BERKELEY LAB

Lawrence Berkeley National Laboratory

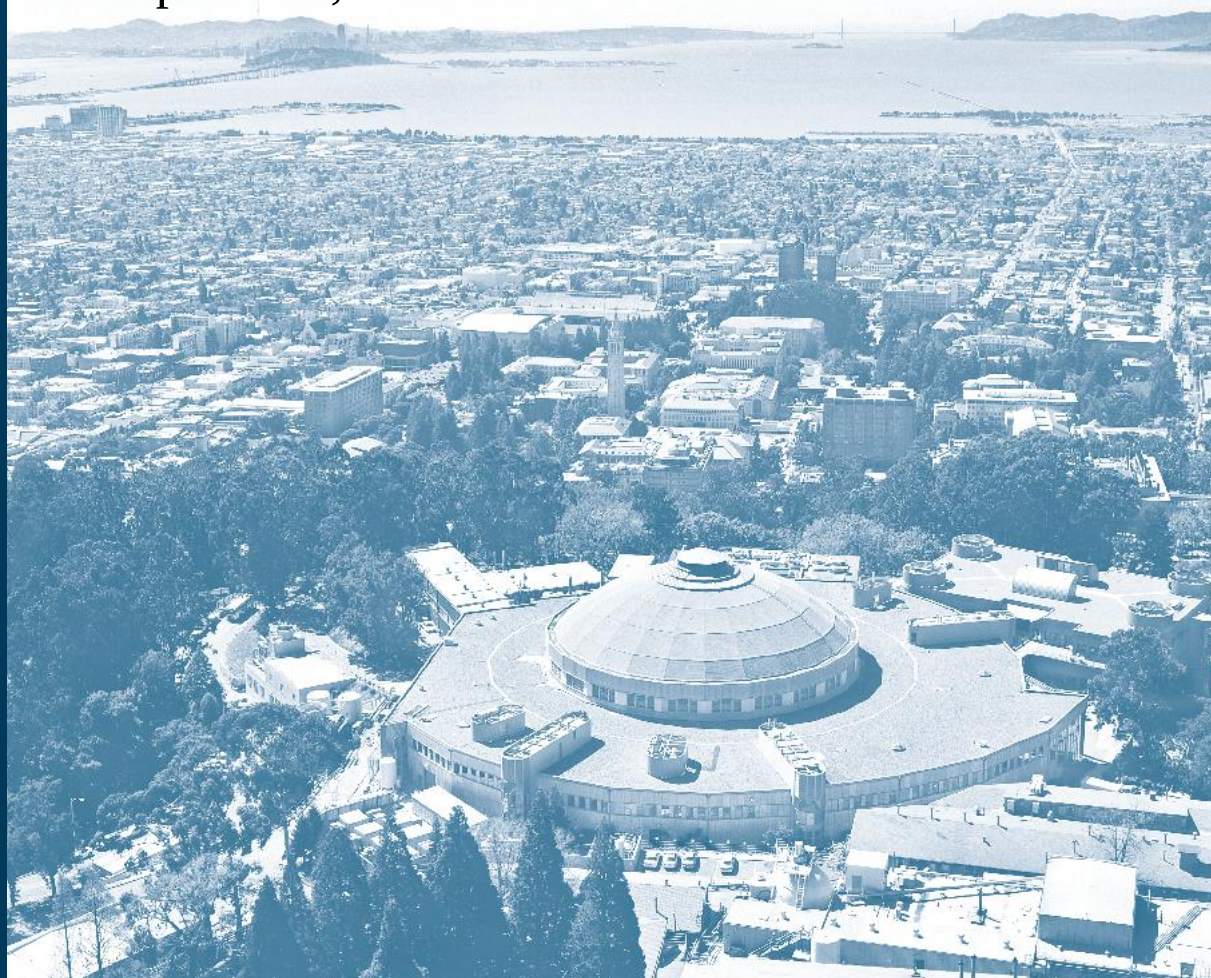
Performance Evaluation of an Agent-based Occupancy Simulation Model

Xuan Luo, Khee Poh Lam, Yixing Chen & Tianzhen Hong

Lawrence Berkeley National Laboratory

Energy Technologies Area

September, 2017



Disclaimer:

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

Performance Evaluation of an Agent-based Occupancy Simulation Model

Xuan Luo[1,2], Khee Poh Lam[1], Yixing Chen[2], Tianzhen Hong[2,*]

[1]School of Architecture, Carnegie Mellon University, Pittsburgh, PA 15213, USA

[2]Building Technology and Urban Systems Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

* Corresponding author (T. Hong). E-mail: thong@lbl.gov, phone: 1(510) 486-7082

Abstract

Occupancy is an important factor driving building performance. Static and homogeneous occupant schedules, commonly used in building performance simulation, contribute to issues such as performance gaps between simulated and measured energy use in buildings. Stochastic occupancy models have been recently developed and applied to better represent spatial and temporal diversity of occupants in buildings. However, there is very limited evaluation of the usability and accuracy of these models. This study used measured occupancy data from a real office building to evaluate the performance of an agent-based occupancy simulation model: the Occupancy Simulator. The occupancy patterns of various occupant types were first derived from the measured occupant schedule data using statistical analysis. Then the performance of the simulation model was evaluated and verified based on (1) whether the distribution of observed occupancy behavior patterns follows the theoretical ones included in the Occupancy Simulator, and (2) whether the simulator can reproduce a variety of occupancy patterns accurately. Results demonstrated the feasibility of applying the Occupancy Simulator to simulate a range of occupancy presence and movement behaviors for regular types of occupants in office buildings, and to generate stochastic occupant schedules at the room and individual occupant levels for building performance simulation. For future work, model validation is recommended, which includes collecting and using detailed interval occupancy data of all spaces in an office building to validate the simulated occupant schedules from the Occupancy Simulator.

Keywords: Occupancy simulation; occupancy pattern; model performance evaluation; verification; occupant presence and movement; occupant behavior

1. Introduction

Occupancy is an important factor driving building performance, and occupants' presence and movement patterns are fundamental to building energy simulation [1,2]. Traditionally, in building performance simulation (BPS), occupancy schedule inputs are static and homogeneous, leading to a lower accuracy in predicting building energy performance [3]. The real occupancy pattern of occupants in buildings may differ significantly from each other [4]. Furthermore, the indeterminacy of occupant behavior in building performance simulation is a leading source of uncertainty in predicting building energy use [5]. To model the influence of human behaviors on building energy consumption and the indoor environment accurately, the occupancy schedule inputs should provide realistic information on the presence and absence status, the number of occupants in rooms, and the diversity of occupant behavior patterns. This issue of realistic representation of occupancy schedules used in building performance simulation has been a recent topic of study and discussion.

Compared to the conventional static schedules, stochastic occupancy simulation models, which better represent the random nature of occupant presence and movement behavior, are recommended for application in BPS programs [6]. The most common way of generating a stochastic occupant schedule is to reproduce an occupancy pattern using selected occupant profiles and then apply statistical models representing the occupant behavior processes [7]. Page, et al. proposed a probabilistic model to predict and simulate occupancy in single-occupancy offices, which generalizes a stochastic model for the occupancy simulation using weekly presence probability statistics and a mobility parameter regarding state change of presence and absence [8]. By considering occupant presence as an inhomogeneous Markov chain, interrupted by occasional periods of long absence, the model generates a time series of the state of presence (absent or present) of each occupant in each room inside a building. Stoppel, et al. also presented a probabilistic occupancy model for occupants' long vacancy activities, which could be further integrated with BPS models [9]. In their simulation results, long vacancy activities such as training, vacation, and other building underutilization are reported in the daily occupancy rates of rooms and buildings.

Incorporating the concept of using occupant profiles and selected properties of occupants to generate corresponding occupancy schedule output, an agent-based building occupancy simulation model based on the homogeneous Markov chain model was introduced to simulate

the stochastic movement of occupants [10]. With detailed building and occupant profile as inputs, the model can be integrated in simulation tools to generate time series location for each occupant and the occupancy of each space in the building [11]. This agent-based algorithm was adopted in a web-based occupancy simulation tool named the Occupancy Simulator, which simulates the stochastic occupant presence and movement in buildings, capturing the spatial and temporal occupancy diversity [12].

However, these simulation models are usually based on theoretical assumptions and mathematical simplification of occupancy-related behaviors. In the real world, occupancy patterns are diverse and complex [2], and it is important that the building energy model can effectively represent the features of various occupants' behavior in real [13]. Therefore, to generate applicable occupant schedule inputs for BPS, the stochastic occupancy model should be evaluated according to realistic occupancy properties of a variety of typical occupant types. Moreover, it is necessary to verify whether the implementation of the agent-based model is suitable for modeling different simulation scenarios, and to verify if the model satisfies a range of accuracy consistent with the intended applications.

Existing studies adopted a large number of evaluation metrics to assess the performance of the occupancy simulation models to determine the accuracy of simulation results compared to observed ground truth data. A study by Tahmasebi & Mahdavi validated the stochastic realization of occupant profiles as a representation of occupants' presence patterns by comparing the occupancy model outputs with the actual occupancy levels at the building level [14]. To conduct a quantitative evaluation, three statistical metrics are considered, namely mean error, root mean squared error (RMSE), and Jensen–Shannon distance [15]. Similarly, in Mahdavi's research regarding occupancy pattern analysis, a set of specific evaluation statistics was deployed for the comparison of the model performance—such as the cumulative distribution of first arrival time errors, departure time errors, and number of transitions errors [16,17].

The Occupancy Simulator, developed by Lawrence Berkeley National Laboratory and made freely available for public at occupancysimulator.lbl.gov, is a web-based application running on multiple platforms and devices to simulate occupant presence and movement in buildings. It generates hourly or sub-hourly occupant schedules for each space and for individual occupants in CSV files and EnergyPlus IDF files for building performance simulation [18]. Object-oriented

design is adopted in the Occupancy Simulator, which represents the objects of building, space, occupant, and their movment behavior in a hierarchical structure (Figure 1). Each object is modeled mathematically with its properties and behaviors. For example, the movement behavior object is modeled as an aggregation of occupancy-related events. Under the structure, the simulator performs an agent-based simulation for each occupant to produce individual occupancy by generating a time series of occupant locations that renders occupancy-related events such as arrivals, departures, and random movement from one space to another inside the building.

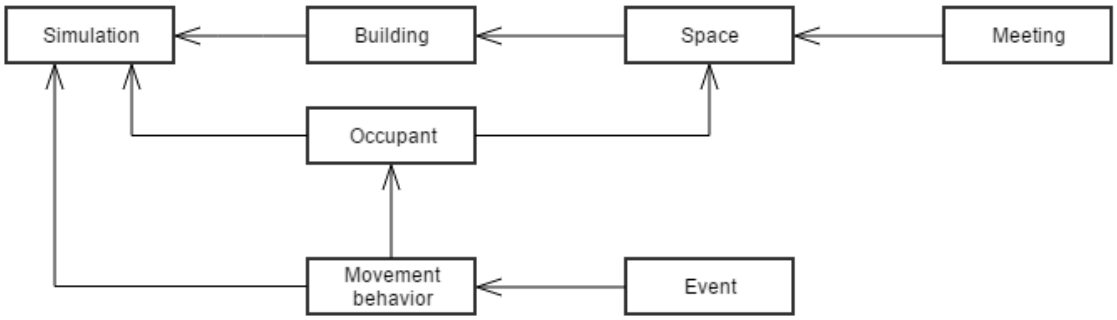


Figure 1 Object structure of the Occupancy Simulator

As the model input, the Occupancy Simulator adopts a set of descriptive occupant profiles to reproduce the occupancy patterns for each type of occupant in buildings using the movement behavior simulation model. The profile is formatted based on obXML, an XML schema that standardizes the representation and exchange of occupant behavior models for building performance simulation [19]. In the profile, each *Space* is assigned a *SpaceType,* which defines the occupancy density, occupancy composition and meeting events in the space. Similarly, each *Occupant* has an *OccupantType*, which is defined by the *MovementBehavior* of the occupants. For occupant movement behaviors simulation, the *MovementBehavior* object includes the detailed description of regular occupancy-related events and their properties, such as arrival and departure events, and the percentage of their presence time in each space category (Figure 2). The profile serves as the model input for the simulation.

The current implementation of the simulation model aims to generate the regular working schedules in office buildings, considering the stochastic nature of the occurrence of occupancy related events. The space- and building-level occupancy schedules are generated by simulating each occupant separately as an agent and then aggregating the produced patterns of presence.

Users can choose an individual space or the whole building to visualize or export the simulated occupancy results. Currently, the model simulates regular work days, while handling public holidays at the building level. It does not yet consider individual long-term absences (e.g. sick leaves, vacations, work from home, business trips, etc.).

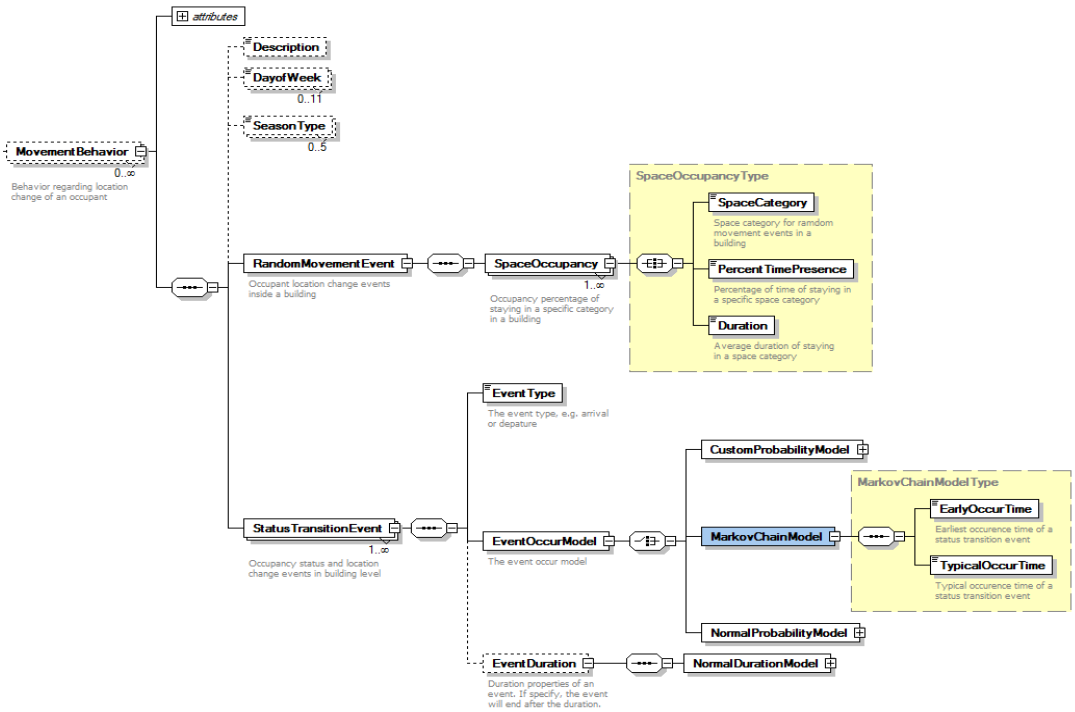


Figure 2 Mathematical models for movement behavior simulation in the Occupancy Simulator

This study aims to evaluate and verify the algorithms incorporated in the Occupancy Simulator for occupant presence and movement simulation. As illustrated in Figure 2, the Occupancy Simulator represents the movement behavior with two kinds of events: the status transition events and the random movement events. The status transition events contain events determining the binary occupancy status in the building, such as occupants' arrival and departure. The random movement events describe occupant location transition in the building for a certain period of time, such as going to the restroom or going to other people's offices. In the simulator, the status transition events are generated by the ''inverse function method'' (IFM), which generates a sample of events in a time series from a given probability distribution function (PDF), as suggested by Reinhart and Page, et al. [8,20]. Between these two types of events, random movement in the building is simulated using the first-order homogeneous Markov chain technique, which simplifies the movement behavior to a Markov process,

assuming the status at the current time step only depends on the state of presence at the previous time step [10].

2. Methodology

The key to evaluating the performance of the Occupancy Simulator is to evaluate the feasibility of using the tool to reproduce the occupancy-related behavioral patterns. Specifically, the main objective of the model evaluation is to test the hypotheses incorporated in the simulation model including: (1) the occurrences of the observed occupancy-related events follow a certain distribution that can be modeled in a statistical way, and (2) for a specific type of occupants, the occupancy pattern can be reproduced based on selected descriptive properties.

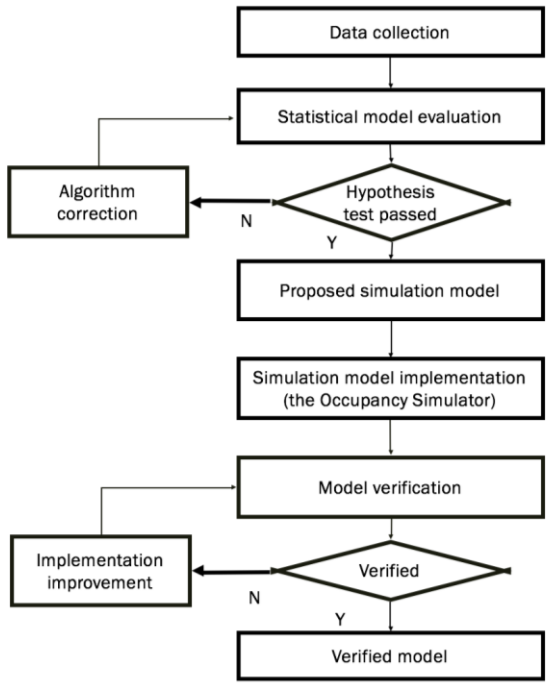


Figure 3 Procedure of model performance evaluation and verification

Figure 3 illustrates the procedure to conduct the model performance evaluation and verification, step by step. To understand the various occupancy patterns of different types of occupants, real occupancy data is collected and categorized based on occupancy type. With the patterns obtained, the theoretical model is evaluated based on whether the model is able to reproduce the statistical features of the observed occupancy patterns. The evaluation is performed by testing the fitness of the theoretical and observed distributions of the selected properties. If passing the hypothesis test, the proposed simulation model is implemented in the Occupancy

Simulator. Then the model implementation is verified to check whether the selected properties of the occupant profiles are successfully reproduced in the Simulator.

2.1 Data collection

To obtain real occupancy data for occupancy pattern analysis in regular office buildings, the office area of the Hillman and Gates' Center building (6th to 9th floor) at Carnegie Mellon University was observed. Categorized by occupation, a total number of 25 researchers, 25 professors, and 16 administrators were monitored by occupancy sensors in each room in the whole studied building area. Generally, the data were collected from private offices, so the occupancy sensor records can be used to determine the individual absence and presence schedule in an office.

The field experiment lasted three months from Oct 1st to Dec 31st 2015. Binary occupancy records in 10-minute intervals were obtained, through the building management system, using the wireless ceiling-mounted passive infrared (PIR) occupancy sensor installed in each monitored office room. Figure 4 shows a sample area plan and location of monitored rooms. Considering that the simulation model aims to generalize the regular occupancy patterns with statistical models, it is suggested that less representative events, such as long-term absence or vacation, be eliminated in the analysis process. Specifically, in this case, days without a regular working pattern were eliminated from the analysis, including weekends, holidays and days with special events. During data processing, days were not considered in the analysis when the daily room occupancy rate is less than 10% (when the working hour is less than 2.4 hours). On average for all observed occupants, 67% of the observed days are considered as the regular working days, and are used in regular occupancy pattern analysis.

Figure 4 Plan view of a sampled area of observed offices in the GHC building

2.2 Metrics for evaluation

Properties describing the occupancy patterns of different occupant types serve as model performance indicators during the model evaluation and verification process. For the studied agent-based occupancy simulation model, occupancy-related events are reproduced based on the probability distribution of the event occurrence time and duration, while random movement between the events is simulated with the homogenous Markov chain model according to space occupancy and one-time stay duration. The model is evaluated in two procedures: statistical models evaluation and model implementation verification.

Statistical models evaluation

The model is firstly evaluated by if the statistical distributions of the four occupancy traits fit the theoretical distributions implemented in the Occupancy Simulator, as listed in Table 1. For occupant behavior model evaluation, previous studies adopted evaluation metrics such as the distribution of prediction errors for observed behavioral properties, distance of observation and predictions, and relative deviation of predicted actions from observed behavior [14-17]. In this study, the feasibility of adopting the theoretical model for simulation is evaluated by the fitness of the observed distributions of the selected properties and their mathematical simplification.

Table 1 Event pattern properties for statistical models evaluation

Observed behavior	Occupancy pattern property	Evaluation metrics
Arrival	First arrival time (FA)	Fitness of distribution
Random movement in the building	Average one-time stay duration (SD)	Fitness of distribution
	Average one-time leaving duration (TD)	Fitness of distribution
Departure	Average last departure time (LD)	Fitness of distribution

Specifically, as labeled in Figure 5, the first arrival time (FA) is defined as the first time in a day an occupant arrives at his or her office, and the last departure time (LD) is defined as the last time in a day an occupant leaves the office. One-time stay duration (SD) represents the period of staying at the office without leaving it, while one-time leaving duration (TD) represents the period of absence before coming back to the office again.

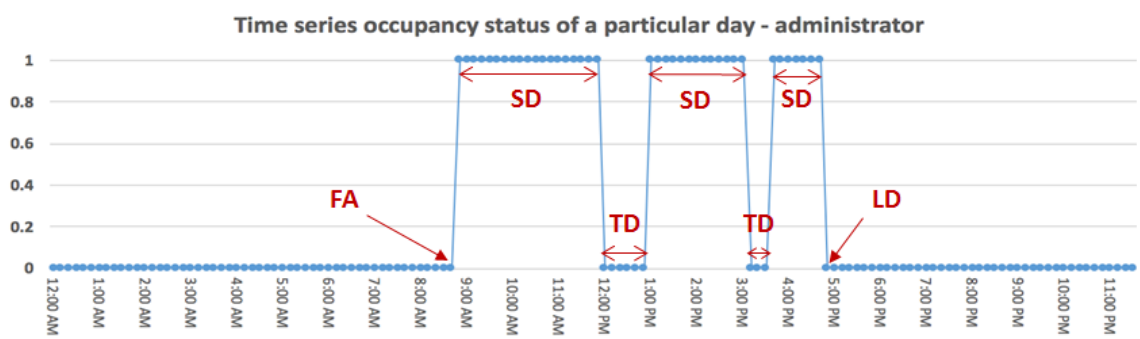


Figure 5 Descriptive properties for occupancy related events in a typical office day

For model evaluation, taking FA as an example, the procedure of the hypothesis test is to compare the observed distribution of the arrival time with several mathematical distributions adopted in the agent-based simulation model, such as the normal distribution, Geometric distribution. If a pattern exists and matches a certain distribution, the simulation model can be used to reproduce the corresponding behavioral pattern of a certain group of occupants. Otherwise, we reject the hypothesis. Also, to draw representative occupancy properties from the raw occupancy data, outliers are excluded in the behavior feature analysis. Specifically, for the four observed behaviors listed in Table 1, their statistical metrics are calculated and any data point more than 1.5 interquartile ranges (IQRs) below the first quartile or above the third quartile are eliminated.

The Kolmogorov-Smirnov (KS) test is adopted to test the fitness of distributions of the assumed and observed patterns. The KS test is a non-parametric statistical test technique commonly used

to compare two distributions and assess their fit [21]. It is particularly indicated when one of the distributions is the theoretical distribution that the observations are estimated to follow and the other is the actual empirical distribution computed.

Model implementation verification

To verify the overall performance and accuracy of the implementation of the simulation model, the model is further evaluated by whether the Occupancy simulation generation can maintain the original occupancy patterns. The six evaluation metrics used in this study are listed in Table 2. For each measured feature, the verification criterion is whether the simulation results successfully reproduce the occupancy pattern based on model inputs evaluated by the corresponding metrics.

Table 2 Event pattern properties for model implementation verification

Simulated behavior	**Occupancy pattern properties (Model inputs)**	**Evaluation metrics**
Arrival	First arrival (FA)	Fitness of distribution
Random movement in the building	Average one-time stay duration	Error rate
	Average one-time absence duration	Error rate
	Percentage of time stay in own office	Error rate
	Percentage of time of transient absence	Error rate
Departure	Average last departure time (LD)	Fitness of distribution

3. Results

3.1 Statistical model performance evaluation

Properties of occupant event patterns used for the statistical model evaluation are processed and analyzed. As shown in Figure 6, the observed FA and LD for the three occupant types fall into different probability distributions with multiple centers and ranges, while the observed SD and TD fall into similar probability distributions with minor differences. In particular, since the raw occupancy data are recorded in a 10-minute interval, arrival and departure times are rounded to the nearest interval, and the transient absences of less than 10 minutes are neglected.

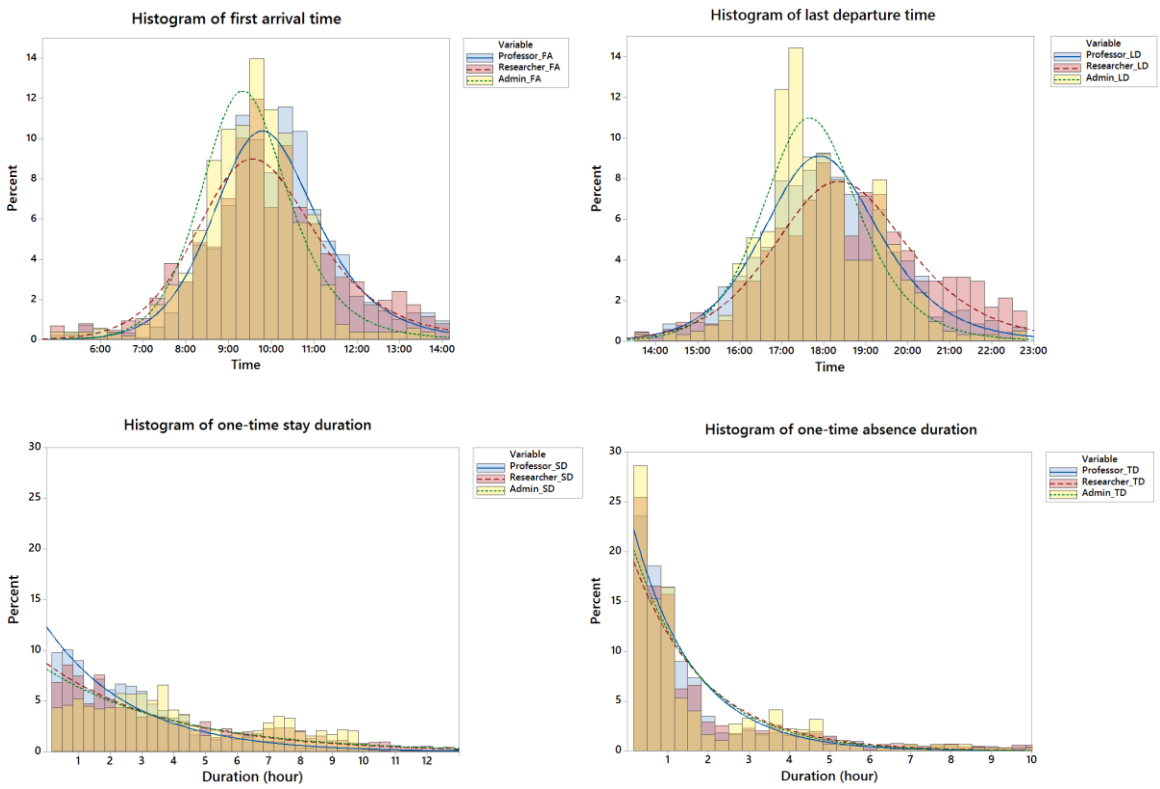


Figure 6 Probability distributions of the observed event pattern properties (The brown area is the overlapped area of the three types of occupants)

To further demonstrate this, one sample KS hypothesis test at a rejection level of 1% is conducted, and cumulative distribution function (CDF) plots are used to further analyze whether the actual pattern matches with or deviates from the theoretical distribution.

Figure 7 shows the CDF plots of the FA and LD of three groups of occupants, professors, researchers, and administrators, with the comparison of the observed distribution of the collected data and the normal distribution. For the professors and researchers, the first arrival time primarily falls into a normal distribution with a center from 9:00 am to 10:00 am and a range from 6:00 am to 14:00 pm. When tested with the one-sample KS test with the normal distribution, it fails to reject the null-hypothesis at the level of 1%, indicating that the empirical distribution matches with the proposed normal one. This confirms that the FA and LD for these two groups of occupants can be reproduced with a normal distribution. However, the CDFs of the FA and LD of administrators both deviate from the normal distribution, and they reject the null-hypothesis. Considering this, for reproducing the arrival and departure event of various groups of occupants, various statistical distributions may need to be incorporated in the

simulation model. Particularly for this case, results of the study suggest that the customized probability model be applied to simulate the occupancy pattern of the administrator group.

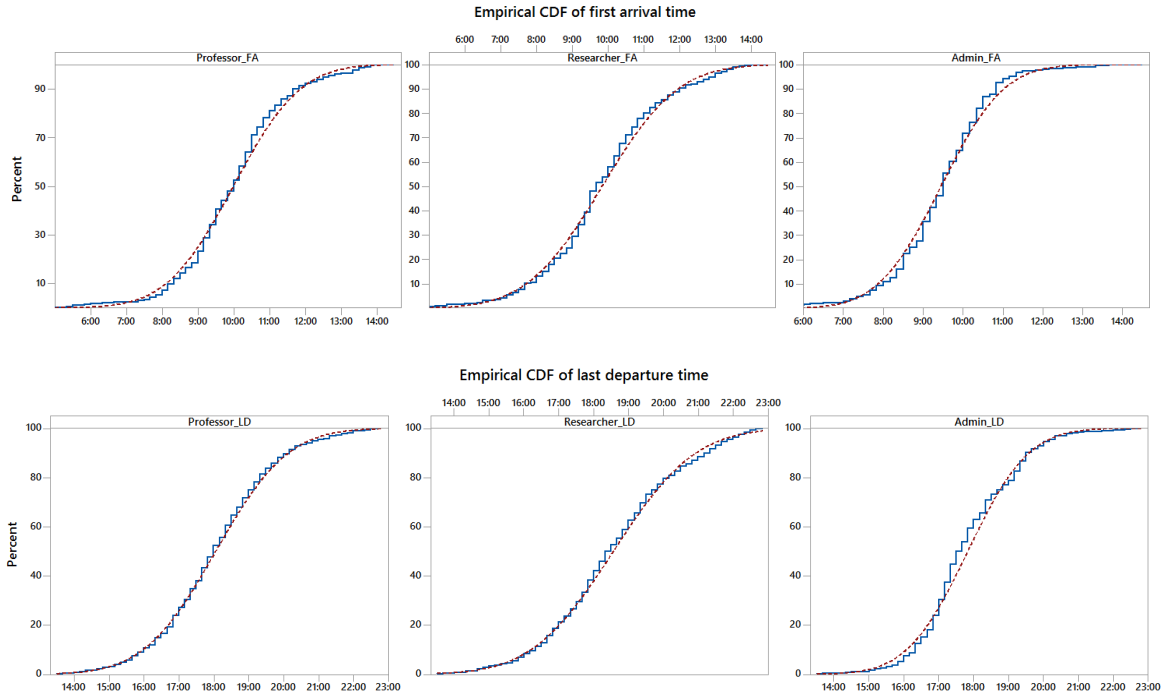


Figure 7 Cumulative distribution functions of the observed FA and LD times

Further, in the proposed homogeneous Markov chain model, it is assumed that the duration of the one-time stay and the one-time absence in a space is geometrically distributed. In this case, the collected data allows the parameter of the average one-time stay and absence from the office room to be determined and the fitness of their distributions with the proposed distribution in the simulation model to be tested. The hypothesis tests fail to reject the null-hypothesis at the level of 1% for the SD for professors and researchers, indicating that the distribution matches the geometric distribution. However, tests for the TD distribution and SD distribution for the administrator group are rejected.

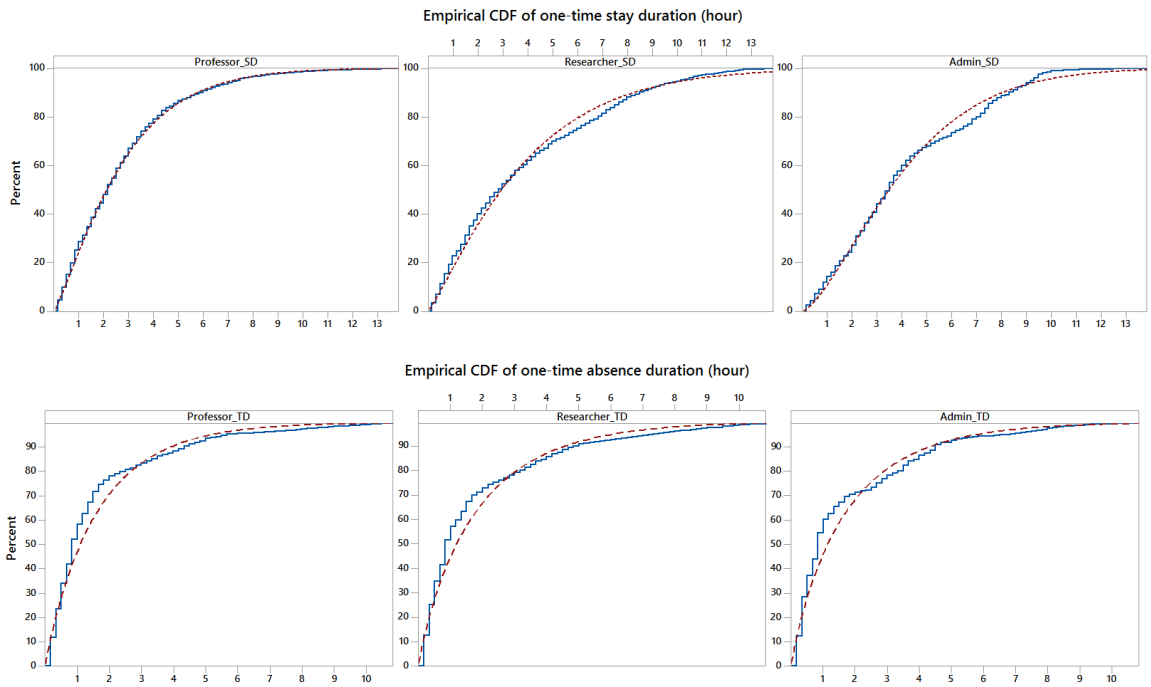


Figure 8 Cumulative distribution functions of the observed SD and TD durations

As illustrated in Figure 8, the results of the one-time stay duration primarily fall into the expected theoretical distribution, while the distribution curves of the one-time absence duration tend to deviate from the proposed distribution due to a higher probability of short absence when less than 60 minutes is observed. However, in the simulation model, the input of the one-time stay duration in each space would mutually affect the results. The one-time stay duration regulates the percentage of time occupants stays in the domain space, such as occupants' own offices. In the simulation, the occupancy rate of the domain spaces is of more concern, performing as the domain constraints in the simulation model. Since the homogeneous Markov chain sucessfully reproduces the observed behaviors in the domain spaces, it is a feasible and appropriate model to be used for the random movement events.

In general, occupancy pattern analysis results show that the occupancy simulation model should be built to be adaptable to various cases of different types of occupants in office buildings. For the status transition events such as arrival and departure, a variety of probabilistic models should be incorporated and applied according to different simulation cases. For the random movement behavior of occupants after arrival and before departure, the homogeneous Markov chain model is acceptable for modeling the stochastic nature of occupant movement.

3.2 Simulation model verification

The occupancy simulation tool is also verified to ensure its appropriate applications in various simulation cases. A case study was conducted to demonstrate the simulation performance of the Occupancy Simulator. A small one-story office building with 44 m (144.4 ft) (L) × 20 m (65.6 ft) (W) × 3.5 m (114.8 ft) (H) was selected for the case study. The building has a total of 13 occupants, among them eight are researchers, two are professors, and three are administrators. The occupant profiles were constructed based on observed data for the three occupant types. The simulation was run for a year; overall occupancy patterns on workdays for each occupancy group were derived from the results based on the above-mentioned metrics.

For the status transition events, three calculation methods—namely the normal probability model, the geometric probability model, and the customized probability model—were tested for modeling the occurrence of an arrival event. As shown in Figure 9, the results verify the feasibility of using a variety of probability models for event occurrence calculation. This allows the users of the simulator to capture behavior pattern features of the specific groups of occupants in a more flexible way.

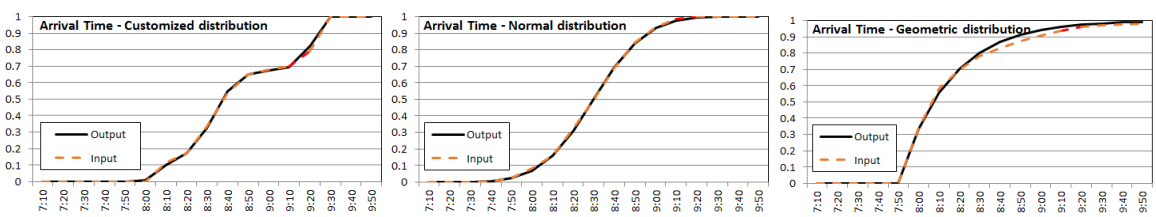


Figure 9 Verification of event occurrence time distribution

In the simulation model, random movement events are calculated with the constraints defined by the average one-time stay duration and the average transient leaving duration. In order to verify whether the simulator can correctly reproduce the input features, the model is tested with a range of inputs for each constraint feature. Table 3 summarizes the input duration and the yearly average of output simulation results, indicating that the statistics match with the input when the input ranges from 60 minutes to 150 minutes. The output deviates from the expected values when the input duration is too short or too long, resulting from two main reasons. On the one hand, the simulation is discrete and the time step in the test cases is 10 minutes, which means that the minimum sojourn time at a location is at least 10 minutes. On the other, the behavior of staying in one space would be interrupted by other events such as

lunch and meetings, so it is very unlikely that the one-time stay duration can be longer than three hours (180 minutes).

Table 3 Verification of the one-time stay duration simulation results

Case	**a**	**b**	**c**	**d**	**e**	**f**
Input duration (minutes)	30	60	90	120	150	180
Output the average duration (minutes)	24	56	87	105	124	146
Error rate	20.0%	6.7%	3.3%	12.5%	17.3%	18.9%

Similarly, a range of one-time leaving durations from 20 minutes to 60 minutes are also applied to test the feasibility of the model to simulate a variety of cases. The test results are listed in Table 4, and the cases **a**-**f** represent the different simulation scenarios with different one-time stay duration as the model input. The results also indicate that the simulator is suitable to model regular cases when the average one-time leaving duration is less than 45 minutes.

Table 4 Verification of the one-time leaving duration simulation results

Case	**a**	**b**	**c**	**d**
Input duration (minutes)	20	30	45	60
Output the average duration (minutes)	17	24	39	46
Error rate	15%	20%	13.3%	23.3%

For regular types of occupants in an office building, observed occupancy data reveals that the typical average one-time stay duration in one's own office is around 80 – 120 minutes, and the one-time absence around 30 minutes (Table 5). The cases **a**-**d** represent the different simulation scenarios with different one-time leaving duration as the model input, and all cases lie in the acceptable ranges of the simulation input as tested.

Table 5 Statistical features of the SD and LD durations

Variable	**One-time stay duration (minutes)**			**One-time leaving duration (minutes)**		
	Q1	Median	Q3	Q1	Median	Q3
Professor	30	78	132	18	30	66
Researcher	48	102	156	12	30	84
Administrator	60	126	216	12	30	102

Occupancy percentage in an occupant's domain space serves as the constraint condition for the random movement behavior simulation. For verification, we also examined the yearly average of occupancy percentage that each occupant stays in his or her own office. Although this metric may be affected by other daily events and may vary from day to day, in a statistical view, the

output shows a high accuracy. Therefore, model implementation is acceptable for simulating regular cases of the random movement events in office buildings.

Table 6 Verification of the occupancy percentage time in own offices

Occupant type	Input	Output						Error rate
		Mon	Tue	Wed	Thu	Fri	Workdays average	
Professor	74.0%	74.0%	70.4%	69.8%	71.0%	76.4%	72.6%	1.9%
Researcher	79.3%	78.0%	78.3%	75.3%	74.2%	74.0%	76.5%	3.5%
Administrator	83.2%	76.0%	77.8%	78.8%	79.7%	76.0%	78.6%	5.5%

4. Discussion

This study revealed, for different types of occupants with different occupation and behavior preference, that occupancy behavior patterns can vary one from another. Due to limited available measured occupancy data, the study focuses on performance analysis as well as verification of modeling regular occupancy behavior patterns in office buildings. If more data are obtained in the future, the performance evaluation of the simulation could be expanded following the similar procedure suggested in this study. To ensure that the model be more generalized to reproduce a range of occupancy patterns properly for a larger scope of occupant types and building types, a variety of mathematical models should be considered and integrated into the simulator to simulate the occupancy presence and movement behaviors in office buildings. Moreover, the study focuses on the evaluation and verification of the presence and movement behavior model at the occupant level. With sufficient data collected and analyzed, it is also suggested that group events at the room level, such as meetings, be analyzed in a similar way. Further, the current model implementation does not consider individual long-term absences (e.g. sick leaves, vacations, work from home, business trips, etc.), since the behavior of an occupant taking a long-term absence is not fully observed in the study. A new category of personal absent workdays can be added to the occupant profile in future work.

As the probabilistic and stochastic simulations are based on random number generators, which depend on the seeds that can vary between simulation runs, the simulated occupant schedules will be different in each simulation run. It is suggested that reasonable repetition be taken into consideration in simulation for analyzing occupancy behavior, especially when analyzing the impact of occupant behavior on building energy consumption [22]. Therefore, the generality of

the evaluation and verification demonstrated in this study remains to be further tested with long-term observation of occupancy patterns.

It should be noted that occupancy models derived from data mining and machine learning can be incorporated in the occupancy simulation model with similar approaches as suggested in this study. Considering the variety and indeterminacy of occupant behaviors in buildings, using measured occupancy behaviors and activities is recommended to derive the occupancy models with data mining and machine learning methods, rather than purely by theoretical assumptions and mathematical simplifications. Extrapolation of patterns from big data streams is a powerful analysis technique for this kind of study [23] . Further investigations are suggested to deploy more robust models to reproduce patterns of human movement and actions into behavioral models and to uncover the impact of typical occupancy patterns on the energy performance of buildings.

5. Conclusions

This study recommends that for status transition events such as occupant arrival and departure, a variety of probabilistic models should be incorporated and applied according to different simulation cases. To simulate various types of occupants in different types of office buildings, a variety set of probability distribution functions should be adopted in the model. Currently three distribution functions, the geometric, the normal and the user-custom, were implemenetd in the Occupancy simulation model, which can be chosen by users depending on their applications. The study also recommends that the homogeneous Markov chain model is acceptable for modeling the stochastic nature of occupant movement after daily arrival and before daily departure. Result of this verification of the statistics of the simulation occupancy results against the simulation inputs show that the Occupancy Simulator is capable of simulating a range of occupancy patterns and behavioral preference in office buildings.

This study shows that the Occupancy Simulator does accurately simulate occupant presence and movement in office buildings and that the generated occupant schedules accurately represent the temporal and spatial diversity of real occupancy.

Future work is recommended that uses detailed measured interval occupancy data from all spaces of an office building to validate the simulated occupant schedules from the Occupancy

Simulator, and to further identify potential limitations and enhancements that need to be addressed for Occupancy Simulator's practical applications.

Acknowledgment

This work is sponsored by the United States Department of Energy (Contract No. DE-AC02-05CH11231) under the U.S.-China Clean Energy Research Center for Building Energy Efficiency. The work is also part of the research activities of the International Energy Agency Energy in Buildings and Communities Program Annex 66, Definition and Simulation of Occupant Behavior in Buildings.

References

[1] A. Al-mumin, O. Khattab, and G. Sridhar, "Occupants ' behavior and activity patterns influencing the energy consumption in the Kuwaiti residences," vol. 35, pp. 549–559, 2003.

[2] X. Liang, T. Hong, and G. Qiping, "Occupancy data analytics and prediction : A case study," *Build. Environ.*, vol. 102, pp. 179–192, 2016.

[3] M. Bonte, F. Thellier, and B. Lartigue, "Impact of occupant's actions on energy building performance and thermal sensation," *Energy Build.*, vol. 76, pp. 219–227, 2014.

[4] D. Yan, W. O'Brien, T. Hong, X. Feng, H. Burak Gunay, F. Tahmasebi, and A. Mahdavi, "Occupant behavior modeling for building performance simulation: Current state and future challenges," *Energy Build.*, vol. 107, pp. 264–278, 2015.

[5] P. Hoes, J. L. M. Hensen, M. G. L. C. Loomans, B. de Vries, and D. Bourgeois, "User behavior in whole building simulation," *Energy Build.*, vol. 41, no. 3, pp. 295–302, 2009.

[6] W. Parys, D. Saelens, and H. Hens, "Coupling of dynamic building simulation with stochastic modelling of occupant behaviour in offices – a review-based integrated methodology," *J. Build. Perform. Simul.*, vol. 4, no. August 2012, pp. 339–358, 2011.

[7] J. Virote and R. Neves-Silva, "Stochastic models for building energy prediction based on occupant behavior assessment," *Energy Build.*, vol. 53, pp. 183–193, 2012.

[8] J. Page, D. Robinson, N. Morel, and J. L. Scartezzini, "A generalised stochastic model for the simulation of occupant presence," *Energy Build.*, vol. 40, pp. 83–98, 2008.

[9] C. M. Stoppel and F. Leite, "Integrating probabilistic methods for describing occupant presence with building energy simulation models," *Energy Build.*, vol. 68, pp. 99–107, 2014.

[10] C. Wang, D. Yan, and Y. Jiang, "A novel approach for building occupancy simulation," *Build. Simul.*, vol. 4, pp. 149–167, 2011.

[11] X. Feng, D. Yan, and T. Hong, "Simulation of occupancy in buildings," *Energy Build.*, vol.

87, pp. 348–359, 2015.

[12] T. Hong, H. Sun, Y. Chen, S. C. Taylor-Lange, and D. Yan, “An occupant behavior modeling tool for co-simulation,” *Energy Build.*, vol. 117, pp. 272–281, 2016.

[13] C. Duarte, R. Budwig, and K. Van Den Wymelenberg, “Energy and demand implication of using recommended practice occupancy diversity factors compared to real occupancy data in whole building energy simulation,” *J. Build. Perform. Simul.*, vol. 1493, no. January 2016, pp. 1–16, 2014.

[14] F. Tahmasebi and A. Mahdavi, “The sensitivity of building performance simulation results to the choice of occupants’ presence models: a case study,” *J. Build. Perform. Simul.*, vol. 1493, no. January, pp. 1–11, 2015.

[15] A. Mahdavi and F. Tahmasebi, “Predicting people’s presence in buildings: An empirically based model performance analysis,” *Energy Build.*, vol. 86, pp. 349–355, 2015.

[16] F. Tahmasebi and A. Mahdavi, “Advances in Building Energy Research Stochastic models of occupants ’ presence in the context building systems control,” vol. 2549, no. September, 2016.

[17] A. Mahdavi and F. Tahmasebi, “On the quality evaluation of behavioural models for building performance applications,” *J. Build. Perform. Simul.*, vol. 0, no. 0, pp. 1–11, 2016.

[18] Y. Chen, X. Luo, and T. Hong, “An Agent-Based Occupancy Simulator for Building Performance Simulation,” *ASHRAE Annu. Conf.*, no. April, 2016.

[19] T. Hong, S. D’Oca, S. C. Taylor-Lange, W. J. N. Turner, Y. Chen, and S. P. Corgnati, “An ontology to represent energy-related occupant behavior in buildings. Part II: Implementation of the DNAS framework using an XML schema,” *Build. Environ.*, vol. 94, no. P1, pp. 196–205, 2015.

[20] C. F. Reinhart, “Lightswitch-2002: A model for manual and automated control of electric lighting and blinds,” *Sol. Energy*, vol. 77, no. 1, pp. 15–28, 2004.

[21] F. J. M. Jr, “The Kolmogorov-Smirnov Test for Goodness of Fit,” vol. 1459, no. October, 2016.

[22] X. Feng, D. Yan, and C. Wang, “On the simulation repetition and temporal discretization of stochastic occupant behaviour models in building performance simulation,” *J. Build. Perform. Simul.*, vol. 0, no. 0, pp. 1–13, 2016.

[23] S. D. Oca and T. Hong, “Occupancy schedules learning process through a data mining framework,” no. February, 2015.