Analysis of the occupants’ behavior related to natural ventilation

Davide Calì¹, Rune Korsholm Andersen², Tanja Osterhage¹, Mark Wesseling¹, Dirk Müller¹, Bjarne Olesen²

¹Institute for Energy Efficient Buildings and Indoor Climate, E.ON Energy Research Center, RWTH Aachen University, Mathieustr. 10, 52074 Aachen, Germany
²International Centre for Indoor Environment and Energy, Department of Civil Engineering, Technical University of Denmark, Nils Koppels Allé, 2800, Kgs. Lyngby, Denmark

¹dcali@eonerc.rwth-aachen.de
²rva@byg.dtu.dk
¹tosterhage@eonerc.rwth-aachen.de
¹mwesseling@eonerc.rwth-aachen.de
¹dmueller@eonerc.rwth-aachen.de
²bwo@byg.dtu.dk

Abstract

The real energy performance of buildings depends both upon deterministic aspects (building’s physics and engineering systems) and probabilistic aspects such as weather and occupant behavior. Occupant behavior is usually not directly considered when calculating the expected energy performance of buildings. In fact, field test studies all over the world have shown discrepancies between expectation and real energy performances of buildings. This gap could be bridged, by embedding stochastic occupants’ behavior models within buildings’ energy performances simulation software. Within this work, an established method to analyze the probability of a state change of the windows, based on logistic regression, was applied to monitored data (measured each minute) from two refurbished residential buildings. The weather as well as the five rooms of each of the 60 apartments located in the buildings were monitored in terms of indoor environmental quality and window operation for four years. The aim of this work is the investigation of the drivers leading occupants to open and close windows. The evaluation of the 300 windows showed: the two most common drivers leading to the opening action were the time of the day and the carbon dioxide concentration in the room. The two most common drivers leading to the closing action were: the daily average outdoor temperature, and the time of the day.

Occupant behavior, Logistic regression, Natural ventilation, Buildings’ energy performance, Case study.

1. Introduction

Buildings are responsible for 40% of the total primary energy consumption in the European Union [1]. The ambitious goal of the EU on the reduction of the primary energy consumption (Energy Roadmap 2050) can only be reached, retrofitting the existing building stock. However field test studies [2–7] show higher observed consumption than expected. Reasons for this discrepancy are technical issues and the
missing integration of realistic occupant behavior models, when calculating expected energy figures. Understanding occupant behavior is the first step towards a realistic model of occupant behavior. Occupants influence the energy performance of buildings in many ways, e.g. changing set points of the heating system, modifying the position of sun-blinds, opening and closing windows.

This work focuses on the evaluation of the drivers [8] leading occupants to open and close windows. In particular, this work aims to analyze the drivers related to thermal comfort (room air temperature, room air relative humidity) indoor air quality (carbon dioxide concentration) and weather (Outdoor temperature, outdoor relative humidity and wind speed).

2. Description of the field test

Measurements of the following variables were monitored every minute in each room of 60 apartments in two refurbished buildings (Figure 1):

1. Air temperature [°C]
2. Relative humidity [%]
3. CO₂ concentration [ppm]
4. Volatile organic compounds (VOC) [-],
5. Light on the ceiling [Lux],
6. Infrared / visible light ratio [-],
7. Window opening position (open/closed).

The buildings have 60 geometrically identical apartments (the floor space of the apartments is shown in Figure 2). Each entrance of the building has an own retrofit layout. Within this work, the following nomenclature was used: “B” for building, e.g. “B2” refers to the group of the 30 apartments in building 2, E for entrance, e.g. B2E1 indicated the 10 apartments located in building 2 and accessible through entrance 1. Various engineering system components, building’s insulation materials and windows were selected and combined, to generate seven different retrofit layouts for three buildings (only building 2 and building 3 are considered in this evaluation). Building 2 is connected to a district heating network, while building 3 is heated through different
types of heat pumps (HP). Depending upon the entrance, radiators (Rad), ceiling heating (CH), floor heating (FH) and ventilation heating (VH) are installed to deliver the heating energy to the rooms. The six retrofit layouts of building 2 and building 3 are schematically described in Table 1. More information about the buildings and the retrofit layout can be found in [10,11].

![Figure 2. Floor space of the apartments.](image)

Table 1. Description of each retrofit layout in each of the 3 entrances in the two buildings: The HVAC system is supplied by either district heating (DH) or heat pump (HP). The ventilation is: exhaust air ventilation (EAV) or ventilation with heat recovery (HR). The apartments are heated through radiators (Rad), floor heating (FH), warm air heating (VH) or ceiling heating (CH). Domestic hot water is produced through central heat exchanger (HX) or fresh water heat exchanger stations (FWHX).

<table>
<thead>
<tr>
<th>Insulation</th>
<th>Windows U-Value</th>
<th>HVAC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>B2E1</strong></td>
<td>16 cm 0.021 W/(m²K)</td>
<td>1.3 W/(m²K)</td>
</tr>
<tr>
<td><strong>B2E2</strong></td>
<td>16 cm 0.021 W/(m²K)</td>
<td>0.8 W/(m²K)</td>
</tr>
<tr>
<td><strong>B2E3</strong></td>
<td>16 cm 0.021 W/(m²K)</td>
<td>1.3 W/(m²K)</td>
</tr>
<tr>
<td><strong>B3E1</strong></td>
<td>Vacuum: 4 cm 0.008 W/(m²K)</td>
<td>0.8 W/(m²K)</td>
</tr>
<tr>
<td><strong>B3E2</strong></td>
<td>Vacuum: 4 cm 0.008 W/(m²K)</td>
<td>0.8 W/(m²K)</td>
</tr>
<tr>
<td><strong>B3E3</strong></td>
<td>Vacuum: 4 cm 0.008 W/(m²K)</td>
<td>1.3 W/(m²K)</td>
</tr>
</tbody>
</table>
To evaluate the energy performances of the refurbished buildings and the occupants’ behavior, a comprehensive high time resolution monitoring system was designed and installed ([12]).

All variables were monitored through a room monitoring unit developed at the University of Applied Science Karlsruhe. The monitoring started during 2010 (The buildings have been completely occupied since spring 2011), the measurements were collected each 60 s, and the data was stored in HDF5 files.

3. Method

As an established method to analyze and model binary dependent variables (such as the state of a window, closed or open, or the change of state of a window), logistic regression analyses (LRA) was chosen to evaluate the occupant behavior related to the changes of state of windows. The here explained method was successfully used in [9].

LRA is based on the logistic function as expressed in (1), where \( p(x) \) expresses the probability function for a certain event (e.g. a window changes its state), and, by definition, \( P(x) \in [0,1] \, \forall \, x \). Equation (1) can be rewritten as in (2).

\[
p = \frac{1}{1+\exp(\alpha + \beta x)} \tag{1}
\]

\[
\ln \left( \frac{p}{1-p} \right) = \alpha + \beta x \tag{2}
\]

Where: \( P(x) \) (or simply \( p \)) is the probability function, \( \alpha \) is the intercept, \( \beta \) is a coefficient, \( x \) is the explanatory variable.

Equation (2) describes the probability of an event depending on one explanatory variable, and is therefore used for simple linear regression analysis. For regression analysis with “n” explanatory variables, the probability function \( p \) can be expressed as in (3). Finally, (4) includes the interaction terms, as suggested in [9]:

\[
\ln \left( \frac{p}{1-p} \right) = \alpha + \beta_0 x_0 + \beta_1 x_1 + \cdots + \beta_n x_n \tag{3}
\]

\[
\ln \left( \frac{p}{1-p} \right) = \alpha + \beta_0 x_0 + \beta_1 x_1 + \cdots + \beta_n x_n + \gamma_{1,2} x_1 x_2 + \cdots + \gamma_{1,n} x_1 x_n + \cdots + \gamma_{n-1,n} x_{n-1} x_n \tag{4}
\]

To evaluate the data, the LRA was executed on each window singularly. In addition, the "forward and backwards" selection of the variables for the regression models was executed, based on the Akaike information criterion (AIC), as already done e.g. in [9]. In this way, the selection of a "best model", containing only the most important explanatory variables (the variables which have a consistent impact on the probability function), is possible.

In practice, the process for the selection of the best model with \( n \) explanatory variables is described as follows:

1. each variable was fitted by the regression model (in a single-variable model), and the AIC calculated for each fit;
2. the variable with the lowest AIC was selected, and the model was fitted n-1 times with the selected variable and each of the n-1 remaining variables;

3. the model based on two variables with the lowest AIC was selected and the AIC of this model compared to the best single-variable model (the single-variable model with the lowest AIC); Then:
   a. If the new model (two-variables model) had a consistently lower AIC, the process went further to step 4,
   b. otherwise the single-variable model was selected;

4. The yet excluded n-2 variables were used to fit the model together with the two variables of the “two variables model” with the lowest AIC, in a “three variables model” (this is the "forward selection"). Further, from each of the three-variables models, three two-variables models, obtained by dropping each of the variables recursively, were fitted (this is the "backward selection"). Then:
   a. If none of the three-variable or "new generated" two-variable models had a consistently lower AIC than the two-variables model with the lowest AIC from step 3, the model with the lowest AIC from step 3 was selected,
   b. Otherwise, the process went further with the same criteria, up to n-variables models.

Further, a k-fold cross validation (with K=10) was executed on the top of the described selection process: Therefore, each data sample was partitioned in ten sub-samples. Nine sub-samples were used for the training of the models, while one sub-sample was used to test the model. This was realized by using the measured input variables of the 10th subsample (the one which was not used for the training) as input to the model, and comparing the model output with the monitored window position. The operation was executed 10 times: each subsample was used once as a test subsample, while the remaining nine sub-samples were used as training samples.

4. **Evaluation of the drivers**

The results of the regression analysis applied at room (window) level are illustrated in this section. The method explained in the previous section was applied to the 300 windows located in the buildings. The only categorical variable which was used is “Time range”, which distinguishes in low, medium and high probability of a state change of a window, and is grouped as follows:

1. Night: Low probability of action;
2. Morning: High probability of action;
3. Rest of the day: Medium probability of action;

In addition, the following continuous variables, measured each minute, were used:

1. Room air temperature;
2. Room carbon dioxide concentration (CO₂), transformed through the reciprocal function to obtain a more suitable distribution for the use in the regression analysis;
3. Room relative humidity;
4. Daily average outdoor temperature;
5. Wind speed;
6. Outdoor relative humidity.

By the selection of the variables used for the LRA, variables which could correlate were avoided. For example, the carbon dioxide concentration was preferred to the volatile organic compounds (VOC) as indicator for the air quality (this choice was justified by the fact that VOC could not be used to discern between good and bad odors, while the CO\textsubscript{2} concentration was a good indicator for human bioeffluents). The daily average outdoor temperature was preferred to the instant value of this, to prevent correlation issues with the indoor air temperature.

5. Results from the LRA

In this section, the results of all the windows are presented. The results are organized in graphics, showing the explanatory variables, selected with the procedure explained in section 3, and the number of LRA equations (models) using each of them. Figure 3 shows the explanatory variables selected for the opening of window. The only used categorical variable, "time range", was included in more than 70\% of the windows. Moreover, the interaction terms with the variable "time" were used by less than 10\% of the models: this means that the variable “time” influences commonly the intercept, but not the coefficients of the continuous explanatory variables. The most common continuous explanatory variable was the CO\textsubscript{2} concentration, present in more than 50\% of the models. Room temperature, relative humidity of the room, daily average outdoor temperature, and outdoor relative humidity were used by more than 40\% of the models. The wind speed turned out to be mostly irrelevant for the opening action. For some of the windows no drivers were found ("None" indicates that no explanatory variables were found for the fitting of the model).

Figure 4 and Figure 5 show the drivers which directly (positively) and inversely (negatively) influenced the probability of the action “window opening” respectively. For example, an increase in carbon dioxide concentration lead to an increase of the probability of window opening for more than 50\% of the models; however, contrary to this, the carbon dioxide concentration negatively influenced the probability of an opening action for approximately 6\% of the models; this is not necessarily a contradiction, and could be related to the occupancy patterns (which was not included in the models since presence of occupants was not monitored) and window opening behavior upon arrival.

The probability of opening windows increased with increasing indoor air temperatures (over 40 \% of the models), and by increasing room relative humidity and daily average outdoor temperature. The only remarkable explanatory variable which negatively influenced the opening of windows (over 30 \% of the windows) was the outdoor relative humidity.

Figure 6 shows the explanatory variables of the logistic regression models of closing action. Figure 7 and Figure 8 show the drivers which positively and negatively influenced the window opening action respectively. The daily average outdoor temperature was the most common driver (in almost 70 \% of the models); in particular, as it can be seen in Figure 7 the probability of closing windows increased with
decreasing daily average outdoor temperature. The variable time was present in more than 50% of the models.

Figure 3 Drivers for opening the window and the number of rooms/models with the driver used as explanatory variable.

Figure 4 Drivers for opening the window and number of rooms/models with a positive correlation between the variable and the probability of opening.

Figure 5 Drivers for opening the window and number of rooms/models with a negative correlation between the variable and the probability of opening.

Further, an increase in the carbon dioxide concentration was associated with increasing probability of the closing action. This may be an effect of correlations between carbon dioxide concentration and presence of occupants, since the presence of occupants is a necessary condition, for the window to be closed. In almost 40% of the models, a decrease of the room temperature corresponded to an increase of the probability of closing the window.
Drivers for closing the window and number of apartments/models with the driver used as explanatory variable.

Drivers for closing the window and number of rooms/models with a positive correlation between the variable and the probability of closing.

Drivers for closing the window and number of rooms/models with a negative correlation between the variable and the probability of closing.

**Conclusion and outlook**

The results from the LRA identified the drivers of opening and closing of windows for 300 windows located in 60 apartments, in two refurbished buildings. The most common drivers for the opening action of windows were: the time of the day (for more than 70% of the modeled windows), and the indoor carbon dioxide concentration (for over 50% of the modeled windows). The most common drivers for the closing action of windows were: the daily average outdoor temperature (for almost 70% of the modeled windows) and the time of the day (for more than 50% of the modeled windows). These results are in agreement with the tendency of the results published by Andersen et al. and Fabi et al. based on the evaluation of Danish dwellings [9,14]. These results could help researchers to model occupants’ window opening behavior.
In a further work based on the presented data, the dependency of the occupant behavior on the engineering system (or in general the retrofit layout) could be investigated. Beside the evaluation of the drivers, the LRA generated two models for each window, one for the opening action, one for the closing action. Those models could be used in dynamic building energy performance simulation software in order to integrate occupant behavior in the simulations. However, the number of models is very high, and there is not yet valid criterion for the choice of one model rather than another. For modeling purposes, the technique of the mixed effect modelling, as proposed in [13], seem to be a better way forward.

Acknowledgment

We gratefully acknowledge the financial support of BMWi (03ET1105A) and of E.ON New Build and Technology

References