PROBABILISTIC WINDOW OPENING MODEL CONSIDERING OCCUPANT BEHAVIOR DIVERSITY: A DATA-DRIVEN CASE STUDY OF CANADIAN RESIDENTIAL BUILDINGS

Jean Rouleau, Louis Gosselin*

Department of Mechanical Engineering, Université Laval, Quebec City, QC, Canada

Article accepté pour publication dans : Energy, Volume 195, 15 mars 2020

Abstract

It was found from monitored data from eight dwellings in a case study building in Quebec City (Canada) that there are clear differences in the window opening behavior between different households. This paper aims to develop from data a probabilistic window opening model that accounts for occupant behavior. Logit regression is employed to predict the state (opened/closed) of windows according to indoor and outdoor temperatures environmental and temporal parameters. To replicate the diversity of behavior, normal distribution functions applied to the logit regression coefficients are used so that simulated occupants respond differently to environmental stimuli. It was found that the model offers good prediction for the monitoring by only using the outdoor and indoor temperatures as predictors. The proposed methodology was tested by simulating 10,000 times a full validation year of the case study building and comparing the results with measured data. The agreement was good. The model overestimated slightly the total frequency of window opening in the dwellings and the number of window changes-of-state. A vast range of window opening behavior was generated by the model, showing its ability to reproduce both the aggregated window opening behavior and the diversity of behaviors of the case study building.

Keywords: Window opening; Occupant behavior; Residential buildings; Occupant diversity; Building energy simulation

Nomenclature

p Probability

* Corresponding author: Louis.Gosselin@gmc.ulaval.ca; Tel.:+1-418-656-7829.

T	Temperature [°C]
x	Independent variable

Greek letters

ω	Logit regression coefficient

Subscripts

clo	Closure of windows
const	Constant
in	Inside temperature
op	Opening of windows
open	Opened window
out	Outside temperature

1. Introduction

Recent studies in social housing buildings found that many occupants experience thermal and olfactory discomfort, and even report health problems related to these issues [1], [2]. Window openings have a high impact on the energy consumed to sustain the desired level of indoor air quality [3], [4]. Window states are one of the required pieces of information for modeling natural ventilation in commercial and residential buildings and its impact on the energy performance [5]. However, the control of windows is a product of a complex combination of physical drivers and behaviors of occupants [6]. This is especially true in residential buildings where occupants are entirely free to use their windows as they wish. During the design phase of a building, it is consequently cumbersome to forecast when windows will be opened and when they will be closed, which can result in the building not performing as expected. For example, if a building design relies on natural ventilation to prevent overheating in summer, there could be a problem of high indoor temperatures if occupants open their windows less frequently than expected. On the other hand, if the windows openings are more frequent than expected during the heating season, the building consumption of heat could dramatically increase.

In situ observations have revealed that window opening patterns can vary significantly among occupants, making it difficult to assess the potential range of impact of natural ventilation on energy performance and comfort. Therefore, efforts have been devoted over the years to developing window opening models that employ probabilistic or machine learning approaches [7]–[11]. Probabilistic models generally calculate the probability of an event's occurrence through empirical correlations created from measurements or aggregated statistical data. Since these models are based on probabilities instead of a purely static behavior (e.g., fixed schedules), they allow the representation of various occupant behaviors and in a simulation, every day of the year can have a different window opening schedule.

One of the most popular modeling approach to create probabilistic window opening models is the "logit" equation. Calì et al. analyzed window behavior in residential buildings, and investigated drivers that lead occupants to interact with windows, while focusing on how these actions can be modeled [12]. Jones et al. conducted a field study in ten UK dwellings over a year, and used multivariate logistic regression to investigate the probability of window opening and closing in the main bedroom, based on indoor and outdoor environmental factors, considering the time of the day and the season as well [13]. Meanwhile, Yao and Zhao conducted an investigation on the factors influencing occupant window behavior in 19 residences in Beijing, based on the monitored state of windows and eight environmental parameters [14]. In their study, multivariate linear logistic regression was also used to establish predictive models for occupant window behavior, and the results indicated that outdoor temperature was the most influential factor. Stazi et al. investigated the relationship between window use and environmental stimuli in an Italian classroom, and developed a window behavior model using logistic regression [15]. In their study, it was found that indoor temperature was the best predictor for both opening and closing windows and that outdoor temperature also had a significant impact on the window states, but not as strong as the indoor temperature.

In spite of the large number of window opening models that have been developed in the last decades, most models do not consider the diversity in occupant behavior, assuming that everyone follows the behavior (from a statistical point of view) of the "average occupant". In reality, all people do not necessarily interact in the same way with the built environment systems and do not necessarily follow the same behavioral patterns. It has been shown in other studies that in low-energy buildings, the practice of using an average behavior can be a major cause of the gap between the actual and predicted energy use of the building [16]–[19]. Not considering the diversity in occupant behavior also prevents the prediction of extreme values of energy demand or the assessment of the robustness of a given building design in front of particular behavioral patterns. Previous studies created behavioral classifications [20]–[24], typically consisting of three classes of behaviors: "low-energy" behavior, "medium" behavior and "high-energy" behavior. These classifications are useful to have a glimpse at the range of potential energy use of the building, but they do not provide the full picture of observed behaviors.

A noticeable recently proposed model that includes window opening diversity is that of Haldi et al. [25] for apartments located in Germany and Denmark. Logit models with probability functions for the weights of the regression were used successfully to represent occupant behavior diversity. The model employs temperatures and CO_2 data as predictors. However, the latter is an information that is not necessarily available in all buildings (e.g., the case study building of this paper). This could make the model difficult to apply to some buildings. Furthermore, only one set of European data was used to test this approach, in such a way that it is still unknown whether the method could be extended to other countries where climates and behaviors are different.

Based on literature, it is also found that most window opening models so far were developed using data obtained in Europe or East Asia. For instance, in their 2019 literature review on window behavior [8], Pan et al. provided information regarding stochastic window opening models. Including their own model, a total of 18 window opening models are reported [6], [8], [9], [13]–[15], [23], [24], [26]–[34]. Out of these 18 models, 4 were created with observations made from buildings located in UK, 4 from buildings located in China and the remaining 10 from buildings located in other parts of Europa and East Asia. No model based on North American data was reported, suggesting an underrepresentation in terms of available window opening behavior data and models for this region.

The first objective of this study is to develop a window opening model for residential buildings that considers the diversity of occupant behaviors. Considering diversity in occupant behavior means that when simulating the use of windows in multiple households with the model, results should be different between the simulated households, but still be consistent with what is observed in real buildings. This means that there should be some households that are low “window users”, some that are heavy users and the rest between these two extremes. The model should also include difference regarding how a environmental stimulus influences occupant behavior. For instance, the window related behavior of some people could be heavily influenced by the indoor temperature, whereas for others, it could be mainly driven by the indoor humidity or the time of the day. The purpose of this model is to reflect these differences in behavior for energy simulations and not to explain why these differences exist. The model is meant to be part of a unified occupant behavior model that simulate multiple aspects of occupant behavior (occupancy, set point temperatures, windows…) and that is to be coupled with energy simulation tools, so the window model should also be coherent with other parts of the unified occupant behavior tool. This means that if the “occupancy” aspect of the overall occupant behavior tool predicts that there is no one at home, then there cannot be any change in window states. This work is in line with the needs and challenges expressed by Yan et al. [35], among others, for the development of better methodologies to model and account for occupant behavior. To build this model, four years of data measured from eight dwellings in a case study building in Canada was used. Again, very limited information was found to be available on window opening monitored data and advanced models for the Canadian context. Bourgeois’ thesis reports on 48 eye observations of window states of a building on Université Laval’s campus [36], but these were for an institutional building and the time-resolution is quite low. Therefore, the second objective of this work is to present and analyze recent data on window openings in Canadian dwellings. The experimental dataset that was used in the present work is described in Section 2, as well as the building from which it was obtained. Section 3 explains the various steps that were followed to develop the probabilistic model. Simulation results and their validation are detailed in Section 4.

2. Window opening behavior from experimental data

The data used in the present study comes from a high-performance social housing building that was built in 2015 and that is located in Quebec City, Canada. Out of the 40 apartments of the building, eight dwellings have been thoroughly monitored since the beginning of the occupancy period of the building (August 2015). These eight dwellings are the apartments located on the corners of the building, so that all "extreme positions" in the building are covered: four on the first floor, four on the top floor (the building is four story high), four on the south façade, four on the north façade, four on the east façade and finally four on the west façade. The apartments located on the north façade of the building have three 80 cm × 140 cm windows in addition to a 91 cm × 214 cm patio door for a total window area of 5.31 m^2 per dwelling (window-to-wall ratio of 8.5%). On the other side of the building, dwellings have a fourth window, increasing the total window area to 6.43 m^2 (10% of WWR). The eight monitored dwellings are inhabited by a total population of ~18 occupants. Monitored data considered in this study include opened/closed state of the windows (1-min frequency), indoor temperature and relative humidity (10-min), outdoor temperature and relative humidity, wind speed and direction and solar radiation (1-hour). Linear interpolation was used to synchronize all data at a 1-min frequency.

Note that 73.9% of the data (from August 2015 to May 2018) was used to train the model while the remaining 26.1% (June 2018 to May 2019) was used for validation. By taking a whole year of data for validation, it is possible to demonstrate the ability of the model to replicate window opening behavior in both winter (defined in this paper as going from October to April) and summer (May to September). The observations made in this subsection come from the training data and are summarized in Table 1. Since dwellings have different numbers of windows and some windows are larger than others, data provided in Table 1 and throughout the paper are the average values found within each dwelling – these averaged values are weighted according to the surface areas of the windows. During that period of 34 months, windows were opened across the dwellings 21.0% of the time (5.04 hours per day) – 10.2% in winter versus 36.1% in summer. Since there is no mechanical cooling in the building, natural ventilation is the main overheating mitigation strategy available to the occupants, hence the need for opening windows in summer. Since the building is very airtight (0.6 air change per hour at 50 Pa) and considering that the mechanical ventilation system in place is often not fully used by households (data has shown that it is used on average 56.6% of the time), another potential reason for opening windows is to improve indoor air quality by increasing the rate of ventilation. This might explain the relatively high use of window openings during winter – a season in which from a thermal standpoint, the windows should be closed as much as possible. A window is opened on average 1.75 times per day (the mean window opening duration is 71.5 minutes) for the heating season. In the summer, this value goes up to 2.20 with an average duration of 247.5 minutes per opening. In other words, the number of openings is relatively similar between winter and summer, but once the window is opened, occupants tend to leave it opened three to four times longer in summer than in winter.

Table 1: Global statistics on the overall use of windows in the eight monitored dwellings.

Dwelling	Winter		Summer	
	Frequency of opened window [%]	Average number of window openings [day^{-1}]	Frequency of opened window [%]	Average number of window openings [day^{-1}]
1	7.4	1.41	54.0	1.87
2	1.1	1.21	36.2	2.32
3	10.3	1.87	44.3	2.01
4	6.9	1.64	49.2	2.22
5	2.0	1.43	15.9	1.03
6	3.3	1.62	22.2	2.87
7	15.0	2.07	27.8	2.12
8	35.9	2.93	39.4	3.14
Total	10.2	1.77	36.1	2.20

Looking at the various patterns found across the sample of eight monitored apartments, there are clear evidences that households act differently regarding their window control. For example, Fig. 1 presents a colormap of the probability of a window being in an opened state for different combinations of indoor and outdoor temperatures based on the measurements. In dwelling #8, there is a relatively high probability of having an opened window and this probability is weakly related to the outdoor and indoor temperatures. The probability of opened windows in dwelling #5 is also weakly linked with the temperatures, because its windows are nearly always closed. On the other hand, temperatures have a high impact on the probability of observing an opened window in other dwellings. For instance, both the outdoor and indoor temperatures have a large effect in dwelling #1. When it is relatively cold outside and inside the dwelling, the windows are typically closed (probability of opened windows below 15%). If the outdoor temperature increases above 15°C but the indoor temperature remains below 25°C, the probability of a window being opened in dwelling #1 increases up to 35%. For cases where both the outdoor and indoor temperatures are relatively high, this value can go up to 70%. In dwelling #2, the outdoor temperature appears important, but not the indoor temperature – there is a clear shift in the frequency of opened windows on the outdoor temperature axis, but no such shift on the indoor temperature axis.

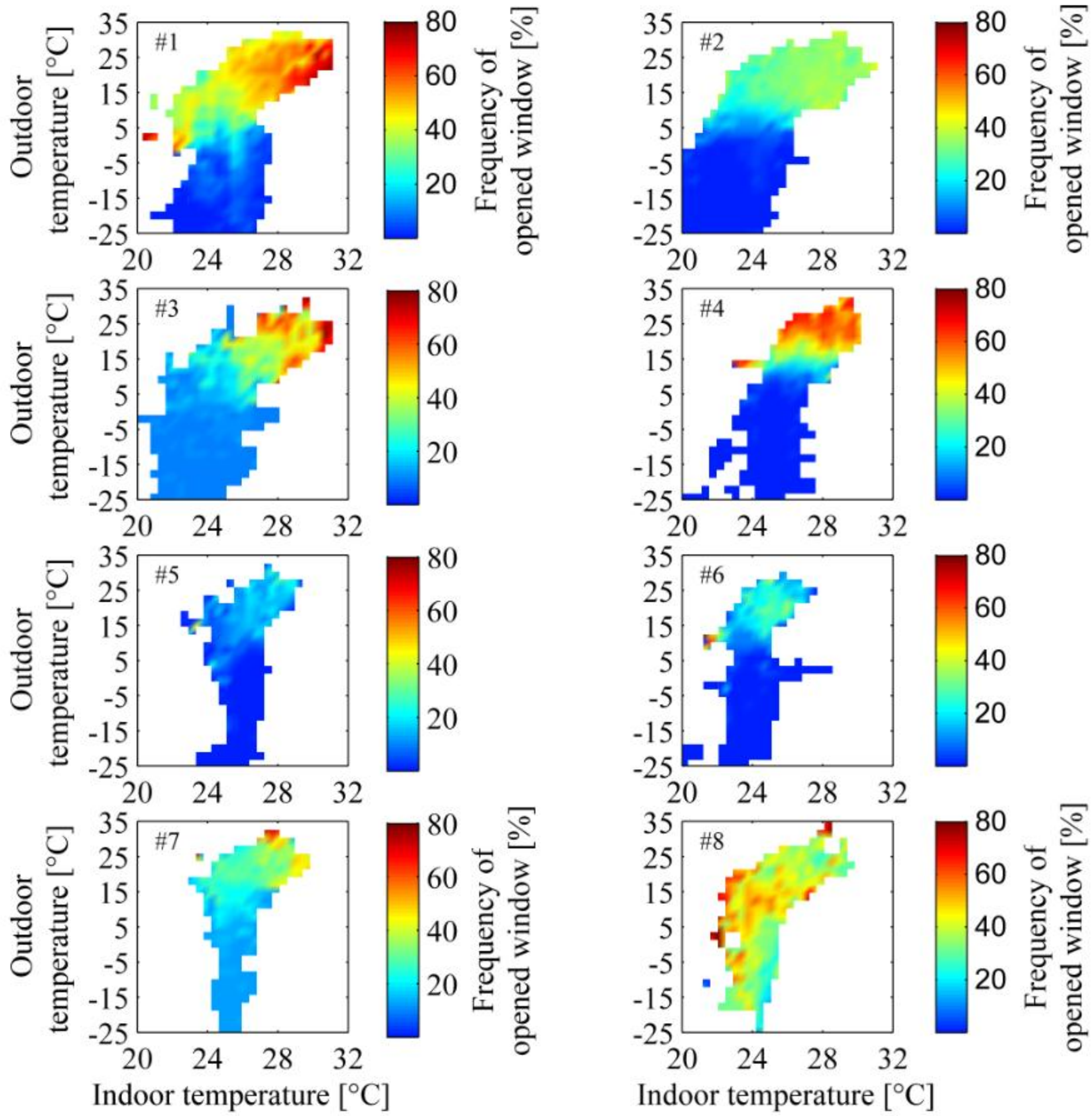

Figure 1: Colormap describing the link between the indoor and outdoor temperature and the probability of observing an opened window for eight apartments in the case study building based on measurements.

3. Development of the window opening model

3.1 Presentation of logit regression

A logit regression is chosen to compute the probability of observing an open window given a certain set of independent variables:

$$\mathrm{logit}\left(p_{opened}\right) = \ln\left(\frac{p_{opened}}{1-p_{opened}}\right) = \omega_0 + \sum_{i=1}^{M} \omega_i x_i \quad (1)$$

where p_{opened} is the probability of a window being opened, x_i represent the M independent variables considered by the model and ω_i, their related regression coefficients. The logit regression is one of the most typical approaches for developing a probabilistic window opening model. The probabilistic nature of the model comes from the fact that, for a given set of parameters x_i (e.g., indoor and outdoor temperatures, hour of the day, indoor air quality…), Eq. (1) only provides a probability of finding the window in an open state. In other words, using such a model can lead to different window states for the exact same conditions. However, this approach often assumes that all occupants act the same way regarding window control. When applying a model based on Eq. (1) on multiple households, the window opening schedule of each household will be different because the timings of window openings will not necessarily match, but on aggregate they will be highly similar with the same total frequencies of window openings, same reaction towards high indoor temperatures, etc. As seen in Table 1, very different window opening patterns can be observed in reality. To replicate this diversity, it is suggested to employ regression coefficients ω_i computed from probability density functions (assumed to follow normal distribution) instead of having deterministic/fixed values. This approach is in line with that proposed by Haldi et al. [25] which, so far, has only been tested once for European dwellings. Apart from the set of data on which they rely, there are also some other differences between the present model and that of Haldi et al. To name a few: in the present model, CO_2 data is not included as this information is not always available in many buildings; in Haldi's model, the probability of action were different in each room typology (e.g., kitchen, living room, etc.) whereas the present model is for the entire dwelling to facilitate its implementation in energy simulation tools; as described below, correlations between weights in the regressions were used to limit the number of variable weighting factors and account for observed relations between them (see Section 3.4).

3.2 Selecting appropriate predictors for the window control model

The influence of nine different parameters on the probability of observing an opened window were assessed: the indoor and outdoor temperatures, the indoor and outdoor relative humidities, the minute of the day, the day of the year, the total horizontal solar radiation, and the wind speed and direction. These nine predictors were first ranked according to their importance on the window state. This was done by building univariate regression models for each of the predictors to understand the relationship between the probability of finding an opened window versus individual studied predictors. These univariate models were created by merging all eight monitored dwellings into a single dataset of window opening behavior. Fig. 2 displays the various plots obtained when measuring the average window states for different values of each of the predictors. Predictors in Fig. 2 are ranked from the most significant parameters (top row of the figure) to the ones that have the lowest impact (bottom row). The criterion used to rank the parameters was the F-test of overall significance of the regression equations, which was computed for each univariate regression models displayed in Fig. 2.

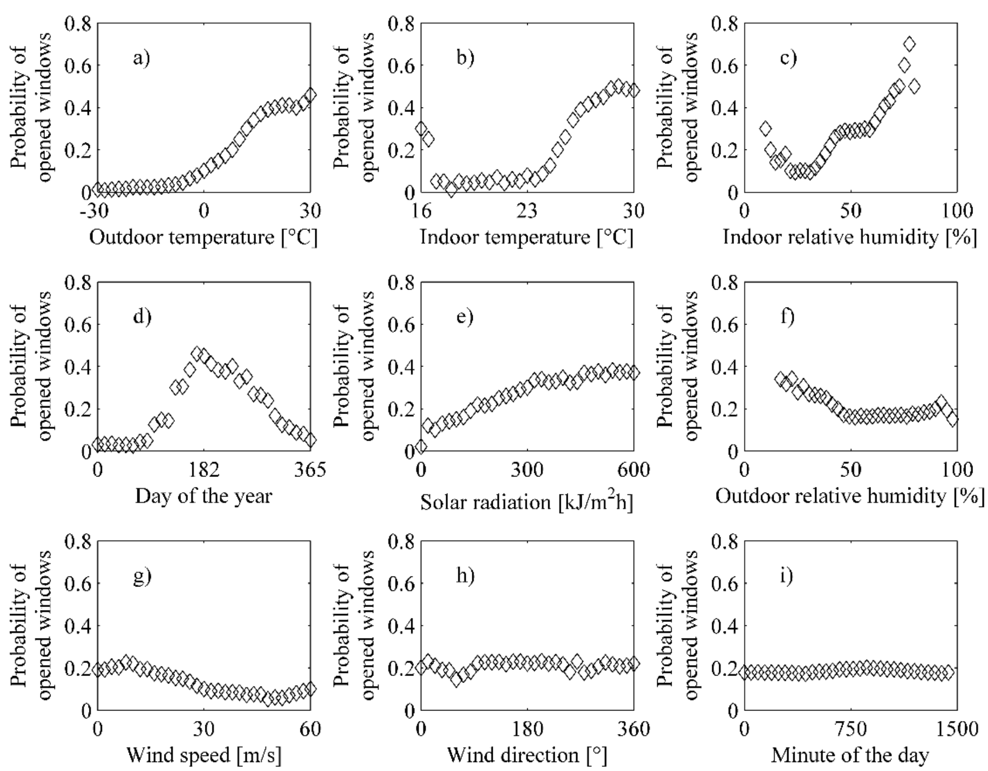

Figure 2: Observed probability of opened windows versus each of the considered predictors for the window opening model individually.

Both indoor and outdoor temperatures are highly significant, with the probability of opened windows following quite closely a logit regression. Indoor relative humidity also appears important with the probability of observing an opened window being lower for dry indoor environment. However, this correlation could be explained by the fact that in Canada dry indoor conditions typically occur during the winter, when the outdoor air is very cold and thus contains a minimal amount of vapor. Since it is cold outside during the winter, occupants limit window openings hence lowering the probability observed for dry indoor environments. On the other hand, very humid indoor conditions happen during the warmest days of the summer, when one should expect occupants to open their windows. A similar explanation can be provided for the relatively strong correlation observed between window states and the day of the year. Fig. 2 shows that there is a peak of window opening happening during the summer, when it is warmer. Window openings are linearly proportional to solar radiation – radiation is another variable that is highly correlated with the outdoor temperature. As for outdoor relative humidity, wind speed, wind direction and the minute of the day, their impact on the occupants' control of their windows seems to be insignificant. Note that a figure similar to Fig. 2 was produced for individual windows on different façade, and the low impact of these variables was also observed.

To quantifiably evaluate which parameters must be considered to obtain an adequate window opening model, nine different multivariate models were created for each of the eight monitored dwellings. First, a univariate model was implemented using only the most significant predictor (outdoor temperature). Then, the second regression model was developed by adding the second most significant predictor (indoor temperature). This process was repeated until all nine predictors were included in the regression analysis. The goodness-of-fit of each model was calculated with Nagelkerke's R^2 value, which is suited for the comparison between a logistic regression model and a baseline. Table 2 provides the Nagelkerke's R^2 values obtained for all 72 regression models (8 dwellings times 9 models per dwelling). When only considering outdoor temperature to predict the state of windows, the average R^2 value found across the eight dwellings is 0.335. With the indoor temperature, this value increases up to 0.415 versus a value of 0.469 when all predictors are considered. Considering that seven other predictors had to be added into the regression models to increase the average R^2 value from 0.415 to 0.469, it appears that the two-variable regression models are more efficient. In fact, the adjusted R^2 values of the model, which consider the number of predictors employed by the model, goes from an average of 0.384 for the two-predictor models to an average of 0.309 for the models that use all variables. Therefore, since it is able to produce adequate results, the two-variable regression analysis strictly based on temperatures was chosen as the most appropriate approach to build a stochastic window state model as simple as possible that still accounts for occupant diversity.

Table 2: Goodness-of-fit (Nagelkerke's R^2) of the logistic regression models for window states in each dwelling according to the considered predictors.

Model	Dwelling 1	Dwelling 2	Dwelling 3	Dwelling 4	Dwelling 5	Dwelling 6	Dwelling 7	Dwelling 8
Only outdoor temperature	0.372	0.185	0.402	0.267	0.358	0.409	0.240	0.447
Adding indoor temperature	0.509	0.249	0.484	0.316	0.412	0.494	0.362	0.495
Adding indoor relative humidity	0.529	0.272	0.486	0.327	0.413	0.510	0.386	0.515
Adding day of the year	0.540	0.281	0.504	0.334	0.423	0.519	0.412	0.532
Adding solar radiation	0.550	0.292	0.515	0.345	0.427	0.530	0.428	0.543
Adding outdoor relative humidity	0.557	0.300	0.523	0.356	0.433	0.537	0.435	0.549
Adding wind Speed	0.561	0.305	0.529	0.362	0.439	0.543	0.437	0.553
Adding wind direction	0.562	0.308	0.529	0.363	0.441	0.546	0.438	0.555
Adding minute of the day	0.562	0.309	0.530	0.364	0.442	0.546	0.439	0.556

3.3 State model versus change-of-state model

Two types of window opening model can be created: one that predicts the state of the window (i.e., for a given time step, what is the probability of finding the window in an opened or closed state?) and one that predicts changes of the state of the window (i.e., for a given time step, will there be a window opening (if the window is closed), a window closure (if opened) or no change of state?). For the "state model", the logit equation is expressed as:

$$\mathrm{logit}\left(p_{open}\right)=\ln\left(\frac{p_{open}}{1-p_{open}}\right)=\omega_{in}T_{in}+\omega_{out}T_{out}+\omega_{const} \tag{2}$$

Eq. (2) is used for all time steps in the simulations to compute the probability of observing an opened window according to the indoor and outdoor temperatures. A random number is then drawn and compared to p_{open} to determine if the window is opened or not. In the "change-of-state" or "action" model, two logit equations are implemented. The first one provides the probability of observing a window opening and thus is only used for time steps when the window is closed:

$$\mathrm{logit}\left(p_{op}\right)=\ln\left(\frac{p_{op}}{1-p_{op}}\right)=\omega_{op,in}T_{in}+\omega_{op,out}T_{out}+\omega_{op,const} \tag{3}$$

Then, when the window is predicted to be opened, the model shifts to another logit equation to forecast when the window will eventually be closed:

$$\mathrm{logit}\left(p_{clo}\right)=\ln\left(\frac{p_{clo}}{1-p_{clo}}\right)=\omega_{clo,in}T_{in}+\omega_{clo,out}T_{out}+\omega_{clo,const} \tag{4}$$

A test was made with the training data to compare the performance of both methodologies. For each of the eight apartments, both the "state" and "change-of-state" models were sequentially used to predict window openings in the apartment. The measured indoor and outdoor temperatures from the apartment were used as inputs to both models. The two models yielded similar predictions for the total duration of opened windows (across the eight dwellings, "state" model overestimated the frequency of opened window by 5.1% versus an overestimation of 5.4% for the "change-of-state" model), but there was a major difference concerning the number of window opening events. The "state" model overestimated the number of window openings by a factor of nearly 100 – the window schedule for this model kept shifting between the "opened" and "closed" states with unrealistic delays between the changes of state. In comparison, the "change-of-state" model only overestimated window openings by 6.8%.

It was thus decided to develop change-of-state (window opening and closure) models instead of state (opened or closed window) models for the integrated occupant behavior model introduced above. The tests made in Section 3.1 were repeated for window openings and closures and yielded similar results to the ones obtained with the state model, so only the indoor and outdoor temperatures were used as predictors to compute the probability of a window event. Eqs. (3) and (4) were employed in a regression analysis of the measured data for each dwelling, which allowed for the estimation of the ω coefficients (see Table 3). The regression coefficients had different values from a household to another, again illustrating the diversity of window opening behavior. For instance, $\omega_{op,const}$ goes from a value of -2.07 in dwelling #8 up to -10.09 in dwelling #7. In dwelling #5, the indoor temperature appears to have no effect on the probability of opening a window. The normal distribution assigned to each coefficient was then defined by each coefficient's mean value and standard deviation.

Table 3: Logit regression coefficients for the calculation of the probability of window openings and closures for each monitored dwelling.

Dwelling	Window openings			Window closures		
	$\omega_{op,in}$	$\omega_{op,out}$	$\omega_{op,const}$	$\omega_{clo,in}$	$\omega_{clo,out}$	$\omega_{clo,const}$
1	0.10	0.03	-7.68	-0.11	-0.03	-1.83
2	0.02	0.04	-5.90	-0.15	-0.03	1.08
3	0.06	0.02	-6.72	-0.08	-0.02	-1.79
4	0.06	0.04	-6.69	-0.06	-0.03	-1.63
5	0.00	0.04	-5.66	-0.06	-0.05	-0.51
6	0.01	0.03	-4.92	-0.06	-0.05	-0.54
7	0.19	0.04	-10.09	-0.14	-0.01	0.04
8	0.03	0.02	-2.07	-0.07	0.00	-1.79
Mean	0.059	0.033	-6.216	-0.091	-0.028	-0.871
Standard deviation	0.062	0.009	2.296	0.037	0.018	1.073

To verify the accuracy of the eight regression models, the eight different sets of ω coefficients in Table 3 were provided to the window opening model along with the corresponding indoor and outdoor temperatures measured during the training period. For each dwelling, the model was used to perform 100 yearly simulations of the window opening schedules, and then the frequency of opened windows in the summer and in the winter were evaluated for each simulation. The results of these 100 simulations per dwelling are displayed in Fig. 3 by grey crosses and compared to the real behavior of occupants (red diamonds). All eight measured points fall into the scatter plot generated by their corresponding set of simulations. This demonstrates that it is possible to replicate the general window behavior of occupants when using appropriate values for the ω coefficients.

Figure 3 also provides the comparison between the average frequency of opened window measured in the case study building and the outputs of the window opening model when using the mean values of the regression coefficients presented in Table 3. Measurements and simulations are once again in agreement, so one can reproduce the aggregated behavior of the occupants when using averaged regression coefficients. However, the diversity of behaviors is not captured at all with this approach. There is thus a need to develop a methodology to reproduce dwelling-to-dwelling diversity, as will be explained below.

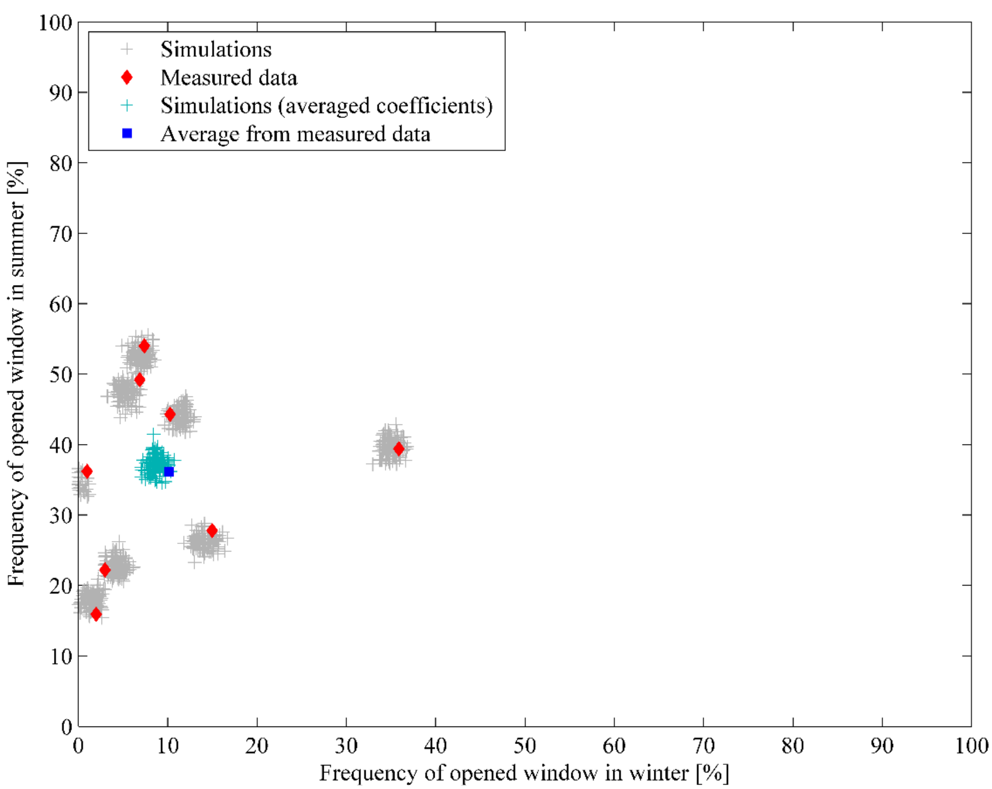

Figure 3: Comparison of the window opening model (when using specific regression coefficients for each dwelling) with measurements - Scatter plot of the frequency of opened window in winter and summer for the simulated and monitored dwellings. 100 data points were simulated for each dwelling.

3.4 Effect of the number of randomly selected coefficients for the window control model

When simulating window operations in a dwelling, one must specify the six ω-values appearing in Eqs. (3)-(4). In order to generate different possible window control behaviors, one could randomly select all six coefficients based on the distributions found above and simulate the window operation accordingly. However, doing so might induce unrealistic window control behaviors if an unrealistic or unfeasible combination of regression coefficients was randomly selected by the model. When studying the regression coefficients observed in each of the eight monitored dwellings, correlations were also found between some of the coefficients. By randomly selecting all coefficients, the model cannot reproduce these correlations and consequently does not reflect well what is observed in reality. For these reasons, it was considered preferable to randomly select only some of the regression coefficients and use observed correlations to compute the remaining coefficients that were not randomly chosen.

An analysis was made to decide how many coefficients (and which ones) should be picked randomly and then use to calculate the others based on correlations. The number of ω coefficients to be randomly chosen was incrementally increased from one to six. For the "one randomly selected coefficient" model, $\omega_{op,in}$ was picked as the variable to be assigned a random value from its probability density function since it was the coefficient that had the highest level of correlation with the other ω coefficients. This was determined by calculating the coefficient of determination R^2 for each pair of ω coefficients using the data of Table 3. Then, for the "two randomly selected coefficients" model, the correlation of each pair of the five remaining coefficients combined with $\omega_{op,in}$ were once again computed in a two-variable regression analysis (e.g., using $\omega_{op,in}$ and $\omega_{op,out}$ as independent variables and $\omega_{op,const}$ as the dependent one) and evaluated with the R^2 value. The coefficient that shown the highest level of correlation with other coefficients when combined with $\omega_{op,in}$ was chosen as the second variable to pick at random. This process was repeated until all six coefficients were ranked: $\omega_{op,in}$, $\omega_{op,out}$, $\omega_{clo,in}$, $\omega_{clo,const}$, $\omega_{op,const}$ and $\omega_{clo,out}$.

The six window models were tested by simulating window openings 10,000 times for each model using the outdoor temperature and the indoor temperature profiles in the monitored dwellings. Fig. 4 shows the frequency of opened window in summer and in winter for each of the 10,000 simulations (grey crosses), while the red diamonds show where the eight monitored dwellings stand. By inspecting Fig. 4, it is clear that with only one coefficient selected randomly, it is not possible to reproduce the various behaviors observed in the monitored building as the simulation outputs produce a very narrow shape that is off compared to measured data. The five other models match relatively well with measured results, with the “two randomly selected coefficients” and “three randomly selected coefficients” models offering the best matches. The models using four, five or six randomly selected coefficients generate “shapes” in the figures such that the measured behaviors (red diamonds) are located at the margins or even appear to be outliers (the two upper-left diamonds in Fig. 4d and rightmost diamond in Fig. 4e and 4f). The main difference with “two randomly selected coefficients” and the “three randomly selected coefficients” models is that the models with three random coefficients generates a larger frequency of very high window opening behaviors. 5.9% of the simulations with this model yielded more opened windows in winter than the maximum data that was measured (i.e. higher than 35.9% of window opened) and 23.4% of simulations had more opened windows in summer than the maximum measured data (i.e. higher then 54.0%). For the models with two random coefficients, these values respectively are 3.4% and 5.1%, which appears closer to the data and to typical observations in the Canadian context.

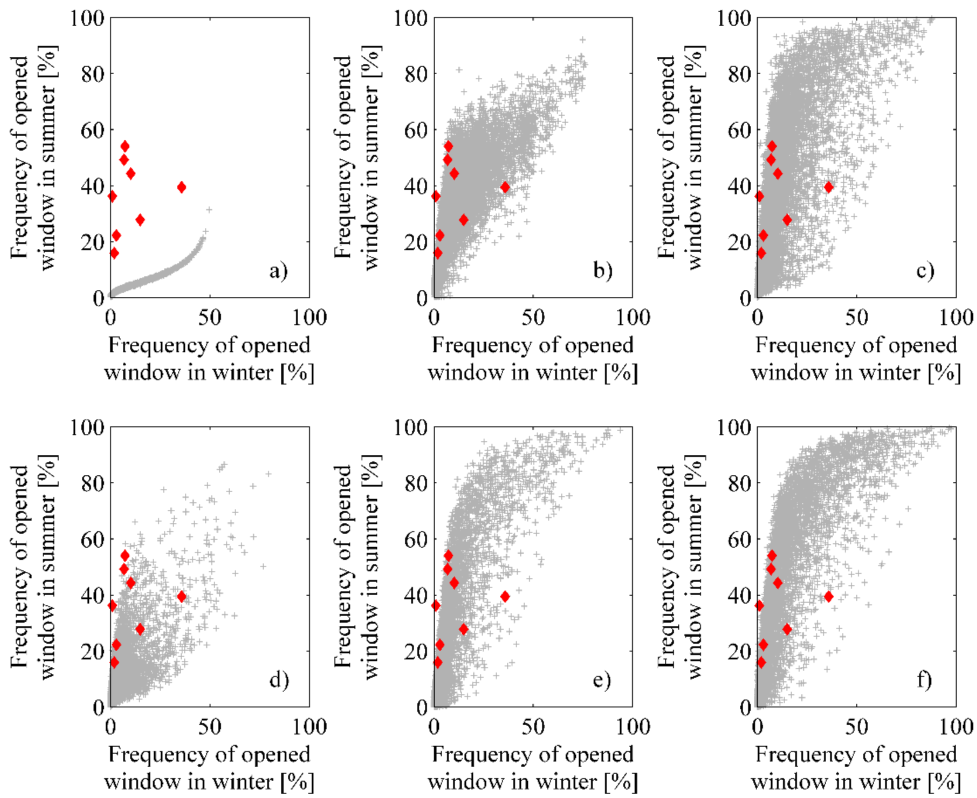

Figure 4: Scatter plot of the frequency of opened window in winter and summer for the simulated and monitored dwellings when randomly selecting a) one, b) two, c) three, d) four, e) five and f) six regression coefficients in Eq. (8). Grey crosses represent the 10,000 simulations per scatter plot and the eight red diamond represent measured data.

The "two randomly selected coefficients" methodology was then chosen for the proposed window opening models. It should be noted, though, that the other methodologies to generate the ω-coefficients can also offer acceptable results and that the best approach might depend on the dataset. In this model, $\omega_{op,in}$ and $\omega_{op,out}$ are randomly selected from their probability density functions and values for other coefficients are directly calculated afterwards using these correlations found by regression analysis:

$$\begin{aligned} \omega_{op,const} &= -27.2\omega_{op,in} - 98.5\omega_{op,out} - 1.42 \\ \omega_{clo,in} &= 0.30\omega_{op,in} + 1.07\omega_{op,out} + 0.04 \\ \omega_{clo,out} &= -0.17\omega_{op,in} + 1.01\omega_{op,out} \\ \omega_{clo,const} &= -2.18\omega_{op,in} + 80.7\omega_{op,out} - 3.37 \end{aligned} \tag{5}$$

Figure 5 shows different probability curves obtained with the proposed methodology for selecting the ω coefficients in order to display how they influence the probability of observing a window change of state. The black lines are the probability obtained with Eqs. (3) and (4) when using the mean regression coefficients from Table 3. Red curves represent five different random draws using normal distributions for $\omega_{op,in}$ and $\omega_{op,out}$ and Eq. (5) for the remaining regression coefficients. Graphs on the left exhibits the probability curves for a window opening, whereas the right-hand side graphs represent the probabilities for a window closures. The top ones are for a constant outdoor temperature, but with a varying indoor temperature. The bottom graphs depict the opposite, with the indoor temperature being constant and the outdoor temperature changing. The fixed temperatures were set at a high value (30°C) to emphasize the chances of windows being opened. The differences between the curves in each graph reveal the high variability of window behavior that the model can produce – some generated households are highly influenced by temperatures in terms of window opening behavior whereas in some others, probabilities of a change of state of windows is mostly unaffected and nearly constant for all temperatures.

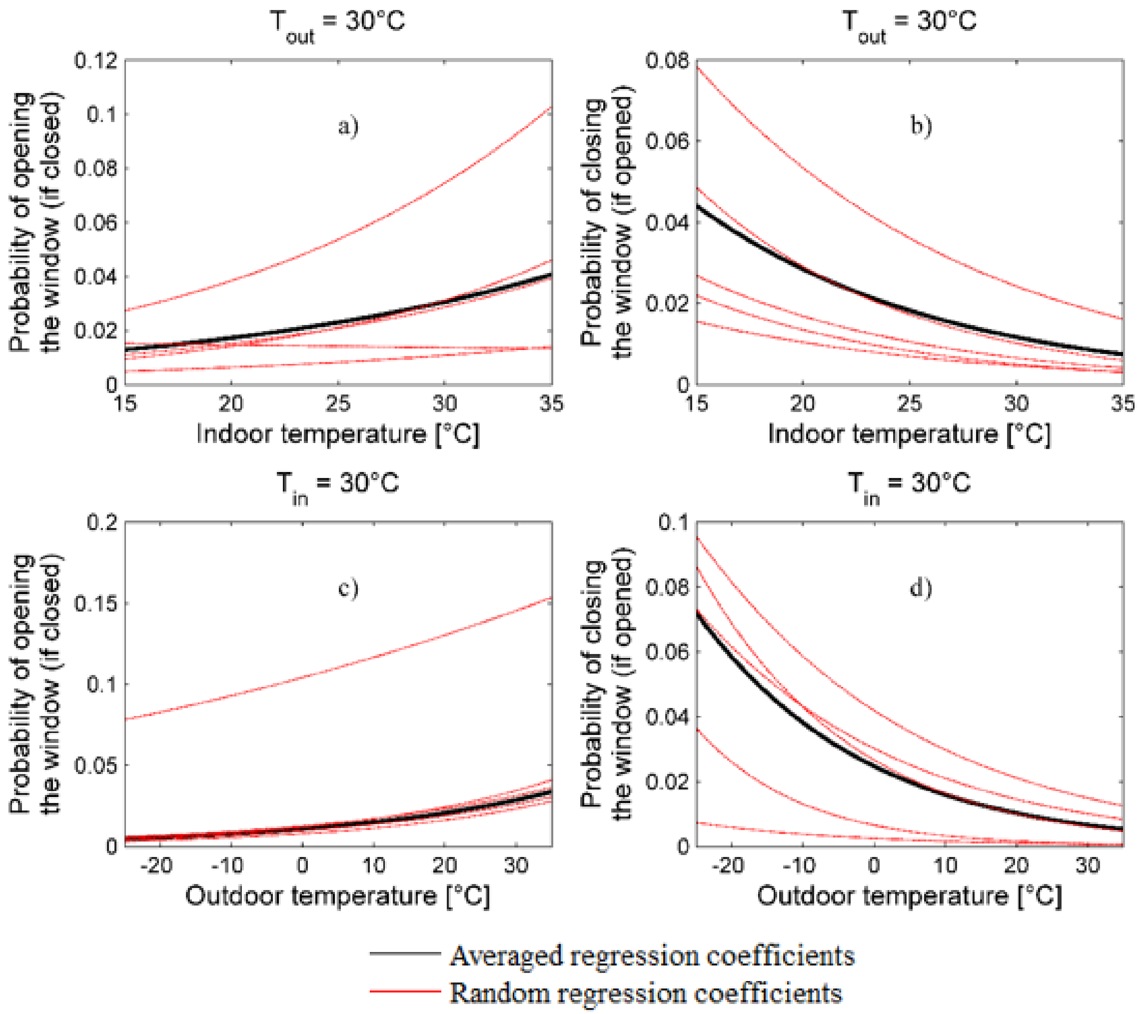

Figure 5: Example of the impact of randomly drawing regression coefficients on the probability of opening or closing a window. Black lines are generated using averaged regression coefficients and red lines when using the proposed methodology.

3.5 Coupling with active occupancy model

Since the window operation model is intended to be part of a unified occupant behavior model, modifications were made to ensure that the unified model remains as comprehensive as possible with no contradiction. For example, predicting a window change of state when there is no active occupancy in a dwelling (i.e., no one is awake and at home in the dwelling) would be a contradiction. Therefore, the window opening part of the unified occupant behavior model first receives the active occupancy schedule generated by the occupancy part of the overall model which is described extensively here [37], [38]. This active occupancy schedule is produced by first assigning a household size for a simulated dwelling and then using Markov-chains based on data from time-use surveys. These surveys provide the times when people arrive and leave their home and are comprised of tens of thousands of journal entry. For each simulated household, a random “type of occupants” factor modifies the Markov-chains to reflect the fact that some families spend more time at home than others. When the window behavior model gets the active occupancy schedule, it sets the probability of a window event occurring at zero for time steps in which no one is predicted to be actively at home. On average 26.1% of the times steps of the occupancy schedules generated by the model have zero active occupants. Limiting the changes of window state to time steps with active occupants would thus reduce the number of change-of-state events by 26.1%. To balance this, probabilities are multiplied by a factor of 1/(1 – 0.261) = 1.35 and the new regression equations become:

$$\begin{aligned} \mathrm{logit}\left(\frac{p_{op}}{1.35}\right) &= \omega_{op,in} T_{in} + \omega_{op,out} T_{out} + \omega_{op,const} && \text{if } Occ > 0 \\ p_{op} &= 0 && \text{if } Occ = 0 \end{aligned} \tag{6}$$

Equations for the probability of a window closure have the same form as Eq. (6). An example of this adjustment is illustrated in Fig. 6 for a single day in January. In the original probability curve calculated with Eq. (3), the probability of opening a window oscillates between 0.82% (at 02:10) and 1.04% (at 15:50) depending on the temperatures observed during that day. However, in this example there is no active occupant before 07:00, between 13:20 and 21:50 and after 22:30. There cannot be a window opening during these periods of time, so the new probability curve is set to zero for these time steps. The resulting curve is then multiplied by 1.35 to ensure that the number of window events throughout the year is unaffected by this change. The maximal probability to open a window during that day is now 1.36% and happens at 13:10.

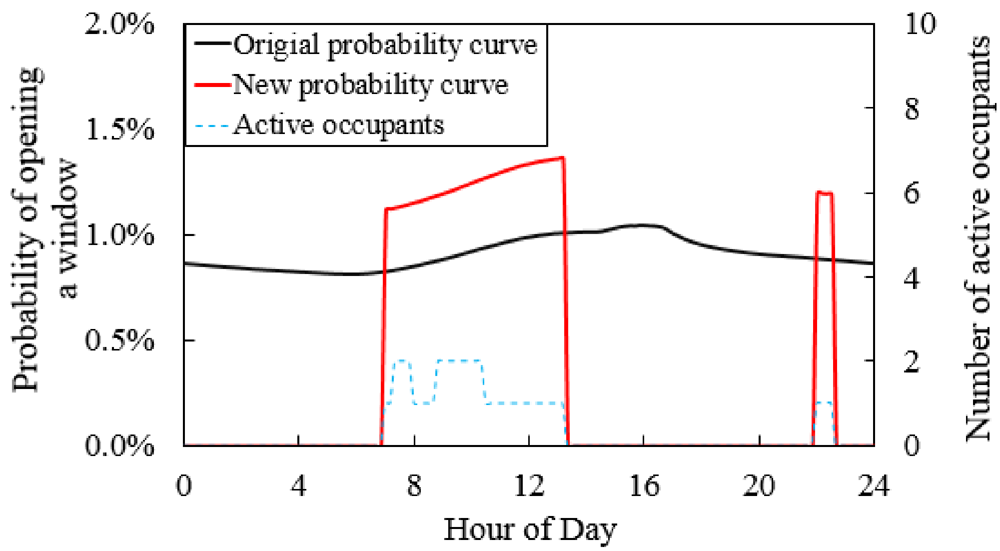

Figure 6: Example of an adjustment of the probability of opening a window when considering active occupancy.

4. Validation

As mentioned, a complete year of data (June 2018 to May 2019) was devoted to the validation of the window opening model. The aim of the validation is to ensure that the window opening model is able to replicate both the aggregated behavior of occupants living in the case study building and the diversity of behaviors observed between various households. The ability of the model in reproducing the diversity of behaviors was verified by simulating window openings 10,000 times using the outdoor and indoor temperature profiles in the monitored dwellings during the validation year. The overall frequency of opened windows in winter and summer from these 10,000 simulations are shown in Fig. 7a, along with the measurements made from the monitored dwellings, each red or blue mark corresponding to one monitored dwelling. The eight dwelling window control behaviors are mostly covered by the scatter plot formed by simulation results.

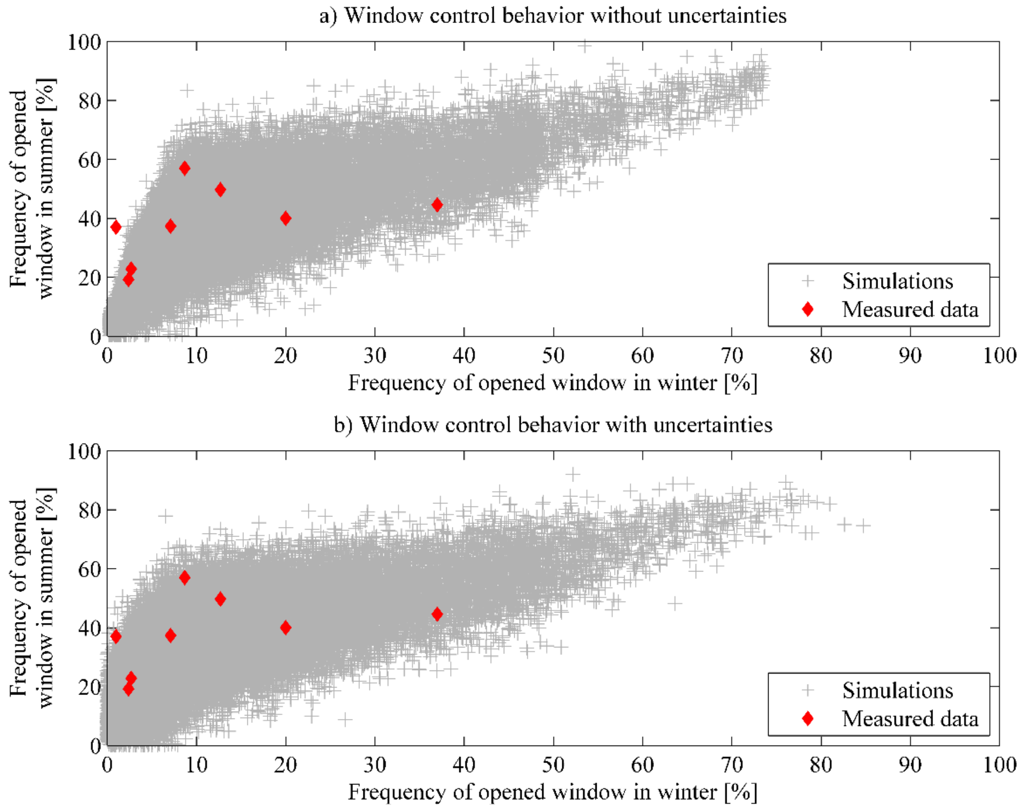

Figure 7: Comparison of the proposed window opening model with measurements - Scatter plot of the frequency of opened window in winter and summer for the simulated and monitored dwellings.

One data point is slightly out of the scatter plot. This measured data point has a frequency of opened window of 1.5% in winter versus ~4.0% for the minimal value found in simulations for the same summer opening frequency. This is caused by the fact the correlations between the ω-values expressed in Eq. (5) are not "perfect", i.e. the winter and summer behaviors are not perfectly correlated. As in any correlation, there is a certain level of uncertainty on the result returned by the correlation, but in this case, this uncertainty is hard to assess due to the small sample size (i.e., 8). This uncertainty is not considered by the model, i.e. the results returned by the correlations are directly provided to Eq. (6). To test the impact of the uncertainty of the correlations, random variations within ±15% was assigned to the coefficients of Eq. (6) and the process behind Fig. 7a (e.g. simulating the window control behavior of 10,000 households) was repeated. The resulting figure is presented in Fig. 7b. It can be seen that with the additional uncertainties, the out-of-range measured data point is within the cloud of points generated by the model. However, since data is limited to eight dwellings, it is difficult to have a thorough estimation of the uncertainties related the coefficients of Eq. (6). Consequently, it was deemed preferable not to include these uncertainties in the rest of the simulations presented in this paper, in particular since the offset between simulations from the current model and the single out-of-range data point is small.

As for the aggregated window opening behavior, the average frequency of window opening in the eight monitored dwellings during the validation year was 8.7% in winter and 39.3% in summer. The corresponding values computed out of the 10,000 simulations respectively are 9.2% (relative overestimation of 5.7%) and 42.4% (relative overestimation of 7.9%). During the validation year, there was a total of 2.04 openings of window per window per day in the case study building versus a prediction of 2.10 openings from the model.

In terms of the duration of window openings, a window in the case study building stays opened on average 149.0 minutes per opening (66.7 minutes in winter and 239.8 minutes in summer). Outputs of the model predict that windows are opened 69.8 minutes per opening in winter and 255.8 minutes in summer for an overall average of 158.3 minutes per opening (overestimation of 6.2%). The distributions of the duration of window openings from monitored data and simulations are provided in Fig. 8. Both distributions follows the lognormal law with the plurality of openings having relatively low durations (less than 20 minutes). In the case study building, 32.3% of openings had a duration smaller than 20 minutes during the validation year versus 26.9% in simulations. The model underestimated the number of very short window openings, which could explain the slight overestimation of the mean duration for a window opening. An inspection of the measured dataset shows a relatively high number of window openings (7.2%) that have lasted a single minute. These really short openings should have negligible impact on the indoor environment and the consumption of energy, and could potentially be linked to a misreading by the sensors. If these 1-minute openings are taken off the measured dataset, the average duration for a window opening in the case study building reaches 162.3 minutes per opening, which is closer to the simulate average. Note that the x-axis of Fig. 8 cuts off at 500 minutes for the sake of visibility. In both measured and simulated datasets, there are few window openings that last up to 20,000 minutes (approximately 14 days). 1.4% of the window openings in the case study building lasts more than 1,440 minutes (a day). In the simulated dataset, 1.2% of openings lasts that long. In short, the simulated distribution of window opening durations follows the same pattern as the one from the measured data, except for an underestimation of very short window openings. The mean window opening duration is similar in the measured and simulated datasets.

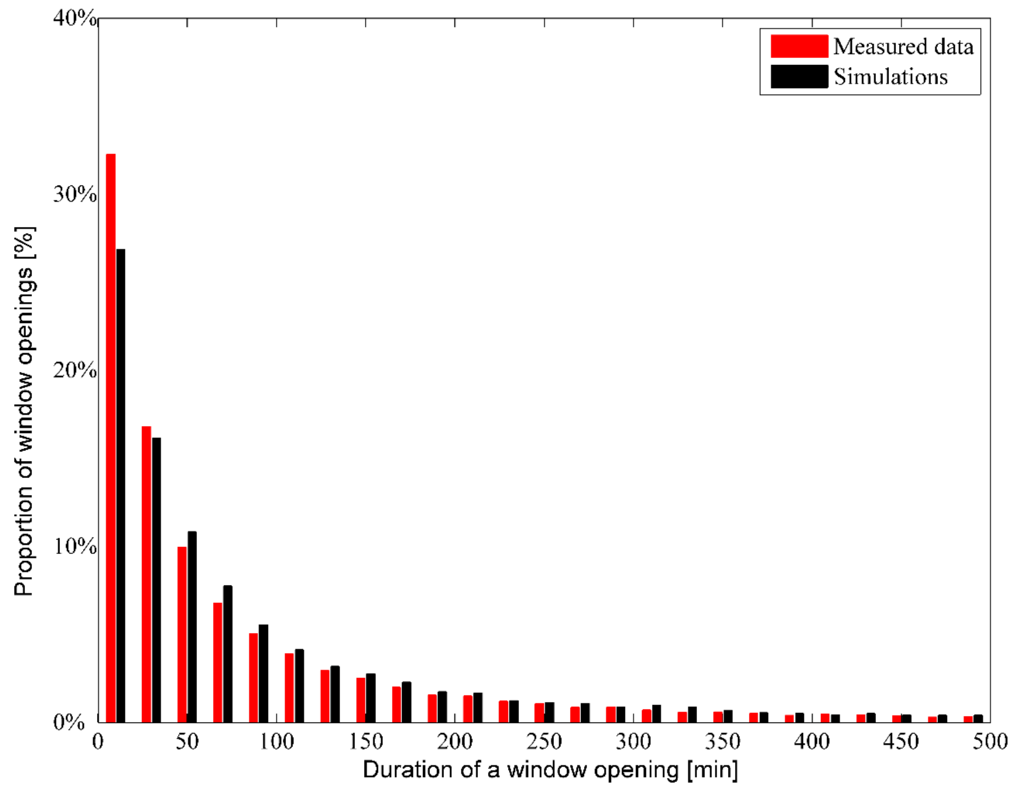

Figure 8: Measured and simulated distribution of the duration of window openings during the validation year.

Another point to validate is the decision made to consider only the indoor and outdoor temperatures as predictors for the openings and closures of windows in the model. In Section 3, it was argued that correlations observed between other predictors (relative humidity, solar radiation…) and the probability of observing a window opened are mostly due to the fact that these other predictors are highly correlated with either the indoor or the outdoor temperature. To make sure that this statement is true and that no pattern was missed by excluding other predictors, the correlations between these predictors and the state of the windows were computed for the validation year from both measured and simulated data. These correlations are shown in Fig. 9, with the red diamonds representing measured data from the eight monitored apartments and the black crosses the results of all simulations. The curve outlined by the data points of measurements and simulations highly fit together, proving that the correlations observed with measurements and simulations are mostly the same, so cutting most of the predictors did not leave out patterns that simulations cannot replicate. Table 4 provides the coefficients of determination between both data series from the nine graphs of Fig. 9. The only two predictors that yielded relatively poor fits between simulations and measurements are the wind direction (R^2 = 0.382) and the minute of the day (R^2 = 0.483). This is due to the fact that these two variables are weakly correlated with indoor and outdoor temperatures. Since both wind direction and minute of the day have very small influence on the state of the windows, these poor fits between simulations and measurements are not expected to affect the overall performance of the model. For other predictors, the coefficients of determination between measurements and simulations are very high, with five of them having R^2 values above 0.9. The model is thus able to replicate the univariate behavior of window opening in accordance to each of the important predictors.

Table 4: Coefficient of determination between the logistic regression models for window states from measured and simulated data according to the considered predictor.

Predictor	R^2
Outdoor temperature	0.989
Indoor temperature	0.959
Indoor relative humidity	0.946
Day of the year	0.967
Solar radiation	0.923
Outdoor relative humidity	0.794
Wind speed	0.702
Wind direction	0.382
Minute of the day	0.483

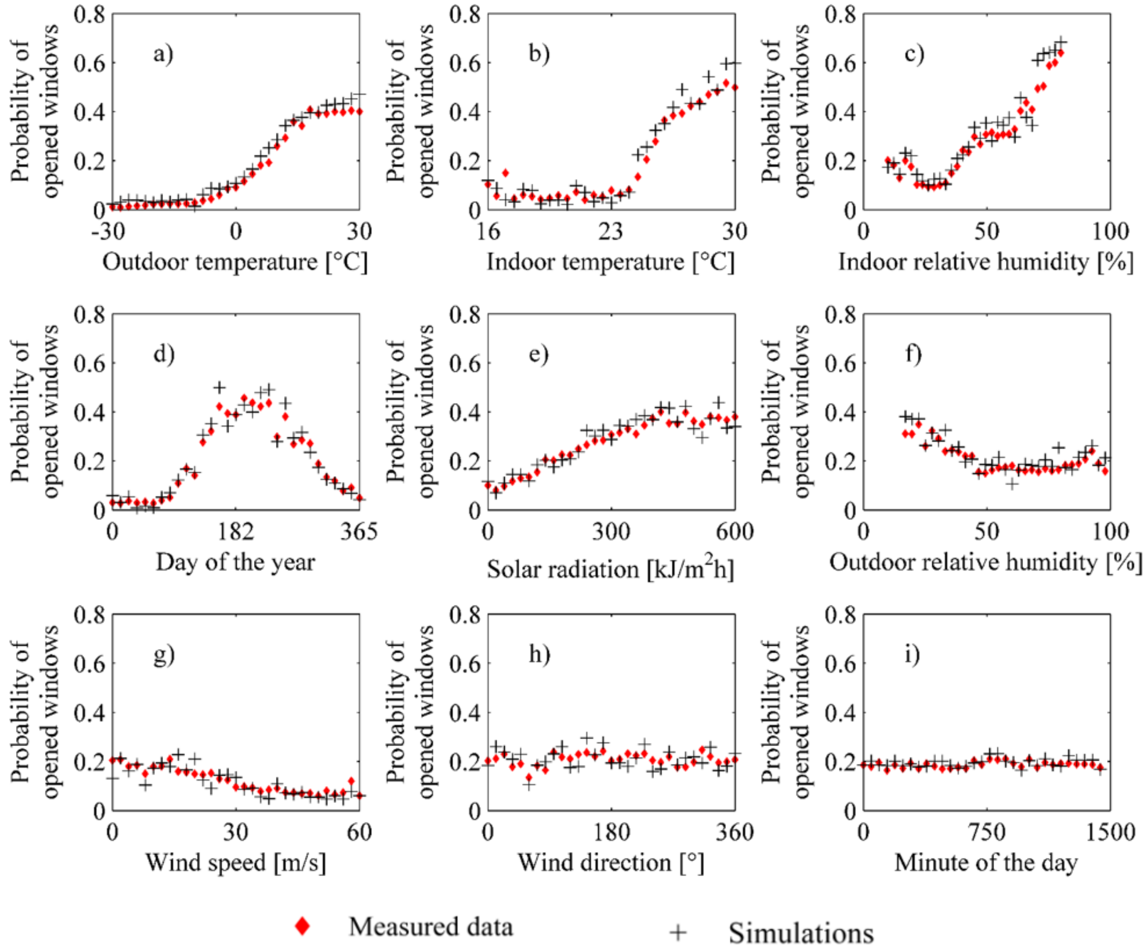

Figure 9: Simulated (black crosses) and observed (red diamond) probability of opened windows versus each of the recorded predictors for the window opening model individually during the validation year.

5. Conclusions

Since windows were found to play a significant role on indoor air quality in social housing buildings, developing a better understanding of how windows are controlled in low-income buildings and how to take that into account in the design and operation would be impactful for this population. A probabilistic window opening model for such buildings is thus developed in this paper, which includes randomly drawn coefficients to account for occupant behavior diversity. Observations made from a case study building clearly showed that there is a high variability of window opening behavior across different households.

The model developed in this paper is to be included in a unified occupant behavior tool. In order for this tool to be fully coherent, it was ensured that there could not be window opening or closure in times when no one is predicted to be awake at home. Data coming from eight dwellings in a high-performance social housing building was used for the development of the window opening model, which is a change-of-state model that predicts window openings and closures based on the outdoor and indoor temperatures. These two predictors were chosen as their effect on the frequency of open windows in a dwelling is much larger that of other variables such as relative humidity or solar radiation. Based on the logit equation, this window opening model employs a total of six coefficients (three to predict window openings, three to predict window closures). It was deemed preferable to randomly draw two of these six coefficients from normal density functions and to calculate the remaining terms from correlations observed between the coefficients. This choice was made to ensure that observed correlation between the coefficients measured from the case study building remain intact in the model and that extreme combinations of coefficients that lead to unrealistic window opening behaviors cannot be reached by the model.

A comparison was made between the outputs of the model and the measured window behavior observed during a validation year. This comparison showed that the model was able to adequately simulate the "average" window opening behavior and the diversity of behaviors observed across the eight monitored dwellings during both winter and summer. The frequency of opened windows found in the monitored building matched with the predictions of the model. The same can be said for the distribution of the durations of window openings. Only using the outdoor and indoor temperatures as predictors for the change of state of windows appears sufficient, since the correlations between the state of windows and other variables, such as relative humidity or solar radiation, are similar in measured and simulated datasets.

Like most probabilistic window opening models, the model developed in this paper employs logit regression to compute the correlation between the probability of a window opening and environmental stimuli. The novelty of this paper comes from the fact that these correlations are not based on fixed regression coefficients, but on coefficients that are randomly derived from normal distributions and thus are variable from a dwelling to another. This additional feature means that the correlations driving the window openings and closures are different between households, so that they follow a different window opening behavior, which is closer to what is observed in reality. A similar approach was used by Haldi et al., but their methodology did not study the interrelation between the regression coefficients as was done here. This paper also describes a window opening model that is developed from Canadian data, which is rare in the literature where most studies on window behavior are from Europe and East Asia. Since occupant behavior is known to be highly related to culture, it is important to have access to data from different regions. Finally, another beneficial feature of the model is that there are only two predictors that are used, which makes it easy to apply to other studies.

Other variables, such as CO_2 levels or the zone where a window is located (living room, bedroom, kitchen…), could have improved the accuracy of the window opening model. With data coming from eight dwellings, it is impossible to quantify the full range of possible window opening behaviors. Although the model is able to cover all behaviors observed in the case study building, further studies would be needed to assess whether the model underestimates or overestimates the dwelling-to-dwelling variability of window opening in other buildings. In particular, it would be interesting to measure window openings in buildings with different architectural features (e.g., WWR, etc.) to evaluate how these affect occupant behavior.

Acknowledgements

This work is part of the research program of the Natural Sciences and Engineering Research Council of Canada (NSERC) Industrial Research Chair on Eco-responsible wood construction(CIRCERB) through programs IRC (IRCPJ 461745-12) and CRD (RDCPJ 445200-12). Authors are also grateful to the industrial partners of the NSERC Industrial Research Chair on Eco-responsible wood construction (CIRCERB). The first author is also supported by the Fonds de recherché du Québec – Nature et technologies

(FRQNT). The authors are also grateful to the Société d'Habitation du Québec (SHQ) and the Office municipal d'habitation de Québec (OMHQ) for providing the data for the residential building studied in this work.

References

[1] E. Diaz Lozano Patino and J. A. Siegel, "Indoor environmental quality in social housing: A literature review," *Build. Environ.*, vol. 131, pp. 231–241, Mar. 2018.

[2] D. Vakalis, M. Touchie, E. Tzekova, H. L. MacLean, and J. A. Siegel, "Indoor environmental quality perceptions of social housing residents," *Build. Environ.*, vol. 150, pp. 135–143, Mar. 2019.

[3] V. Fabi, R. V. Andersen, S. Corgnati, and B. W. Olesen, "Occupants' window opening behaviour: A literature review of factors influencing occupant behaviour and models," *Build. Environ.*, vol. 58, pp. 188–198, Dec. 2012.

[4] M. J. Sorgato, A. P. Melo, and R. Lamberts, "The effect of window opening ventilation control on residential building energy consumption," *Energy Build.*, vol. 133, pp. 1–13, Dec. 2016.

[5] R. Markovic, E. Grintal, D. Wölki, J. Frisch, and C. van Treeck, "Window opening model using deep learning methods," *Build. Environ.*, vol. 145, pp. 319–329, Nov. 2018.

[6] S. Pan, Y. Xiong, Y. Han, X. Zhang, L. Xia, S. Wei, J. Wu and M. Han, "A study on influential factors of occupant window-opening behavior in an office building in China," *Build. Environ.*, vol. 133, pp. 41–50, Apr. 2018.

[7] Y. Wei, H. Yu, S. Pan, L. Xia, J. Xie, X. Wang, J. Wu, W. Zhang, Q. Li "Comparison of different window behavior modeling approaches during transition season in Beijing, China," *Build. Environ.*, vol. 157, pp. 1–15, Jun. 2019.

[8] S. Pan, Y. Han, S. Wei, Y. Wei, L. Xia, L. Xie, X. Kong and W. Yu, "A model based on Gauss Distribution for predicting window behavior in building," *Build. Environ.*, vol. 149, pp. 210–219, Feb. 2019.

[9] R. Andersen, B. Olesen, and J. Toftum, "Modelling window opening behaviour in Danish dwellings," presented at the Proceedings of Indoor Air 2011: the 12th International Conference on indoor air quality and climate, Austin, Texas, 2011.

[10] V. Fabi, R. K. Andersen, and S. Corgnati, "Verification of stochastic behavioural models of occupants' interactions with windows in residential buildings," *Build. Environ.*, vol. 94, Part 1, pp. 371–383, Dec. 2015.

[11] V. M. Barthelmes, Y. Heo, V. Fabi, and S. P. Corgnati, "Exploration of the Bayesian Network framework for modelling window control behaviour," *Build. Environ.*, vol. 126, pp. 318–330, Dec. 2017.

[12] D. Calì, T. Osterhage, R. Streblow, and D. Müller, "Energy performance gap in refurbished German dwellings: Lesson learned from a field test," *Energy Build.*, vol. 127, pp. 1146–1158, Sep. 2016.

[13] R. V. Jones, A. Fuertes, E. Gregori, and A. Giretti, "Stochastic behavioural models of occupants' main bedroom window operation for UK residential buildings," *Build. Environ.*, vol. 118, pp. 144–158, Jun. 2017.

[14] M. Yao and B. Zhao, "Factors affecting occupants' interactions with windows in residential buildings in Beijing, China," *Procedia Eng.*, vol. 205, pp. 3428–3434, Jan. 2017.

[15] F. Stazi, F. Naspi, and M. D'Orazio, "Modelling window status in school classrooms. Results from a case study in Italy," *Build. Environ.*, vol. 111, pp. 24–32, Jan. 2017.
[16] B. Bordass, R. Cohen, and J. Field, "Energy performance of non-domestic buildings: closing the credibility gap," presented at the Building performance congress, 2004.
[17] R. V. Jones, A. Fuertes, and P. de Wilde, "The gap between simulated and measured energy performance: a case study across six identical new-build flats in the UK," presented at the 14th Conference of International Building Performance Simulation Association, Hyderabad, India, 2015.
[18] P. de Wilde, "The gap between predicted and measured energy performance of buildings: A framework for investigation," *Autom. Constr.*, vol. 41, pp. 40–49, May 2014.
[19] Y. Zhang, X. Bai, F. P. Mills, and J. C. V. Pezzey, "Rethinking the role of occupant behavior in building energy performance: A review," *Energy Build.*, vol. 172, pp. 279–294, Aug. 2018.
[20] V. Fabi, R. V. Andersen, S. P. Corgnati, and B. W. Olesen, "A methodology for modelling energy-related human behaviour: Application to window opening behaviour in residential buildings," *Build. Simul.*, vol. 6, no. 4, pp. 415–427, Dec. 2013.
[21] W. F. van Raaij and T. M. M. Verhallen, "Patterns of residential energy behavior," *J. Econ. Psychol.*, vol. 4, no. 1, pp. 85–106, Oct. 1983.
[22] V. Haines and V. Mitchell, "A persona-based approach to domestic energy retrofit," *Build. Res. Inf.*, vol. 42, no. 4, pp. 462–476, Jul. 2014.
[23] S. D'Oca and T. Hong, "A data-mining approach to discover patterns of window opening and closing behavior in offices," *Build. Environ.*, vol. 82, pp. 726–739, Dec. 2014.
[24] S. T. Moghadam, F. Soncini, V. Fabi, and S. Corgnati, "Simulating Window Behaviour of Passive and Active Users," *Energy Procedia*, vol. 78, pp. 621–626, Nov. 2015.
[25] F. Haldi, D. Calì, R. K. Andersen, M. Wesseling, and D. Müller, "Modelling diversity in building occupant behaviour: a novel statistical approach," *J. Build. Perform. Simul.*, vol. 10, no. 5–6, pp. 527–544, Nov. 2017.
[26] F. Haldi and D. Robinson, "Interactions with window openings by office occupants," *Build. Environ.*, vol. 44, no. 12, pp. 2378–2395, Dec. 2009.
[27] D. Calì, R. K. Andersen, D. Müller, and B. W. Olesen, "Analysis of occupants' behavior related to the use of windows in German households," *Build. Environ.*, vol. 103, pp. 54–69, Jul. 2016.
[28] H. B. Rijal, P. Tuohy, M. A. Humphreys, J. F. Nicol, A. Samuel, and J. Clarke, "Using results from field surveys to predict the effect of open windows on thermal comfort and energy use in buildings," *Energy Build.*, vol. 39, no. 7, pp. 823–836, Jul. 2007.
[29] F. Naspi, M. Arnesano, L. Zampetti, F. Stazi, G. M. Revel, and M. D'Orazio, "Experimental study on occupants' interaction with windows and lights in Mediterranean offices during the non-heating season," *Build. Environ.*, vol. 127, pp. 221–238, Jan. 2018.
[30] Z. Deme Belafi, F. Naspi, M. Arnesano, A. Reith, and G. M. Revel, "Investigation on window opening and closing behavior in schools through measurements and surveys: A case study in Budapest," *Build. Environ.*, vol. 143, pp. 523–531, Oct. 2018.
[31] R. Fritsch, A. Kohler, M. Nygård-Ferguson, and J.-L. Scartezzini, "A stochastic model of user behaviour regarding ventilation," *Build. Environ.*, vol. 25, no. 2, pp. 173–181, Jan. 1990.
[32] Y. Zhang and P. Barrett, "Factors influencing the occupants' window opening behaviour in a naturally ventilated office building," *Build. Environ.*, vol. 50, pp. 125–134, Apr. 2012.
[33] S. Wei, "Preference-based modelling and prediction of occupants window behaviour in non-air-conditioned office buildings," Thesis, © Shen Wei, 2013.
[34] N. Li, J. Li, R. Fan, and H. Jia, "Probability of occupant operation of windows during transition seasons in office buildings," *Renew. Energy*, vol. 73, pp. 84–91, Jan. 2015.

[35] D. Yan, W. O'Brien, T. Hong, X. Feng, H.B. Gunay, F. Tahmasebi, A. Mahdavi, "Occupant behavior modeling for building performance simulation: Current state and future challenges," *Energy Build.*, vol. 107, pp. 264–278, Nov. 2015.
[36] D. Bourgeois, "Detailed occupancy prediction, occupancy-sensing control and advanced behavioural modelling within whole-building energy simulation," Université Laval, Québec, 2005.
[37] J. Rouleau, A. P. Ramallo-González, and L. Gosselin, "Towards a comprehensive tool to model occupant behaviour for dwellings that combines domestic hot water with active occupancy," presented at the 14th Conference of International Building Performance Simulation Association, San Francisco, USA, 2017.
[38] J. Rouleau, A. P. Ramallo-González, L. Gosselin, P. Blanchet, and S. Natarajan, "A unified probabilistic model for predicting occupancy, domestic hot water use and electricity use in residential buildings," *Energy Build.*, vol. 202, p. 109375, Nov. 2019.