors for this negligence were already alluded to in the introductory sections of this paper. Given the variety of domains and application scenarios of behavioral models, the definition of a definitive set of evaluation statistics is indeed unlikely to be a trivial undertaking.

Whereas an ultimate ontology of fit-for-purpose metrics for behavioral model evaluation cannot be provided here (and may be even ultimately unattainable), a potentially important first attempt can be made. Behavioral models typically aim at predictions of "states" and "events" (or "actions"). In this taxonomy (Mahdavi 2011), events can be system-related (e.g., switching lights on/off) or occupancy-related (e.g., entering into – or leaving – a space). States can refer to systems (e.g., position of shades/windows), indoor environment (e.g., temperature, illuminance), outdoor environment (e.g., solar radiation), and inhabitants' presence (i.e., present versus absent).

The central step in model evaluation is of course the comparison of predicted and monitored events and states. We suggest that, from the large number of indicators, which have been used in previous – predominantly internal – evaluation studies of inhabitants behavioral models (as well as in studies in relatively close fields such as thermal comfort), two broad categories can be inferred: The indicators addressing aggregate aspects of models' predictions, and the indicators addressing the interval-by-interval congruence between predictions and measurements. In other words, whereas the first category "vertically" aggregates observations and predictions independently before the overall comparison, the second category compares first "horizontally" time series data pairs, which can then be further processed statistically. Illustrative listings of these two types of indicators are provided in Figure 15. Note that in this framework, we have grouped indicators, which address aggregate traits of the predictions (such as total number of actions, median state durations, etc.) along with indicators, which address the proximity of predicted probability distributions to those of the measured ones (such as Jensen-Shannon divergence).

It can be argued that while a superior performance in terms of aggregate indicators is specifically desired in simulation studies geared at performance levels over longer periods of time (such as conventional use of building performance simulation models for estimation of annual energy demands), the indicators resulting from interval-by-interval contrast of predictions and measurements are of more interest in studies, in which short-term performance predictions play a central role (e.g., predictive building systems control).

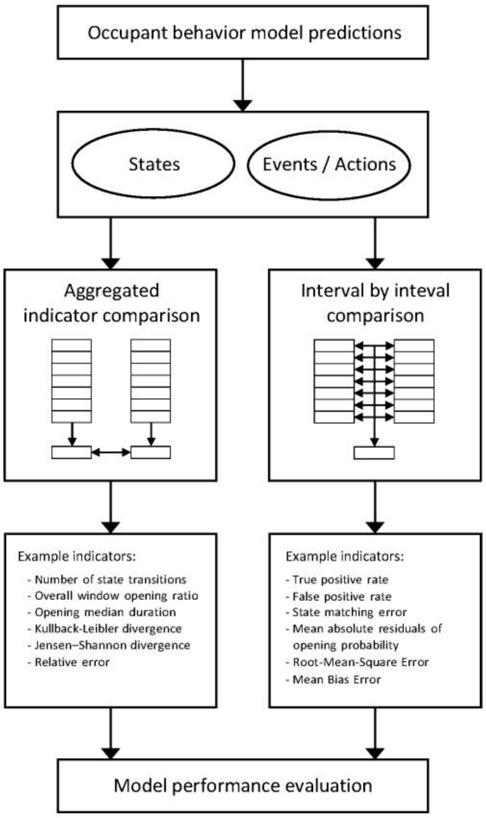


Figure 15: Categories, aggregation structures, and example indicators for occupant behavior model evaluation

For the purpose of the current case study, we used the following indicators to evaluate the predictive performance of window operation models. Note that whereas the first selected indicator in the following list belong to the interval-by-interval comparison category, the last three are typical for the aggregated indicator category:

- Fraction of correct open state predictions [%]: This is the number of correctly predicted open state intervals divided by the total number of open state intervals.
- Overall fraction of open state [%]: This is the total window opening time divided by the observation time.
- Mean number of actions per day [d^{-1}] averaged over the observation time.
- Open state durations' median [hour].

To ensure the robustness, transparency, and integrity of model evaluation procedures, the selection of reliable, expressive, and consistent model performance metrics is indispensable. Related future efforts in this direction are thus of utmost importance.

2.5.7. Results

To better illustrate the performance of models in terms of different evaluation indicators, Figure 16 to Figure 18 show the models' prediction errors under consideration of their feedback. In addition to the graphical representation of data, Table 3 provides a numeric overview of the relative deviations of predictions from corresponding observations.

Table 3: Relative deviation of the predictions from the observed behavior in terms of five evaluation indicators obtained from model executions with feedback

Models	Model type	Coefficients	Adjustment for the absence of nighttime ventilation	Relative deviation from observed behavior [%]			
				Open state predictions	Fraction of open state	Number of actions	Median opening duration
A_o	Stochastic	Original	No	56.0	289.7	81.9	962.9
B_o				58.2	213.6	1775.9	71.4
C_o				45.8	464.1	74.7	2017.4
A_o*	Stochastic	Original	Yes	52.8	20.0	25.9	225.1
C_o*				69.1	9.6	34.4	155.6
A_c	Stochastic	Calibrated	No	58.7	268.6	84.4	1033.1
B_c				55.4	28.3	13.0	57.7
C_c				52.1	323.3	40.8	112.6
A_c*	Stochastic	Calibrated	Yes	55.6	3.1	35.2	209.9
C_c*				64.6	0.1	15.2	84.1
D	Non-stochastic	-	-	64.0	7.3	352.1	85.7
E				45.7	52.7	18.2	285.7
F				55.9	65.0	541.9	85.7

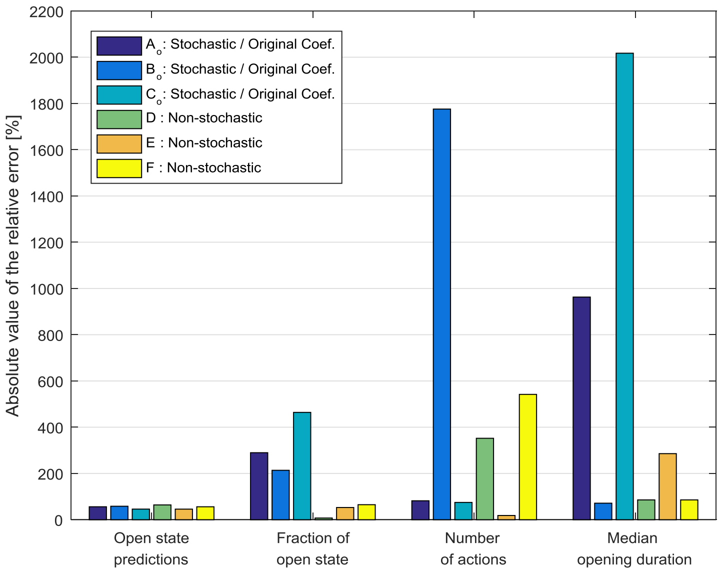


Figure 16: Errors of stochastic window operation models with original coefficients and no adjustment as well as non-stochastic models in terms of five evaluation statistics

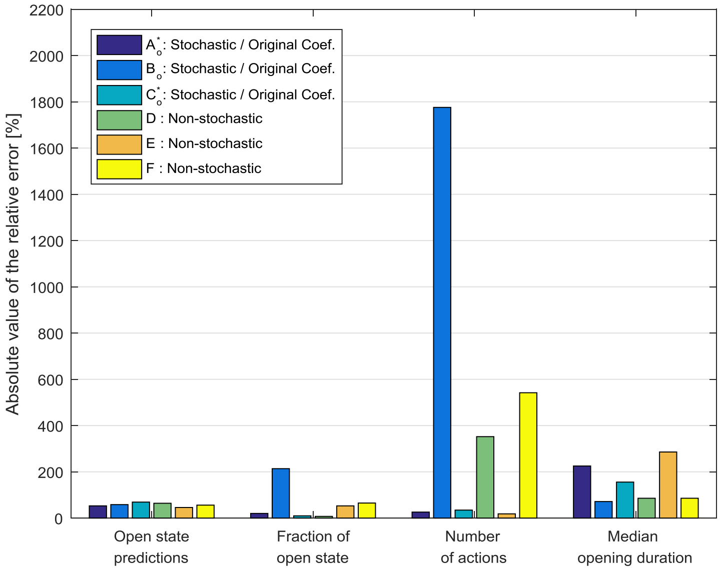


Figure 17: Errors of stochastic window operation models with original coefficients and adjusted to buildings without night time ventilation as well as non-stochastic models in terms of five evaluation statistics

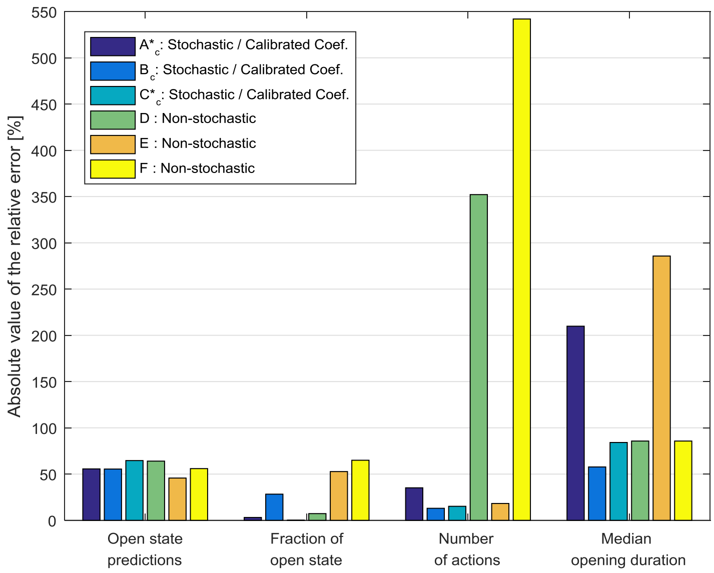


Figure 18: Errors of stochastic window operation models with calibrated coefficients and adjusted to buildings without night time ventilation as well as non-stochastic models in terms of five evaluation statistics

2.5.8. Discussion

A fundamental question with regard to the application of behavioral models concerns their capability in reproducing empirical observations. We may thus first ask if the models could, in the present case, provide acceptable approximations of the observations. Assuming an acceptability threshold of ±20% for the relative error of model predictions as a benchmark, we must conclude that without adjustments (night-time ventilation, calibrated coefficients), none of the studied models performs satisfactorily (see Table 3 as well as Figure 16). However, the nighttime ventilation adjustment markedly improves the performance of the stochastic models A_o* and C_o* (see Figure 17). Furthermore, calibrating the coefficients of stochastic models via observational data results in a significant improvement of their predictive performance (see Figure 18).

Note that, as stressed before, the presented case study was based on a limited set of empirical data obtained from one office area. We consider the underlined shortcomings valid and worthy of serious attention in future studies. However, we do not suggest the findings can be extrapolated to the modelling efforts in different contexts. Ongoing and future – more extensive – cross-sectional investigations in this area are expected to utilize a larger empirical foundation and thus lead to more representative and inclusive model evaluations. Specifically, while calibration of occupant behavior models is not feasible in majority of building performance simulation efforts, similar external validation studies can also contribute toward a repository of coefficients for the use of existing occupant behavior models in different contexts.

Aside from these specific case study results regarding the performance of the selected models, we would like to highlight a number of observations that are relevant to the model evaluation discussion in general:

- As noted earlier, a general problem in both development and evaluation of behavioral models pertains to the paucity of empirical data. For instance, models A and B were solely based on office buildings in UK (15 in case of model A and 1 in case of model B), whereas model C was based on one office building in Switzerland. Moreover, the monitoring period for data collection was rather limited in case of models B (four months).
- Earlier in the paper, we suggested that a sound model evaluation process requires the availability of clear and detailed model documentations. This condition is often ignored and was not also fully met in our case study. For instance, in case of model A, the treatment of night time ventilation was not clearly described. Likewise, in case of model C, it was not clear that the parameter included for closing window upon last departure does not suffice to make the model with original coefficients applicable for buildings without night time ventilation.
- As suggested previously, model developers should ideally conduct an internal validation via separate developmental and evaluative data sets. In the present case study, this was not done in case of models A and B. In case of model C, the publication introducing the model suggests that a "cross-validation" was performed. Note that only the publication related to model C included some model validation metrics. However, the types, coverage, scope, and suitability of performance metrics for behavioral models remains an open challenge.
- We suggested that a sound model documentation should entail comments on the applicability of the proposed models (e.g., with regard to building type, location, climate, deployment scenario). The documentation of the models selected for our case study did not provide such comments.

All in all, the above illustrative external evaluation study underlines a number of challenges in the evaluation process of behavioral models. These include the paucity of underlying empirical information with sufficiently high quality and of representative nature, shortcomings in model documentation, model input requirements that cannot be met in realistic model deployment situations, problems associated with model coefficients and their calibration, lack of a set of comprehensive, adequate, and universally accepted model performance metrics, and – last but not least – the problem of feedback, i.e., the inclusion of the predicted actions' impact on environmentally relevant model input variables.

2.6. Conclusions

Building performance assessment tools and methods can be improved if they are enriched with high-resolution representations of inhabitants. Many recent model development efforts have explored the potential of detailed mathematical formalisms for such representations. However, rigorous external evaluation processes are needed to ensure the usability and reliability of occupancy-related behavioral models. Given the lack of related general procedures and guidelines, we formulated a number of relevant conditions and requirements. Furthermore, we presented a demonstrative model evaluation study involving a number of recently proposed window operation models. Our concern was not only to highlight the observed large deviations from reality underlined in this specific case. Rather, as a paradigmatic model case, the external window operation evaluation study provided us with the opportunity to point to the need for clear documentation of associated uncertainties with existing behavioral models in different deployment scenarios as well as development of more generally applicable occupancy-related models. Definition and pursuit of rigorous model validation procedures in the behavioral modelling field may be seen as work in progress. As a consequence, both model developers and potential users would be well-advised to be careful with regard to introduction and application of behavioral models pertaining to inhabitants' actions in buildings. Specifically, statements concerning models' validity and overall applicability in the building delivery process would be of little credibility without comprehensive empirical backing and careful model testing procedures.

List of figures

.

List of tables

References

Andersen, P. D., A. Iversen, H. Madsen and C. Rode (2014). "Dynamic modeling of presence of occupants using inhomogeneous Markov chains." *Energy and Buildings* 69: 213-223.

ASHRAE (2013). ANSI/ASHRAE/IES Standard 90.1-2013 -- Energy Standard for Buildings Except Low-Rise Residential Buildings, *ASHRAE*.

ASHRAE (2004). "Standard 55-2004, Thermal environmental conditions for human occupancy." *American Society of Heating, Refrigerating and Air-Conditioning Engineering, Atlanta, GA*.

Borgeson, S. and G. Brager (2008). Occupant Control of Windows: Accounting for Human Behavior in Building Simulation.

Boyce, P. (1980). "Observations of the manual switching of lighting." *Lighting Research and Technology* 12(4): 195-205.

Carlucci, S., F. Causone, F. De Rosa, L. Pagliano (2015). "A review of indices for assessing visual comfort with a view to their use in optimization processes to support building integrated design." *Renew Sustain Energy Rev* 47: 1016–33.

Chang, W.-K. and T. Hong (2013). "Statistical analysis and modeling of occupancy patterns in open-plan offices using measured lighting-switch data." *Building Simulation* 6(1): 23-32.

Chen, Y., T. Hong, X. Luo (2017). An agent-based stochastic occupancy simulator. Building Simulation.

Clarke, J., I. Macdonald and J. Nicol (2006). "Predicting adaptive responses-simulating occupied environments."

Deru, M., K. Field, D. Studer, K. Benne, B. Griffith, P. Torcellini, B. Liu, M. Halverson, D. Winiarski and M. Rosenberg (2011). "US Department of Energy commercial reference building models of the national building stock."

Djongyang, N., R. Tchinda, D. Njomo (2010). "Thermal comfort: A review paper." Renew Sustain Energy Rev 14: 2626–40.

D'Oca S., Valentina Fabi, Stefano P. Corgnati, Rune Korsholm Andersen (2014): Effect of thermostat and window opening occupant behavior models on energy use in homes, BUILD SIMUL (2014) 7: 683–694.

Dounis, A.L., C. Caraiscos (2009). "Advanced control systems engineering for energy and comfort management in a building environment—A review." *Renew Sustain Energy Rev* 13 : 1246-61

Duarte, C., K. Van Den Wymelenberg and C. Rieger (2013). "Revealing occupancy patterns in an office building through the use of occupancy sensor data." *Energy and Buildings* 67: 587-595.

Fabi, V., R. V. Andersen, S. P. Corgnati and B. W. Olesen (2013). "A methodology for modelling energy-related human behaviour: Application to window opening behaviour in residential buildings." *Building Simulation* 6(4): 415-427.

Fabi V., Andersen R.K. & Corgnati S., 2015. Verification of stochastic behavioural models of occupants' interactions with windows in residential buildings, Building and Environment, 94(1), pp 371–383, doi:10.1016/j.buildenv.2015.08.016.

Fadzli Haniff, M. H. Selamat, R. Yusof, S Buyamin and F. Sham Ismail (2013) "Review of HVAC scheduling techniques for buildings towards energy-efficient and cost-effective operations." *Renew Sustain Energy Rev* 27 :94–103.

Foster, M. and T. Oreszczyn (2001). "Occupant control of passive systems: the use of venetian blinds." *Building and Environment* 36(2): 149-155.

Fritsch, R., A. Kohler, M. Nygård-Ferguson and J. L. Scartezzini (1990). "A stochastic model of user behaviour regarding ventilation." *Building and Environment* 25(2): 173-181.

Gaetani, I., P.-J. Hoes and J. L. Hensen (2016). "Occupant behavior in building energy simulation: Towards a fit-for-purpose modeling strategy." *Energy and Buildings* 121: 188-204.

Gilani, S., W. O'Brien, H. B. Gunay and J. S. Carrizo (2016). "Use of dynamic occupant behavior models in the building design and code compliance processes." *Energy and Buildings* 117: 260-271.

Gunay, H. B. (2016). Improving energy efficiency in office buildings through adaptive control of the indoor climate. PhD, *Carleton University*.

Gunay, H. B., W. O'Brien and I. Beausoleil-Morrison (2013). "A critical review of observation studies, modeling, and simulation of adaptive occupant behaviors in offices." *Building and Environment* 70: 31-47.

Gunay, H. B., W. O'Brien and I. Beausoleil-Morrison (2015). "Development of an occupancy learning algorithm for terminal heating and cooling units." *Building and Environment* 93, Part 2: 71-85.

Gunay, H. B., W. O'Brien and I. Beausoleil-Morrison (2015). "Implementation and comparison of existing occupant behaviour models in EnergyPlus." *Journal of Building Performance Simulation*: 1-46.

Gunay, H. B., W. O'Brien, I. Beausoleil-Morrison and J. Bursill (2016). "Implementation of an adaptive occupancy and building learning temperature setback algorithm." *Ashrae Transactions* 122(1).

Gunay, H. B., W. O'Brien, I. Beausoleil-Morrison and S. Gilani (2016). "Development and implementation of an adaptive lighting and blinds control algorithm." *Building and Environment*.

Gunay, H. B., W. O'Brien, I. Beausoleil-Morrison, R. Goldstein, S. Breslav and A. Khan (2014). "Coupling stochastic occupant models to building performance simulation using the discrete event system specification formalism." *Journal of Building Performance Simulation* 7(6): 457-478.

Gunay, H. B., W. O'Brien and I. Beausoleil-Morrison (2016). A toolkit for developing data-driven occupant behaviour and presence models. eSim 2016. Hamilton, Ontario.

Gunay, H. B., W. O'Brien, I. Beausoleil-Morrison, S. D'Oca and S. P. Corgnati (2015). On modelling and simulation of occupant models. Building Simulation. India, *IBPSA*.

Gunay, H. B., W. O'Brien, I. Beausoleil-Morrison and S. Gilani (2016). "Modelling plug-in equipment load patterns in private office spaces." *Energy and Buildings*.

Gupta, S., Mahdavi, A. (2004). Exploring performance query space, SimBuild 2004, IBPSA-USA National Conference Boulder, CO.

Haldi, F. (2010). "Towards a Unified Model of Occupants" Behaviour and Comfort for Building Energy Simulation."

Haldi, F. (2013). A probabilistic model to predict building occupants' diversity towards their interactions with the building envelope. *Proceedings of the international IBPSA conference, Chambery, France*.

Haldi, F. and D. Robinson (2008). "On the behaviour and adaptation of office occupants." *Building and Environment* 43(12): 2163-2177.

Haldi, F. and D. Robinson (2009). "Interactions with window openings by office occupants." *Building and Environment* 44(12): 2378-2395.

Haldi, F. and D. Robinson (2010). "Adaptive actions on shading devices in response to local visual stimuli." *Journal of Building Performance Simulation* 3(2): 135-153.

Haldi, F. and D. Robinson (2011). "The impact of occupants' behaviour on building energy demand." *Journal of Building Performance Simulation* 4(4): 323-338.

Haldi, F. and D. Robinson (2011). "Modelling occupants' personal characteristics for thermal comfort prediction." *International journal of biometeorology* 55(5): 681-694.

Harish, V.S.K.V., A. Kumar (2016). "A review on modeling and simulation of building energy systems." *Renew Sustain Energy Rev* 56: 1272–92

Herkel, S., U. Knapp and J. Pfafferott (2008). "Towards a model of user behaviour regarding the manual control of windows in office buildings." *Building and Environment* 43(4): 588-600.

Hoes, P., J. L. M. Hensen, M. G. L. C. Loomans, B. de Vries and D. Bourgeois (2009). "User behavior in whole building simulation." *Energy and Buildings* 41(3): 295-302.

Hong, T., S. C. Taylor-Lange, S. D'Oca, D. Yan and S. P. Corgnati (2015). "Advances in research and applications of energy-related occupant behavior in buildings." *Energy and Buildings*.

Hulley, S.B., S.R. Cummings, W.S. Browner, D.G. Grady, T.B. Newman (2013). Designing clinical research, Lippincott Williams & Wilkins, 2013.

Humphreys, M. A. and J. F. Nicol (1998). "Understanding the adaptive approach to thermal comfort." *Ashrae Transactions* 104: 991.

Hunt, D. (1979). "The use of artificial lighting in relation to daylight levels and occupancy." *Building and Environment* 14(1): 21-33.

Inkarojrit, V. (2008). "Monitoring and modelling of manually-controlled Venetian blinds in private offices: a pilot study." *Journal of Building Performance Simulation* 1(2): 75-89.

Inkarojrit, V. and G. Paliaga (2004). Indoor climatic influences on the operation of windows in a naturally ventilated building. *21th PLEA Conference, Eindhoven, The Netherlands*.

Inoue, T., T. Kawase, T. Ibamoto, S. Takakusa and Y. Matsuo (1988). "The development of an optimal control system for window shading devices based on investigations in office buildings." *ASHRAE Transactions* 94: 1034-1049.

Jia, M., R.S. Srinivasan, A.A. Raheem (2017). "From occupancy to occupant behavior: An analytical survey of data acquisition technologies, modeling methodologies and simulation coupling mechanisms for building energy efficiency." *Renew Sustain Energy Rev* 68: 525–40.

Keyvanfar, A. A. Shafaghat, M.Z. Abd Majid, H. Bin Lamit, M. Warid Hussin, K.N. Binti Ali and A. Dhafer Saad (2014) User satisfaction adaptive behaviors for assessing energy efficient building indoor cooling and lighting environment. *Renew Sustain Energy Rev* 39:277–95.

Khodakarami, J., N. Nasrollahi (2012). "Thermal comfort in hospitals – A literature review." *Renew Sustain Energy Rev* 16: 4071–7

Konstantoglou, M., A. Tsangrassoulis (2016). "Dynamic operation of daylighting and shading systems: A literature review." *Renew Sustain Energy Rev* 60: 268–83.

Langevin, J., Wen, J., Gurian, P.L., (2015). Simulating the human-building interaction: Development and validation of an agent-based model of office occupant behaviors, Building and Environment 88 (2015), 27–45.

Liang, X., T. Hong, G. Shen (2016). Improving the accuracy of energy baseline models for commercial buildings with occupancy data, Applied Energy 179 (2016), 247-260.

Lindelöf, D. and N. Morel (2006). "A field investigation of the intermediate light switching by users." *Energy and Buildings* 38(7): 790-801.

Mahdavi, A. (2011). "People in building performance simulation." *Building performance simulation for design and operation*: 56-83.

Mahdavi, A., A. Mohammadi, E. Kabir and L. Lambeva (2008). "Occupants' operation of lighting and shading systems in office buildings." *Journal of Building Performance Simulation* 1(1): 57-65.

Mahdavi, A. and C. Pröglhöf (2009). Toward empirically-based models of people's presence and actions in buildings. *Proceedings of Building Simulation*.

Mahdavi, A. and C. Pröglhöf (2009). "User behavior and energy performance in buildings." *Wien, Austria: Internationalen Energiewirtschaftstagung an der TU Wien (IEWT)*.

Mahdavi A., (2015). Common fallacies in representation of occupants in building performance simulation, Proceedings of Building Simulation Applications 2015 - 2nd IBPSA-Italy Conference, 1-7, Bozen-Bolzano University Press, ISBN: 978-88-6046-074-5.

Mahdavi, A. and F. Tahmasebi (2015). The inter-individual variance of the defining markers of occupancy patterns in office buildings: a case study. Building Simulation. India, *IBPSA*.

Mahdavi, A. and F. Tahmasebi (2015). "Predicting people's presence in buildings: An empirically based model performance analysis." *Energy and Buildings* 86: 349-355.

Mahdavi, A. and F. Tahmasebi (2016). "The deployment-dependence of occupancy-related models in building performance simulation." *Energy and Buildings* 117: 313-320.

Mahdavi, A., F. Tahmasebi and M. Kayalar (2016). "Prediction of plug loads in office buildings: Simplified and probabilistic methods." *Energy and Buildings* 129: 322-329.

Mahdavi, A., F. Tahmasebi (2016b), On the quality evaluation of behavioral models for building performance applications, Journal of Building Performance Simulation (2016). DOI:10.1080/19401493.2016.1230148.

Martínez-Molina, A., I. Tort-Ausina, S. Cho, J.L. Vivancos (2016). "Energy efficiency and thermal comfort in historic buildings: A review." *Renew Sustain Energy Rev* 61: 70–85.

Masoso, O. T. and L. J. Grobler (2010). "The dark side of occupants’ behaviour on building energy use." *Energy and Buildings* 42(2): 173-177.

Mata, É., A. S. Kalagasidis and F. Johnsson (2014). "Building-stock aggregation through archetype buildings: France, Germany, Spain and the UK." *Building and Environment* 81: 270-282.

McCullagh, P. and J. A. Nelder (1989). Generalized linear models, CRC press.

Menezes, A., A. Cripps, R. A. Buswell, J. Wright and D. Bouchlaghem (2014). "Estimating the energy consumption and power demand of small power equipment in office buildings." *Energy and Buildings* 75: 199-209.

Menezes, A. C., A. Cripps, R. A. Buswell and D. Bouchlaghem (2012). "Benchmarking small power energy consumption in the United Kingdom: A review of data published in CIBSE Guide F." *Building Services Engineering Research and Technology*: 0143624412465092.

Morgan, C. and R. de Dear (2003). "Weather, clothing and thermal adaptation to indoor climate." *Climate Research* 24(3): 267-284.

Newsham, G. R. (1997). "Clothing as a thermal comfort moderator and the effect on energy consumption." *Energy and Buildings* 26(3): 283-291.

Nicol, J. F. (2001). Characterising occupant behaviour in buildings: towards a stochastic model of occupant use of windows, lights, blinds, heaters and fans. *Proceedings of the seventh international IBPSA conference, Rio*.

Nicol, J. F. and M. A. Humphreys (2002). "Adaptive thermal comfort and sustainable thermal standards for buildings." *Energy and Buildings* 34(6): 563-572.

Nicol, J. F. and M. A. Humphreys (2004). "A Stochastic Approach to Thermal Comfort--Occupant Behavior and Energy Use in Buildings." *Ashrae Transactions* 110(2).

Norford, L., R. Socolow, E. Hsieh and G. Spadaro (1994). "Two-to-one discrepancy between measured and predicted performance of a ‘low-energy’office building: insights from a reconciliation based on the DOE-2 model." *Energy and Buildings* 21(2): 121-131.

NRC (2015). National Energy Code of Canada for Buildings, *National Research Council Canada*.

O'Brien, W. and H. B. Gunay (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., Gunay, H.B., Tahmasebi F. & A. Mahdavi (2016). A preliminary study of representing the inter-occupant diversity in occupant modelling, Journal of Building Performance Simulation, DOI: 10.1080/19401493.2016.1261943.

Oleckno, W.A., B. Anderson (2012). Essential epidemiology: principles and applications, Waveland, 2002.

Omer, A.M (2008). "Renewable building energy systems and passive human comfort solutions." *Renew Sustain Energy Rev* 12: 1562–87

Page, J. (2007). Simulating occupant presence and behaviour in buildings, *EPFL*.

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

Parys, W., D. Saelens and H. Hens (2010). Implementing realistic occupant behavior in building energy simulations–the effect on the results of an optimization of office buildings. *Proceedings of the 10th REHVA World Congress Sustainable Energy use in Buildings, Antalya*.

Parys, W., D. Saelens and H. Hens (2011). "Coupling of dynamic building simulation with stochastic modelling of occupant behaviour in offices–a review-based integrated methodology." *Journal of Building Performance Simulation* 4(4): 339-358.

Pigg, S., M. Eilers and J. Reed (1996). "Behavioral aspects of lighting and occupancy sensors in private offices: a case study of a university office building." *ACEEE 1996 Summer Study on Energy Efficiency in Buildings*.

Reinhart, C. F. (2004). "Lightswitch-2002: a model for manual and automated control of electric lighting and blinds." *Solar Energy* 77(1): 15-28.

Rijal, H., P. Tuohy, M. Humphreys, J. Nicol, A. Samuel and J. Clarke (2007). "Using results from field surveys to predict the effect of open windows on thermal comfort and energy use in buildings." *Energy and Buildings* 39(7): 823-836.

Rijal, H., P. Tuohy, F. Nicol, M. Humphreys, A. Samuel and J. Clarke (2008). "Development of an adaptive window-opening algorithm to predict the thermal comfort, energy use and overheating in buildings." *Journal of Building Performance Simulation* 1(1): 17-30.

Roetzel, A., A. Tsangrassoulis A., U. Dietrich, S. Busching (2010). "A review of occupant control on natural ventilation." *Renew Sustain Energy Rev* 14: 1001–13.

Rubinstein, F., G. Ward and R. Verderber (1989). "Improving the performance of photo-electrically controlled lighting systems." *Journal of the Illuminating Engineering Society* 18(1): 70-94.

Sadeghi, S. A., P. Karava, I. Konstantzos and A. Tzempelikos (2016). "Occupant interactions with shading and lighting systems using different control interfaces: A pilot field study." *Building and Environment* 97: 177-195.

Sanati, L. and M. Utzinger (2013). "The effect of window shading design on occupant use of blinds and electric lighting." *Building and Environment* 64: 67-76.

Schiavon, S. and K. H. Lee (2013). "Dynamic predictive clothing insulation models based on outdoor air and indoor operative temperatures." *Building and Environment* 59: 250-260.

Schimschar, S., K. Blok, T. Boermans and A. Hermelink (2011). "Germany's path towards nearly zero-energy buildings—Enabling the greenhouse gas mitigation potential in the building stock." *Energy Policy* 39(6): 3346-3360.

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

Schweiker, M., and Wagner, A. (2016). Exploring potentials and limitations of the adaptive thermal heat balance framework, Proceedings of 9thWindsor Conference, Cumberland Lodge, Windsor, UK, 7-10 April 2016.

Shaikh, P.H., N.B.M. Nor, P. Nallagownden, I. Elamvazuthi, T. Ibrahim (2014). "A review on optimized control systems for building energy and comfort management of smart sustainable buildings." *Renew Sustain Energy Rev* 34: 409–29.

Shipworth, D., G. Huebner, M. Schweiker, B.R.M. Kingma (2016). Diversity in Thermal Sensation: drivers of variance and methodological artefacts, Proceedings of 9thWindsor Conference, Cumberland Lodge, Windsor, UK, 7-10 April 2016.

Sulaiman, H., F. Olsina (2014). "Comfort reliability evaluation of building designs by stochastic hygrothermal simulation." *Renew Sustain Energy Rev* 40: 171–84.

Sutter, Y., D. Dumortier and M. Fontoynont (2006). "The use of shading systems in VDU task offices: A pilot study." *Energy and Buildings* 38(7): 780-789.

Tahmasebi F. & Mahdavi A. (2012). Optimization-based simulation model calibration using sensitivity analysis, 7th Conference of IBPSA-CZ, Brno, Czech Republic.

Tahmasebi, F. and Mahdavi, A. (2015): The sensitivity of building performance simulation results to the choice of occupants' presence models: a case study, Journal of Building Performance Simulation, DOI: 10.1080/19401493.2015.1117528.

Tahmasebi F. & Mahdavi A. (2016). An inquiry into the reliability of window operation models in building performance simulation, Building and Environment 105 (2016), 343–357.

Taleghani, M., M. Tenpierik, S. Kurvers, A. van den Dobbelsteen (2013). "A review into thermal comfort in buildings." *Renew Sustain Energy Rev* 26: 201–15

Turner, C. and M. Frankel (2008). "Energy performance of LEED for new construction buildings." *New Buildings Institute* 4: 1-42.

Veselý, M., W. Zeiler (2014). "Personalized conditioning and its impact on thermal comfort and energy performance – A review." *Renew Sustain Energy Rev* 34: 401–8.

Wang, C., D. Yan and Y. Jiang (2011). "A novel approach for building occupancy simulation." *Building Simulation* 4(2): 149-167.

Wang, D., C. C. Federspiel and F. Rubinstein (2005). "Modeling occupancy in single person offices." *Energy and Buildings* 37(2): 121-126.

Wang, Y., J. Kuckelkorn, F.Y. Zhao, H. Spliethoff, W. Lang (2016). "A state of art of review on interactions between energy performance and indoor environment quality in Passive House buildings." *Renew Sustain Energy Rev*

Warren, P. and L. Parkins (1984). "Window-opening behaviour in office buildings." *Building Services Engineering Research and Technology* 5(3): 89-101.

Yun, G. Y. and K. Steemers (2008). "Time-dependent occupant behaviour models of window control in summer." *Building and Environment* 43(9): 1471-1482.

Zhang, Y. and P. Barrett (2011). "Factors influencing the occupants' window-opening behaviour in a naturally ventilated office building." Building and Environment.

Zhang, Y. and P. Barrett (2012). "Factors influencing occupants' blind-control behaviour in a naturally ventilated office building." Building and Environment 54: 137-147.

Zomorodian, Z.S., M. Tahsildoost, M. Hafezi (2016). "Thermal comfort in educational buildings: A review article." Renew Sustain Energy Rev 59: 895–906