Analysis of occupants' behavior related to the use of windows in German households

Cali, Davide; Andersen, Rune Korsholm; Mueller, Dirk; Olesen, Bjarne W.

Link back to DTU Orbit

1   **Analysis of occupants' behavior related to the use of windows in German households**

2   Davide Calì[1], Rune Korsholm Andersen[2], Dirk Müller[1], Bjarne Olesen[2]

3   Corresponding Author: Davide Calì, DCali@eonerc.rwth-aachen.de
4   [1]Institute for Energy Efficient Buildings and Indoor Climate, E.ON Energy Research Center, RWTH Aachen University,
5   Mathieustr. 10, 52074 Aachen, Germany
6   [2]International Centre for Indoor Environment and Energy, Department of Civil Engineering, Technical University of Denmark,
7   Nils Koppels Allé 402, 2800, Kgs. Lyngby, Denmark

8   **Abstract**

9   Real energy performances of buildings depend not only on deterministic aspects, such as building physics and HVAC systems,
10   but also on stochastic aspects such as weather and occupants' behavior. Typically, occupant behavior is not adequately
11   considered when calculating the expected performance. As a result, field test studies all over Europe have shown discrepancies
12   between real and expected energy performance of buildings. In order to bridge this gap, stochastic occupants' behavior models
13   could be embedded into building energy performance simulation software. In order to make such models, there is a need for a
14   better understanding of occupants' behavior and in particular the reasons of their adjustments of building controls such as
15   window opening, heating set points, etc. The purpose of this paper was to analyze window opening behavior in residential
16   buildings, investigate which drivers lead occupants to interact with windows and how these actions can be modeled. A method
17   to analyze the probability of a state change of the windows, based on logistic regression, was applied to monitored data
18   (measured each minute) from two refurbished demonstration buildings. The weather and the five rooms of the 60 apartments
19   located in the buildings were monitored in terms of air quality and thermal environment (presence of occupants was not
20   monitored) during four years.
21   The most common driver to open a window was the time of the day, followed by the carbon dioxide concentration. The most
22   common driver to close a window was the daily average outdoor temperature, followed by the time of the day.

1.   Introduction

26   During the last decades, a big effort has been made to build new energy efficient buildings and make energy retrofits to the
27   existing building stock. However, new or retrofitted buildings do not always perform as predicted: field test studies all over
28   Europe [1–14] show higher final energy consumption than predicted. The reasons for this discrepancy can be grouped into
29   engineering system issues (e.g. components not performing as expected or wrongly installed, wrong installation of the
30   insulation, bad insulation of thermal bridges, tightness, etc.), weather data as well as occupants' behavior issues (differences in
31   modeled and real behavior regarding the interaction with windows and engineering systems).
32   The real energy performances of buildings depend in general on deterministic aspects, like building physics and the HVAC
33   systems, as well as stochastic (probabilistic) aspects, such as weather and occupants' behavior. While a lot of effort has been
34   spent on deterministic modeling of the buildings as a whole, as well as on the generation of standard weather files (known as
35   test reference years) for many regions in the world, only in the last decade the importance of the dynamic interaction between
36   occupants and buildings has been recognized and researched. A reason for this can be connected to the fact that the relative
37   impact of the occupants' behavior on the building energy performance seems to increase when building standards require lower
38   energy use [15]. Including the stochastic aspects of occupants' behavior into building energy performance simulation software,
39   could hence lead to better predictions of the buildings' final energy consumption, especially for new low energy buildings or
40   energy retrofitted buildings.
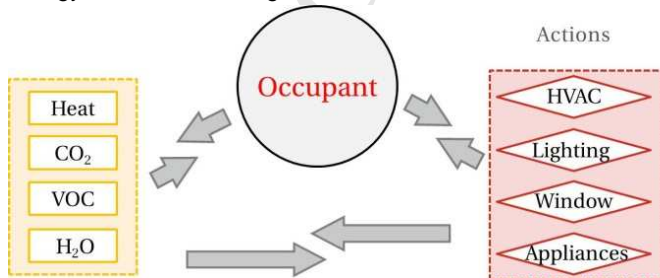
42   Figure 1. General schema of occupants' impact on buildings (previously introduced in [16])

43   In order to improve the energy performance forecasts, there is a need to focus on occupants' behavior, defined as *the set of*
44   *actions, executed by occupants, modifying the physical conditions of the built environment*. Occupants also influence the indoor
45   environment passively through their presence (e.g. through the production of heat, water vapor, carbon dioxide, and volatile
46   organic compounds); thus, a holistic model of the occupants' behavior should include:

1. periods of presence and absence of occupants
2. occupants' interaction with, among others, thermostats/valves, windows and sun-blinds (as schematized in Figure 1)

A consistent stochastic model of occupants' behavior in residential buildings based on a large database of observed data has not been developed yet, mostly due to lack of data. In fact, in studies where the presence of occupants was monitored (e.g. through the use of time-use survey) and modeled [17,18], no information about occupants' behavior was recorded. Vice versa, in studies focusing on the evaluation and modeling of occupants' behavior, based on monitoring data, reliable information about real occupancy profiles is missing. In this context, in 2014, the International Energy Agency started the "IEA-EBC Annex 66 - Definition and Simulation of Occupant Behavior in Buildings" [19]: this aims to bridge the gap between expected and observed energy performance of buildings, by modeling and integrating the occupants' behavior in buildings' simulation software.

The architecture of a holistic occupants' behavior model could consist of a core for the generation of presence patterns (e.g. based on Markov chain as in [17,20,21]), and, on top of this, of a variable number of add ons to reproduce human activities, (e.g. related to settings of the engineering system, change of state of windows, use of appliances, etc.). Many models have already been developed for reproducing occupants' activities: A plenary literature review on occupants' behavior models, up to 2013, is offered in [22]; a more recent review of state of the art models of occupants' behavior is offered in Yan et al. [23].

In the European Union (EU28), space heating [24], which is influenced by the ventilation rate, accounts for more than two thirds of residential buildings' energy consumption. As such, the energy consumption in residential buildings is under heavy influence of the occupants' window opening behavior and any attempt to realistically model indoor environment and energy consumption in residential buildings should include realistic models of the window operation. In order to construct such models, basic knowledge of the reasons for opening and closing windows, is needed.

In 2012 Fabi et al. [25] analyzed more than 70 papers to investigate why occupants opened and closed windows. In addition to the literature review, the authors define the concept of drivers: "Factors influencing occupant behavior, both external and individual, that could be named with the general term "Drivers", are the reasons leading to a reaction in the building occupant and suggesting him or her to act (they namely "drive" the occupant to an action)". From the analyzed literature review, Fabi et al. conclude that:

1. Window opening behavior has a strong impact on both the indoor environmental quality and the energy performances of buildings,
2. there is not a commonly agreed upon approach among scientists about the driving forces of occupants' window opening and closing behavior,
3. most of the existing studies analyze the position of windows instead of the transition from state "Open" to state "Closed" and vice versa, and this might be problematic since the status of the windows influences the indoor environment,
4. more efforts should be made to better understand and model occupants' behavior.

One of the first attempts to mathematically model the occupants' behavior related to natural ventilation was made by Fritsch et al. [26]. They use a Markov Chain process and their model is based on the measured window opening angles recorded each half hour for four windows in four office rooms. Each window is modeled with transition probability matrices with six states respectively (the opening angles are grouped into five opening classes plus a class for the closing status) and adjusted for four different outdoor temperature ranges. The model considers the windows to be closed at night and on weekends, and the generation of the window state profiles starts at the beginning of each working day (and ends therefore at the end of the day). Fritsch et al. [26] argue that the state of windows during night has, except for two cases, always been closed over the entire observed heating period. However, this consideration is not valid for residential buildings, as shown in many field test studies [25], [27].

With their work "Interactions with window openings by office occupants" [28], Haldi et al. set a milestone in the modeling of occupants' behavior related to window operation in office buildings. For seven years, the authors monitored temperature, window position (open/closed, no distinction between opening angles), and presence of occupants (with passive infra-red sensors) of 14 offices in a building, and the outdoor environmental conditions through a weather station located 8 km away from the building. They modeled the behavior following several approaches based on logistic probability distributions, Markov chains, continuous-time random processes, and hybrid combinations of those models. They concluded that "[s]upported by rigorous cross-validation, [they] have demonstrated the superiority of a discrete-time Markov process approach and its strong added value compared with existing models". Further, they wrote: "[f]actors related to indoor air quality (e.g. CO2 or pollutant concentration) should also be treated".

A further contribution to the occupants' modeling for natural ventilation is done by Andersen et al. within several works [27,29,30]. In the last paper, the authors introduce a model for the transition of the state of windows (open to close and vice-versa close to open) based on observations collected from 15 dwellings located in Denmark. Contrarily to previous window operation (or state) models introduced by 2013 and only based on thermal conditions, the authors also monitored indoor carbon dioxide concentration as an indicator for indoor air quality. The models - one for the opening and one for the closing of the windows - are based on the results of logistic regression analysis. The conclusion of the manuscript is that indoor CO_2 concentration and outdoor temperature are the two single most important variables (drivers) determining the window opening and closing probability respectively.

D'Oca and Hong [31] recently evaluated occupants' behavior related to natural ventilation in offices using logistic regression analysis and two different data-mining approaches, cluster analysis and association rules mining. They state that indoor air temperature, outdoor air temperature, time of the day (office arriving time and early morning) and occupancy presence are the

1 top drivers for the window opening operation; indoor air temperature, time of the day (office leaving time and evening),
2 occupancy presence and outdoor air temperature are the top drivers for the window closing operation.
3 From the literature review it is evident, that most analyses focused on office buildings, and that there is a lack of field test
4 studies related to residential buildings. Furthermore, several approaches were used to analyze and model occupants' behavior:
5 logistic regression analysis and mixed effect models seem to be proper instruments for evaluating the drivers, while Markov
6 chains are particularly indicated to model the time dependency of processes (e.g. opening and closing cycles).
7 In many of the mentioned papers [21, 24, 25, 26, 27, 28] there is agreement that factors related to indoor air quality could be a
8 driver for occupants' window operation. There is also agreement that factors related to thermal comfort should be included in
9 models of window opening behavior. However, these conclusions are based on data from a limited number of windows in a
10 limited number of apartments. The objective of this work was to investigate which drivers lead occupants to open or close
11 windows, based on one full year of measurements from 300 windows in 60 apartments.
12 The drivers analyzed were related to thermal comfort (room air temperature, room relative humidity), indoor air quality (carbon
13 dioxide concentration), weather (wind speed, outdoor temperature, outdoor relative humidity) and time of the day. Compared to
14 other studies on modeling the state change of windows, the data used within this work had three main advantages: the
15 availability of both data related to thermal comfort and to indoor air quality, the big number of geometrically identical apartments
16 (however with different engineering systems) and a ten times higher time resolution. This high time resolution allowed for a
17 precise study of the drivers for a window state change, without missing short opening/closing phases (more on this topic in
18 section 3.2)

2. Description of the applied methodology

20 Logistic regression is a well-established statistical method used to analyze and model binary dependent variables (such as the
21 state of a window, closed or open, or the change of state of a window). It is used in many scientific fields, from medicine and
22 economics to engineering. As illustrated in the introduction, the method has been successfully adopted to describe the
23 probability of opening and closing events of windows. The main objective of the statistical analyses was to understand the
24 drivers leading occupants' to act (change the state of windows), thus providing researchers a background about how occupants'
25 make decisions, correlated to thermal comfort, indoor air quality, weather and time of the day. In order to tackle this research
26 question, the logistic regression analysis with multiple explanatory variables was applied singularly to each monitored window.
27 In this way, for each window, it was possible to determine which variables had a major influence on the probability of a change
28 of window state, and which did not. Thus, the variables were classified depending on the number of times they appeared in the
29 models (the logistic regression equations). The variables used most frequently in the models were regarded as more important
30 in general for opening/closing actions.
31 Logistic regression [32] was used as analysis and modeling method. Logistic regression is based on the logistic function as
32 expressed in equation 1. $P(x)$ expresses the probability function for a certain event to happen (e.g. a window changes its state),
33 and, by definition, $P(x) \epsilon [0,1] \forall x$. Equation 1 can be rewritten as in equation 2.

34
$$p = \frac{1}{1+e^{-(\alpha+\beta x)}} \qquad (1)$$

35
$$ln\left(\frac{p}{1-p}\right) = \alpha + \beta x \qquad (2)$$

36 Where:
37   1. $P(x)$ (or simply p) is the probability function,
38   2. α is the intercept,
39   3. β is a coefficient,
40   4. x is the explanatory variable.
41
42 Equation 2 describes the probability of an event depending on one explanatory variable, and is therefore used for simple linear
43 regression analysis. For regression analysis with n explanatory variables, the probability function p can be expressed as in
44 equation 3.

45
$$ln\left(\frac{p}{1-p}\right) = \alpha + \beta_0 x_0 + \beta_1 x_1 + \cdots + \beta_n x_n \qquad (3)$$

46 Andersen et al. [27] suggest to include interaction terms into the probability function. They argue that "the probability might
47 depend differently on x_1 at one level of x_2 as compared to another level of x_2". For example, when inferring the probability of
48 opening (or closing) a window, it may happen that the coefficient β_i of the x_i explanatory variable at a certain time slot (e.g.
49 morning) differs from the coefficient β_i at a different time slot (e.g. night): in an extreme case, an increase of the temperature
50 might result in an increase of the probability of opening a window in the morning, and in a decrease of the probability of opening
51 a window in the evening. Based on their suggestion, in order to include the interaction terms $\gamma_{i,j}$, equation 3 can therefore be
52 written as equation 4. Only interaction terms between continuous and categorical variables were used.

53
$$ln\left(\frac{p}{1-p}\right) = \alpha + \beta_0 x_0 + \beta_1 x_1 + \cdots + \beta_n x_n + \gamma_{1,2} x_1 x_2 + \cdots + \gamma_{1,n} x_1 x_n + \cdots + \gamma_{n-1,n} x_{n-1} x_n \qquad (4)$$

54 The "forward and backwards" selection of the variables for the regression models was executed based on the Akaike
55 information criterion (AIC), as suggested by Schweiker and Shukuya [33]: This allowed for a selection of a "best model"
56 containing only the most important explanatory variables (variables that have a consistent impact on the probability function). In

practice, the process for the selection of the best model was implemented using the 'step' function with the 'glm' function in the statistical language "R", with *n* explanatory variables. It is described by the following steps:

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 is 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 chosen;
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 so-called "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 so-called "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 chosen,
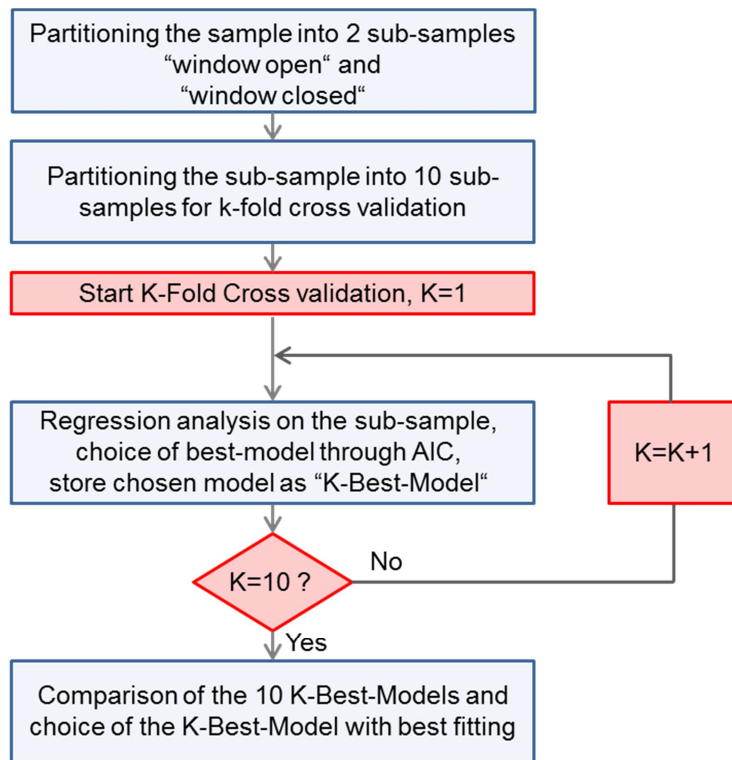   b. Otherwise, the process went further with the same criteria, up to n-variables models.

Figure 2. Flux diagram of the process aimed to generate and obtain the best fitting model

Further, a k-fold cross validation (with K=10) was executed on the top of the described selection process: Therefor, each data sample was partitioned into 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 done 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. The analysis and modeling activity was executed for the window state change "opening" and the window state change "closing" separately. By windows with more panels, p=0 corresponds to both panels closed, while p=1 corresponds to at least one panel open. In order to infer the probability of opening and closing (the "state change probability"), the data sample was partitioned into two sub-samples: sub-sample "window closed", to infer the probability that a window will be opened and sub-sample "window open", to infer the probability that a window will be closed. The complete modeling process is graphically described in the flux diagram of Figure 2. The analyses were executed in the language "R" using the packages "Companion to Applied Regression" (car) and the package "Mixed GAM Computation Vehicle with GCV/AIC/REML Smoothness Estimation" (mgcv), within the programming environment R Studio.

3. Description of the demonstration buildings and of the monitoring system

Measurements were conducted in three refurbished buildings (Figure 3) located in southern Germany. The buildings were built at the end of the nineteen-fifties, with relatively poor materials and unpretentious engineering systems. The three buildings, located next to each other, are geometrically identical. Each building has 3 entrances, and each entrance provides access to 10 apartments over 5 floors. The buildings have a total of 90 apartments. The apartments have the same floor plan. They consist of a kitchen, a bathroom, three rooms and a corridor. In this paper, 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. The buildings are 52 m wide and 10 m long. Before retrofit, the gables had 4 cm insulation material and the windows were double glazed; the roof, the floor to the cellar and the façade were not insulated. The apartments were heated through gas stoves installed in each living room. Domestic hot water was produced through gas flow heaters installed in each apartment.

Figure 3. One building façade before retrofit (a). One back façade after the retrofit (c). Floor space of the apartments, for one entrance with south-gable (b). Relative position and orientation of the buildings (d).

3.1. The Retrofit of the Buildings

Table 1. Description of each retrofit layout in each of the 3 entrances in the three 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).

	Insulation	Windows U-Val	HVAC	DHW
B1, 2160m², 30 apartments	14cm 0.035W/(m²K)	1.3 W/(m²K)	DH, EAV, Rad	Central HX
B2E1, 720m² 10 Apartments	16cm 0.021W/(m²K)	1.3 W/(m²K)	DH, EAV, window frame HR, Rad	Apartment FWHX
B2E2, 720m² 10 Apartments	16cm 0.021W/(m²K)	0.8 W/(m²K)	DH, EAV, Rad	Central HX
B2E3 720m² 10 Apartments	16cm 0.021W/(m²K)	1.3 W/(m²K)	DH, EAV, FH	Central HX
B3E1, 720m² 10 Apartments	Vacuum: 4cm 0.008W/(m²K) 4cm 0.021 W/(m²K)	0.8 W/(m²K)	CO_2-Probe HP, HR central ventilation, FH	Apartment FWHX
B3E2, 720m² 10 Apartments	Vacuum: 4cm 0.008W/(m²K) 4cm 0.021 W/(m²K)	0.8 W/(m²K)	CO_2-Probe master HP + slave HP, apart. HR, VH	Apartment FWHX
B3E3 720m² 10 Apartments	Vacuum: 4cm 0.008W/(m²K) 4cm 0.021 W/(m²K)	1.3 W/(m²K)	Air HP + exhaust air HP, CH	Apartment FWHX

1 Between 2008 and 2010 the buildings were retrofitted. The retrofit layouts were designed together with the municipal society
2 owning the buildings. Various engineering system components and building insulation materials as well as new windows (Figure
3 4) were selected and combined, to generate seven different retrofit layouts: in this way, the retrofit layouts can be compared,
4 and useful knowledge about optimal retrofitting can be gained. Building 1 and building 2 are connected to a district heating
5 network, while building 3 is heated through different types of heat pumps (HP). Depending on the entrance, radiators (Rad),
6 ceiling heating (CH), floor heating (FH) and ventilation heating (VH) were installed to deliver the heating energy to the indoor
7 heated spaces. Standard water heaters and low temperature peripheral domestic hot water (DHW) heaters were installed. In the
8 peripheral solution the DHW is generated in each apartment through so called fresh water heat exchanger stations (FWHX). The
9 seven retrofit layouts are schematically described in Table 1. More information about the buildings and the retrofit layout can be
10 found in [34–38].
11 It should be noted that there are five windows per apartment; two of them (in kitchen and bathroom) have one opening panel.
12 The two windows in the bedrooms and in the children rooms have three opening panels: for those windows, the panels on the
13 left and on the right can be opened independently and were monitored; the central panel can only be open (but not tilted) when
14 the right panel of the window is also completely open. The window (balcony-door) in the living room has three panels: the one on
15 the left can be open be open irrespective of the state of the other two panels and was monitored, the panel in the middle can
16 only be open (but not tilted) only if the left panel is open, and was not monitored. The panel on the right is fixed.

17

18

19 Figure 4. Pictures of the windows and of the façade after retrofit.

20 **3.2. The Monitoring System and the Observed Data**

21 To evaluate the energy performances of the refurbished buildings and the occupants' behavior, a comprehensive high time
22 resolution monitoring system was designed and installed (by the University of Applied Science Karlsruhe). The monitoring
23 system collected energy flows in the heat generation, the heat storage and the distribution system of both domestic hot water
24 (DHW) and heating (HE). In building 2 and building 3, the following variables were monitored every minute in each room of the
25 60 apartments:
26   1. Air temperature [°C]
27   2. Relative humidity [%]
28   3. CO_2 concentration [ppm]
29   4. Volatile organic compounds (VOC) [-],

1 5. Light on the ceiling [Lux],
2 6. Infrared/visible light ratio (to recognize the light source) [-],
3 7. Window opening sensors (open/closed).

4 All variables (except for window position) were monitored through room monitoring units (RMUs) as shown in Figure 5. Further
5 information about the monitoring system can be found in [39]. In building 1, only 7 apartments were equipped with RMUs. The
6 evaluation presented in this manuscript was therefore related to B2 and B3. The windows' positions were monitored through
7 wired reed switches installed in the windows' frames and permanent magnets installed in the frame of the windows' moveable
8 panels. The monitoring started in 2010 (The buildings were completely occupied since spring 2011), the measurements were
9 collected each 60 s, and the data was stored in HDF5 files. Due to crashes of the monitoring system, not all data was always
10 collected correctly. For the year 2012 (chosen for the here proposed analysis), at least 90 % of the data was correctly collected;
11 the crashes of the monitoring system were spread out over the entire monitoring period.
12 Table 2 shows the specification of the sensors adopted to collect the variables used within the LRA (the choice of those
13 variables is explained in the next section).
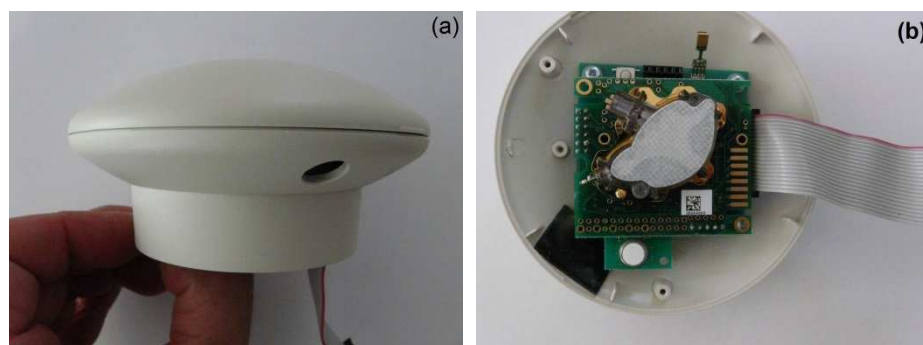
14

15 Figure 5. (a) RMU ready for plugging. The border on the bottom of the RMU is fixed to the wall. The circular opening permits the
16 light penetration. (b) Upper part of an open RMU: VOC sensor (bottom), CO_2 sensor (in the center), temperature and relative
17 humidity sensor (top right outside the board), and illuminance sensor (behind the black tape, bottom left).
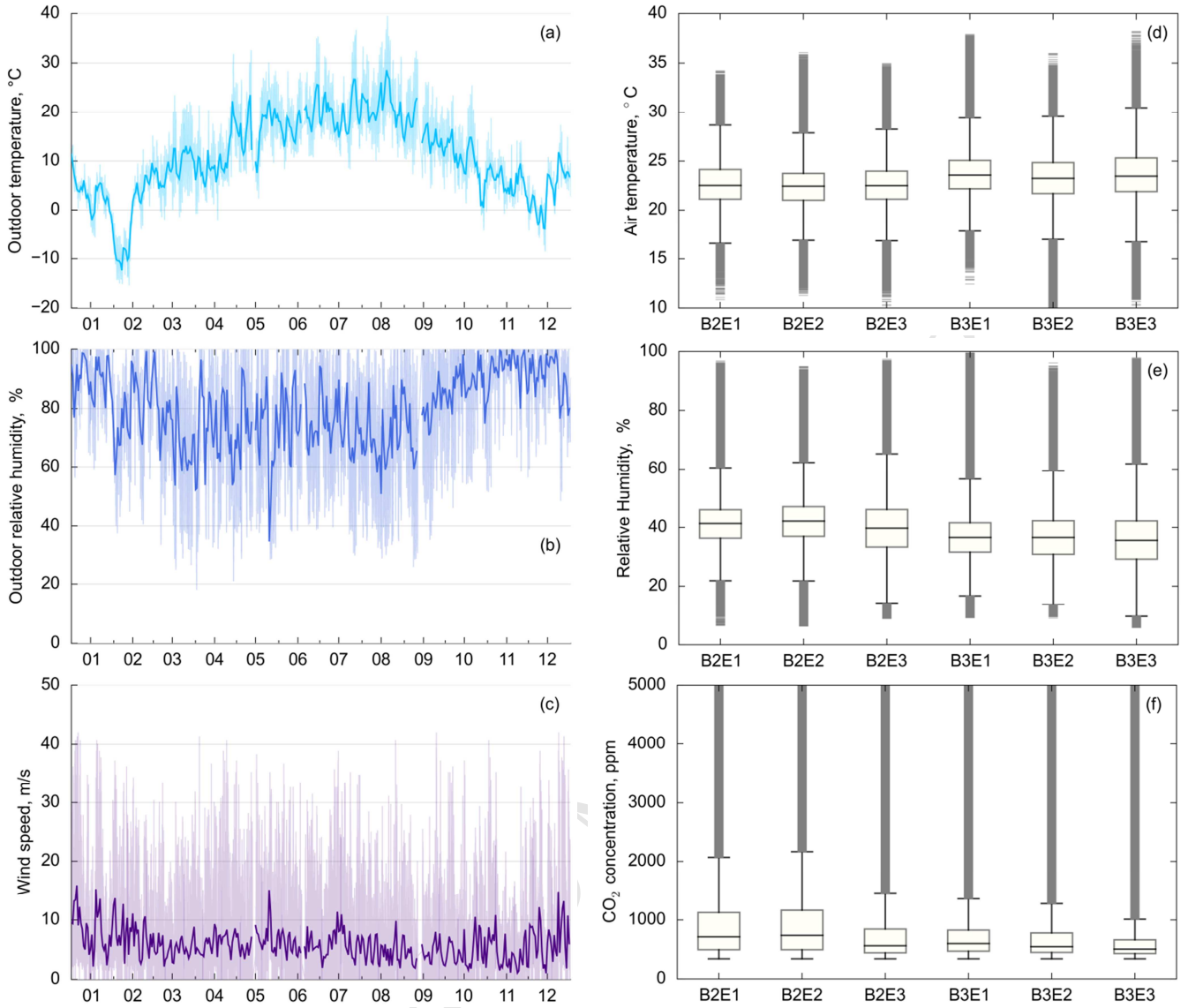
Figure 6. Monitored data (2012). Outdoor measurements (minutes and daily average): outdoor air temperature (a), outdoor relative humidity (b) and wind speed (c). The numbers on the x-axis indicate the months of 2012. Interruptions in the graphs are caused by missing values. Indoor measurements: box plots of the values of air temperature (d), relative humidity (e) and CO_2 concentration, measured in the rooms of each entrance. The boxes contain 50% of the observations; the whiskers are up to 1.5 the size of the boxes, the rest of the data is included in the flyers.

Table 2. Specifications of the installed sensors

	Metric	Sensor	Range	Accuracy
Room	CO_2	SenseAir CO_2EngineTM(K30)	0-5000 ppm	± 20 ppm
	Temperature	Sensirion SHT75	-40 °C - + 124 °C	± 0.3 °C
	Relative humidity	Sensirion SHT75	0% - 100%	± 4%
Weather	Temperature	MWS 5MV / MWS8	-40 °C - + 60 °C	± 0.3 °C
	Relative humidity	MWS 5MV / MWS8	10% - 100%	± 2%
	Wind speed	MWS 5MV / MWS8	0 - 42 m/s	± 5 m/s

A description of the measured data used for the logistic regression analysis is offered in Figure 6. The subfigures (a) to (c) provide information related to the observed weather conditions in 2012. February was particularly cold, with daily average temperature down to -12 °C; On average, the outdoor temperature in 2012 was 11 °C, against the typical value for this region (based on historical data and used e.g. for the generation of the energy labels of buildings) equal to 8.9 °C. The values related to the indoor conditions are plotted in Figure 6 (d) to (f): each box shows the data collected in the 50 rooms (10 apartments) of each building entrance (hence, each box represents over 25 million data points, which explains the high density of the flyers in

1 the figures). In general, building 3 had higher temperatures and lower relative humidity than building 2. The CO_2 concentration
2 was lower in B2E3 and B3, than in B2E1 and B2E2; however, it was not possible to find any dependency between the air quality
3 level and the different retrofit solutions.
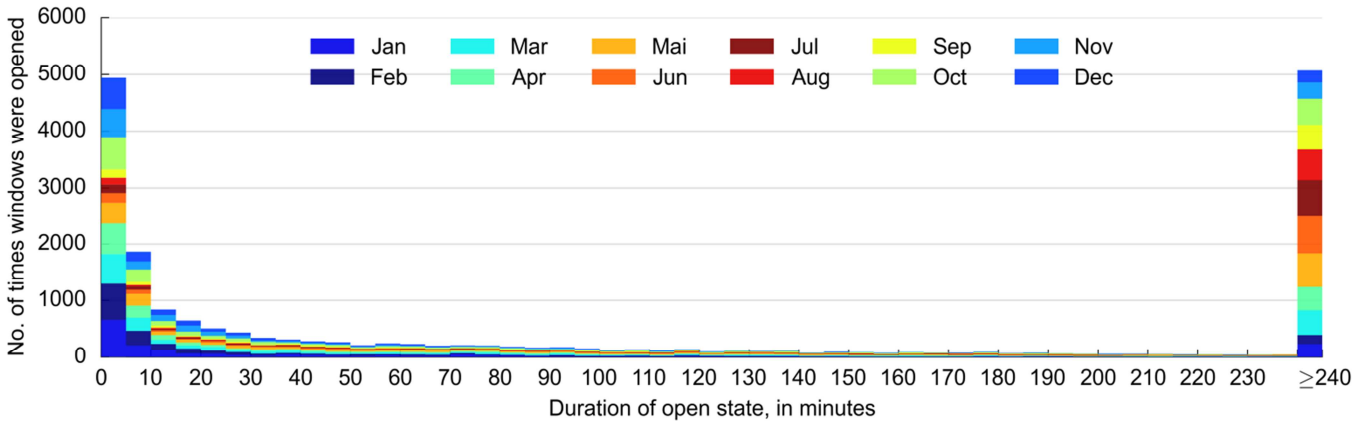
5 Figure 7. Number of times windows were opened, and duration of the state open, by month, for B3E2.

6 As already mentioned in the introduction, it is important to monitor the opening cycles of the windows event-based, or within a
7 short time interval. Figure 7 shows the duration of time with open windows. In colder months, the duration was much shorter
8 than in warmer periods.

9 **4. Evaluation and modeling of the occupants' behavior**

10 In this section, the results of the regression analysis applied at room (window) level are illustrated. The method explained in
11 section 2 was applied to the 300 windows located in buildings B2 and B3. The only categorical variable used was "Time range",
12 which distinguishes between low, medium and high probability of a state change of a window, and is grouped as follows:

       13   1. Night, low probability of action: 7 hours, between 11:00 p.m. and 5:59 a.m.
14   2. Morning, high probability of action: 3 hours, between 7:00 a.m. and 9:59 a.m.
15   3. Rest of the day, medium of action: 14 hours, between 6:00 a.m. and 6:59 a.m. and between 10:59 a.m. and 22:59
16   p.m.

17 The following continuous variables, measured each minute, were used:

  18   1. Room air temperature [°C];
19   2. Room carbon dioxide concentration (CO_2) [ppm], transformed through the reciprocal function to obtain a more suitable
20   distribution for the use in the regression analysis [ppm^{-1}];
21   3. Room relative humidity (RRH) [%];
22   4. Daily average outdoor temperature (DAT) [°C];
23   5. Wind speed (WS) [m/s];
24   6. Outdoor relative humidity (ORH) [%];

25 The categorical variable "Time range" was used in order to take into account the differences in behavior, depending on the time
26 of the day. The categories were chosen based on the observations of the window opening activity in Building 2 entrance 1.
27 Those are illustrated in Figure 8.
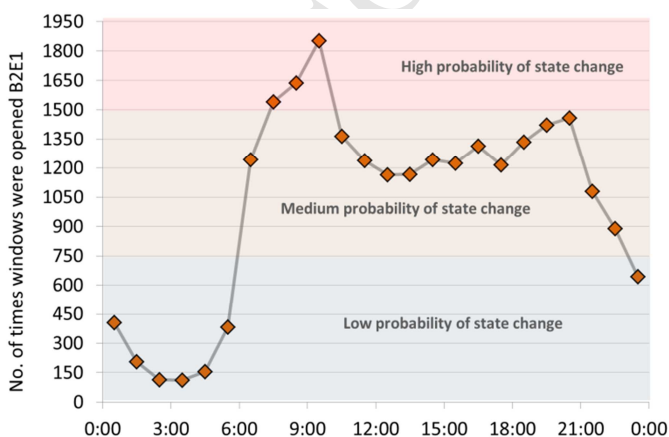
29 Figure 8. Observed no. of times windows were opened in B2E1, depending on the time of the day, and thresholds for the choice
30 of the action probability levels.

1   In the selection of the variables to be used for the LRA, correlating variables were avoided. The carbon dioxide concentration
2   was chosen instead of the volatile organic compounds (VOC) as the indicator for air quality. This choice is justified by the fact
3   that it is not possible to distinguish between good and bad odors through the VOC (some VOCs have pleasant odors e.g. used
4   in cleaning products and perfumes while others have unpleasant odors), while the CO_2 concentration is a good indicator for
5   human bioeffluents [40]: hence, a high concentration of CO2 is an indicator for bad air quality. In addition, the continuous
6   variable "global radiation" was not included in the evaluation since it correlated to the variable "room temperature". The variable
7   "outdoor temperature" was taken into account, but only as daily average outdoor temperature to prevent correlation issues with
8   the time and with the outdoor relative humidity.
9   As previously discussed in the introduction, the presence of occupants is a prerequisite for occupants' actions. A method to
10   reveal occupants' presence using the carbon dioxide concentration in a room, and the mechanical ventilation patterns was
11   developed and introduced in [16]. However, since the door position of the rooms as well as the occupants' settings regarding
12   the ventilation system were unknown, the uncertainty on the results of the proposed algorithm was too big, hence the algorithm
13   was not used. As a consequence, the evaluation and modeling activity was executed for the entire measured time, without
14   distinction between occupied/not occupied timeslots (as those were unknown).
15   In the next section, an example of an opening and a closing model, as result of the LRA, are introduced. Hence, in section 4.2
16   and in section 4.3 the results of all the windows are presented. The results are organized in graphics, showing the explanatory
17   variables, selected with the procedure explained in section 2, and the number of models using each of them. In this way, the
18   occupants' behavior related to the operations of 300 windows (corresponding to 300 analyses resulting in 300 models for the
19   opening operation and 300 models for the closing operation) can be evaluated. The results shown are related to those windows,
20   for which at least one state change per year was observed (some of the sensors were defect and did not register the changes of
21   state): thus, the results were based on 273 windows.
22

4.1. Example of an opening and a closing model of one window

24   One model for the opening and one model for the closing action of one window were selected, with the aim to illustrate the way
25   the models work; the models presented within this section were selected randomly from the 546 models (273 for opening, 273
26   for closing), and therefore those models are not representative. The presented models were obtained from the data monitored in
27   a living room of one apartment located in B2E1. The models were defined by equation 5 and 6, respectively for the opening and
28   the closing action. The terms of the equations were defined as following:
29   1. α is the intercept;
30   2. β are the coefficients, free of interactions;
31

32 $$ln\left(\frac{p_o}{1-p_o}\right) = \alpha_{o,Time} + \beta_{o,f(CO_2)}x_{f(CO_2)} + \beta_{o,RT}x_{RT} \tag{5}$$
33

34 $$ln\left(\frac{p_c}{1-p_c}\right) = \alpha_{c,Time} + \beta_{c,DAT}x_{DAT} + \beta_{c,f(CO_2)}x_{f(CO_2)} + \beta_{c,ARH}x_{ARH} + \beta_{c,RRH}x_{RRH} + \beta_{c,RT}x_{RT} \tag{6}$$

35   Table 3 Intercept α and coefficients β of the explanatory variables for the living room's window model of an apartment in B2E1
36   (n.p. implies that the variable was "not present" in the model).

Expl. variable	B2E1, models of a living room's window					
	Opening action			Closing action		
	Coef.	Conf. interval		Coef.	Conf. interval	
		2.50%	*97,5%*		*2.50%*	*97,5%*
α_{Night}	-10.089	-11.262	-8.916	2.539	0.559	4.519
$\alpha_{Morning}$	-8.214	-9.431	-6.998	3.317	1.309	5.325
$\alpha_{Rest\ of\ the\ day}$	-7.795	-9.002	-6.587	3.955	1.956	5.954
$\beta_{f((CO2^{-1}))}$	-551.15	-696.32	-405.98	-785.70	-934.88	-636.52
β_{RT}	0.134	0.082	0.187	-0.268	-0.348	-0.188
β_{RRH}	n.p.	n.p.	n.p.	-0.058	-0.071	-0.044
β_{DAT}	n.p.	n.p.	n.p.	-0.089	-0.105	-0.074
β_{ARH}	n.p.	n.p.	n.p.	0.027	0.022	0.031

37   The coefficients to solve equation 5 and equation 6, and their confidence interval, are listed in Table 3. When the coefficients β
38   were positive, an increase of the related explanatory variable caused an increase of the probability of the opening/closing
39   action, while, when negative, an increase of the explanatory variable caused a decrease of the probability of the opening/closing

1 action. Since the CO_2 concentration was transformed through the reciprocal function, an increase of the CO_2 concentration
2 caused a decrease of the reciprocal of the CO_2 concentration: hence a negative sign of the coefficient of $f(CO_2)$ meant that with
3 increasing CO_2 concentration in the room, the probability of opening/closing a window increased.

4
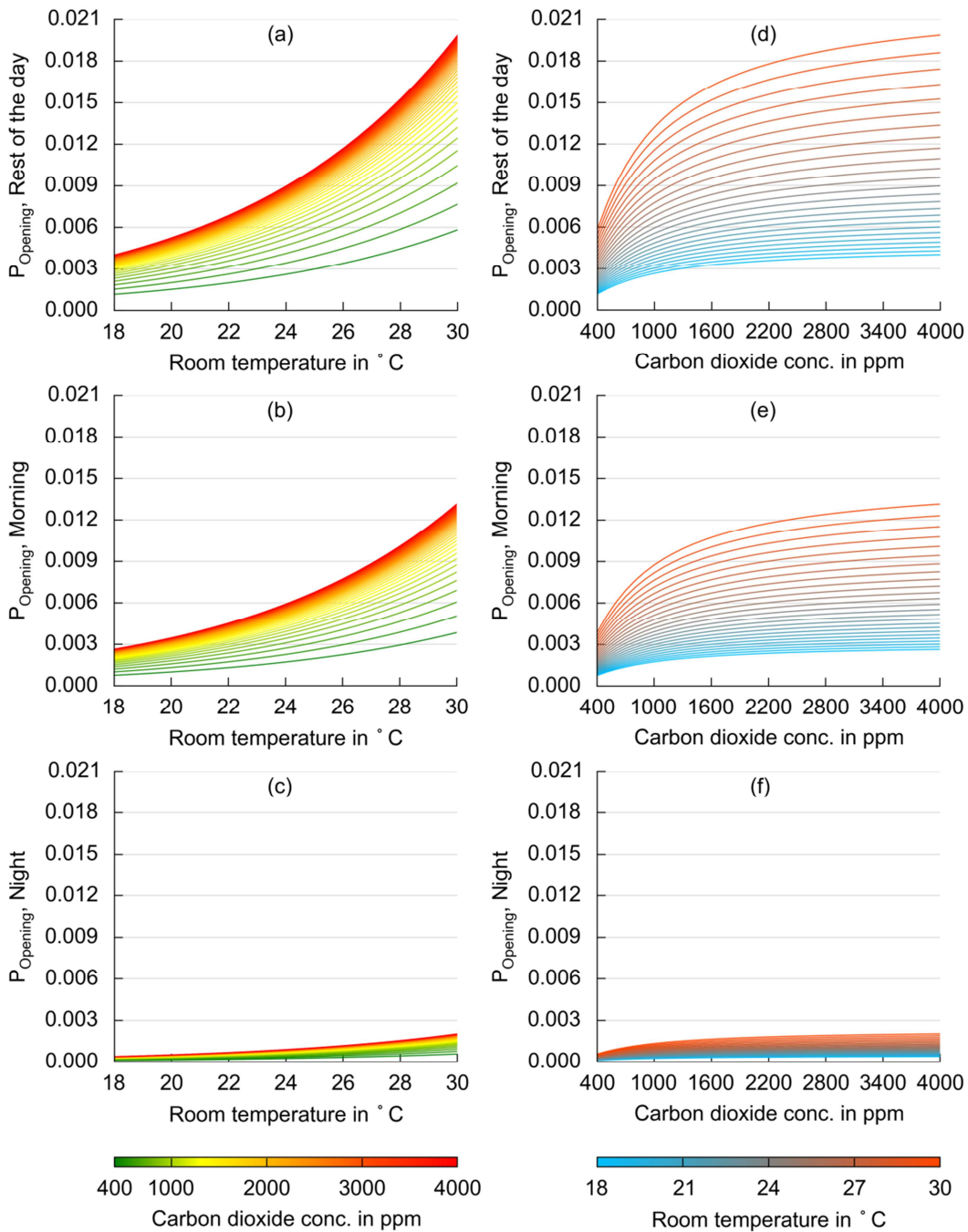
7
8 Figure 9. Probability of opening action for the window in the living room of an apartment located in B2E1, at night (c and f), in
9 the morning (b and e) and during the rest of the day (a and d).

10 Which time slot is analyzed, depends on the chosen intercept: the explanatory categorical variable "time" only affected the
11 intercept, since no interaction terms between this variable and other continuous variables were present in these particular
12 models. For the presented opening action model, the probability of opening a window grew with increasing CO_2 concentration
13 and with increasing room temperature. In the closing action model, the probability of closing a window grew with:
14   1. increasing CO_2 concentration,
15   2. decreasing room temperature,
16   3. decreasing relative humidity in the room,
17   4. decreasing daily average outdoor temperature,
18   5. decreasing outdoor relative humidity.

1 The probability of the opening action in the model described in equation 5 is graphically illustrated in Figure 9. The probability of
2 the opening action for this particular window was very low at night, high in the morning and very high during the rest of the day.
3 The probability of opening varied between 0 and 0.02. The small values were a consequence of the very short measuring
4 interval (60 s). In essence, the model calculated the probability of opening and closing a window within the next minute. This
5 window was in "open status" 32.1 % of the time during 2012 (The analyses of this room were based on a time period of 329
6 days and had a 10% failure rate of the monitoring system) and the window was closed (and opened, when it was in status
7 closed) 1014 times. On average, the probability of opening that window in 2012 was equal to 0.0032 (1014 opening events in
8 *321683 minutes,* 67.9 % of 329 days), while the probability of closing that window was equal to 0.0067 (1014 closing events in
9 152077 minutes, 32.1 % of 329 days).

10   4.2. Evaluation of the window opening action

11 Figure 10 shows the explanatory variables of the models of opening a window. The only categorical variable used, "time", was
12 included in over 70% of the windows. All the interaction terms with the variable "time" were used by less than 10% of the models
13 with the consequence that the variable time commonly influenced the intercept, but not the coefficients of the continuous
14 explanatory variables.
15 The most common continuous explanatory variable was the carbon dioxide concentration, present in over 50% of the models.
16 The room temperature, the relative humidity of the room, the daily average outdoor temperature, and the outdoor relative
17 humidity were used by over 35% of the models. The wind speed turned out to be mostly irrelevant for the opening action.
18 "None" indicates that no explanatory variables were found for the fitting of the model.
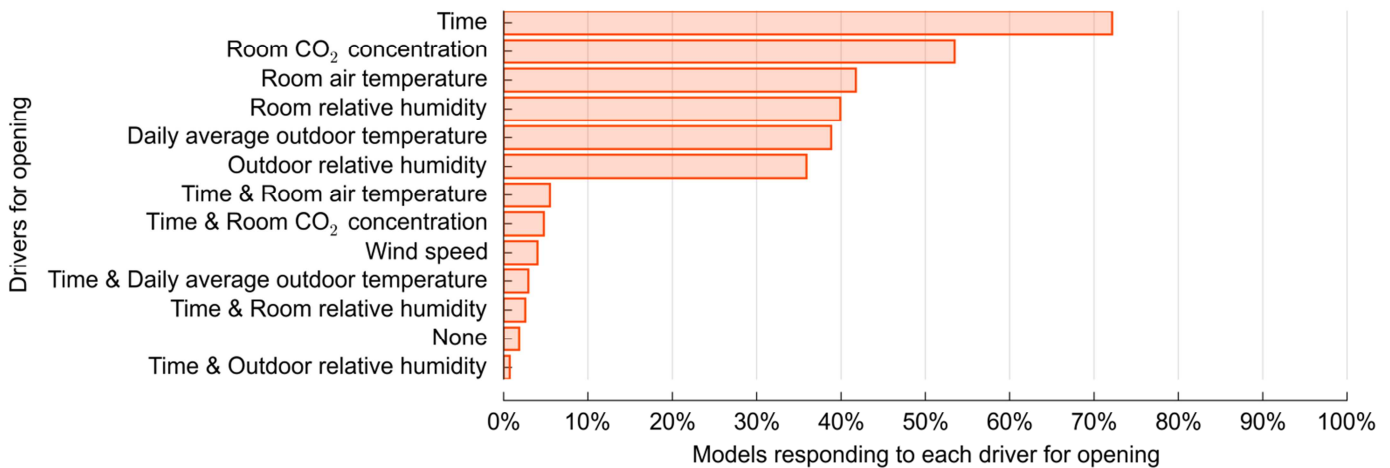
19
20 Figure 10. Drivers for opening and the number of rooms/models with the driver used as explanatory variable.
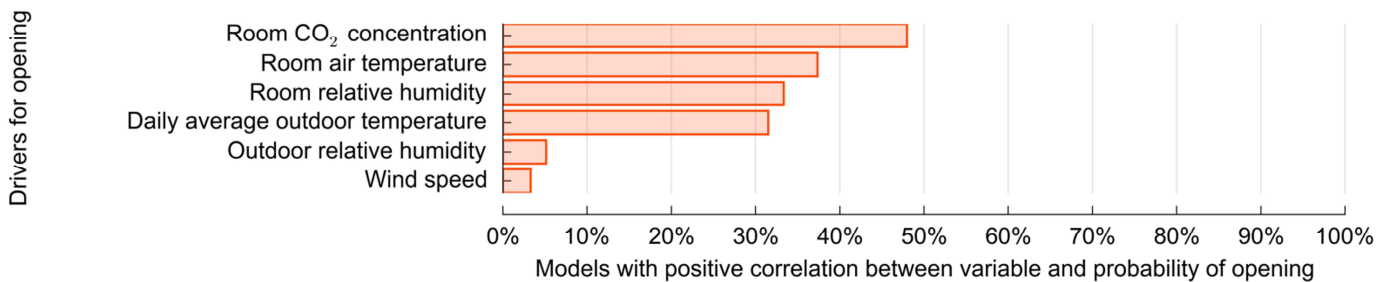
21
22 Figure 11. Drivers for opening and number of rooms/models with a positive correlation between the variable and the probability
23 of opening.

24 Figure 11 and Figure 12 show the drivers which directly (positively) and inversely (negatively) influenced the probability of the
25 action "window opening" respectively (this can be verified based on the sign of the coefficients: a negative coefficient shows a
26 negative correlation to the probability of action). An increase in carbon dioxide concentration leads to an increase of the
27 probability of window opening for more than 45% of the windows; However, contrary to this result, the carbon dioxide
28 concentration negatively influenced the opening action for approximately 6% of the models; this is not necessarily a
29 contradiction, and could be related to the occupancy (which is not included in the models) and window opening behavior upon
30 arrival. The probability of opening windows increased with increasing indoor temperatures, and by increasing room relative
31 humidity and daily average outdoor temperature (each of these explanatory variables was present in more than 35% of the
32 models). The only remarkable explanatory variable that negatively influenced the opening of windows (more than 30 % of the
33 windows) was the outdoor relative humidity.

Figure 12. Drivers for opening and number of rooms/models with a negative correlation between the driver and the probability of opening.

4.3. Evaluation of the window closing action

Figure 13 shows the explanatory variables of the logistic regression models of closing action. Figure 14 and Figure 15 show the drivers that positively and negatively influenced window opening action respectively. The daily average outdoor temperature was the most common driver (in almost 70% of the models); in particular, as Figure 15 shows, the probability of closing windows increased with decreasing daily average outdoor temperature. The variable time was present in almost half of the models.

Further, an increase in the carbon dioxide concentration was associated with increasing probability of the closing action: the carbon dioxide concentration correlated with the presence of occupants, and 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.

Figure 13. Drivers for closing and number of apartments/models with the driver used as explanatory variable.

Figure 14. Drivers for closing and number of rooms/models with a positive correlation between the variable and the probability of closing.
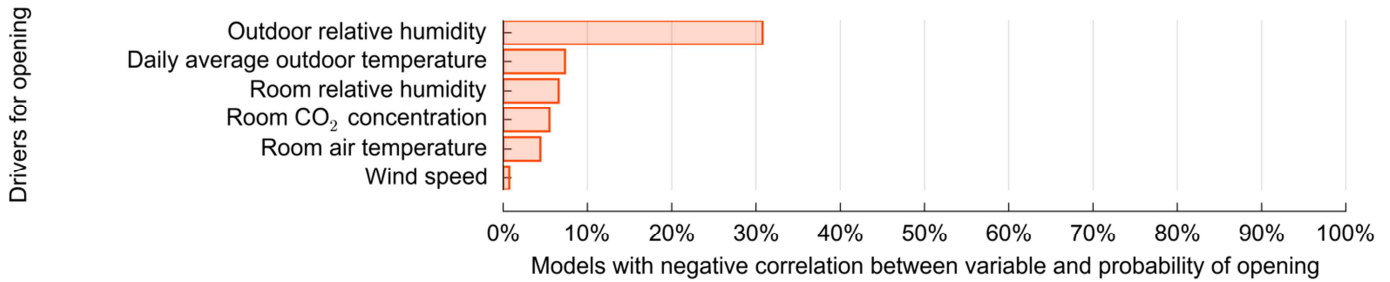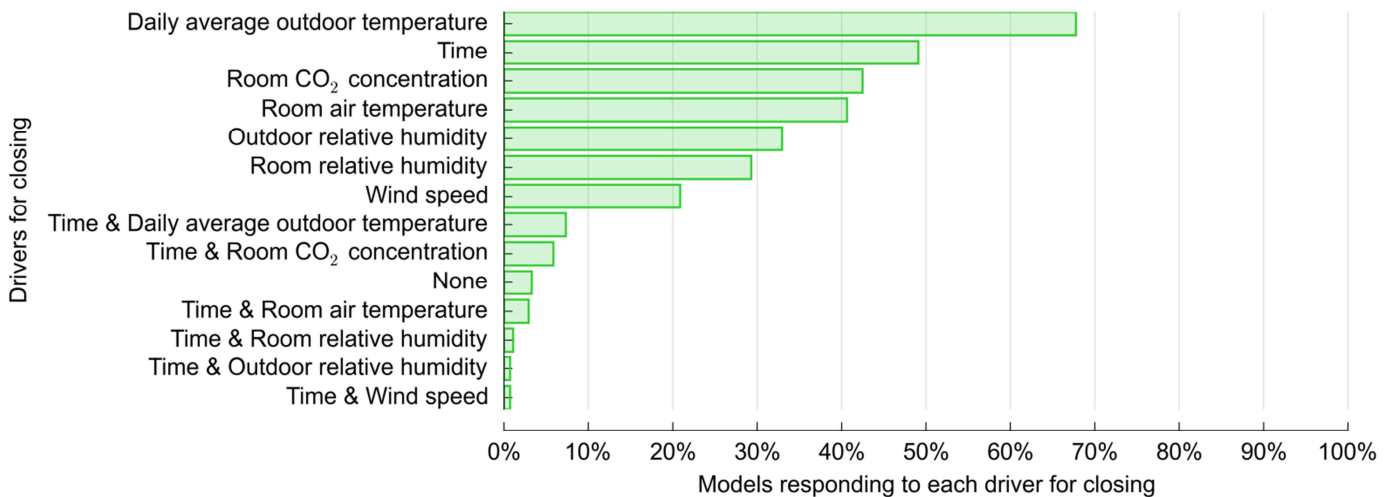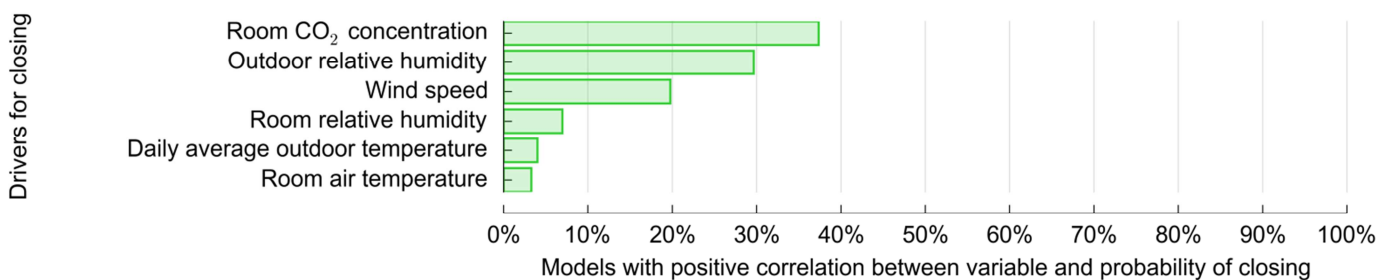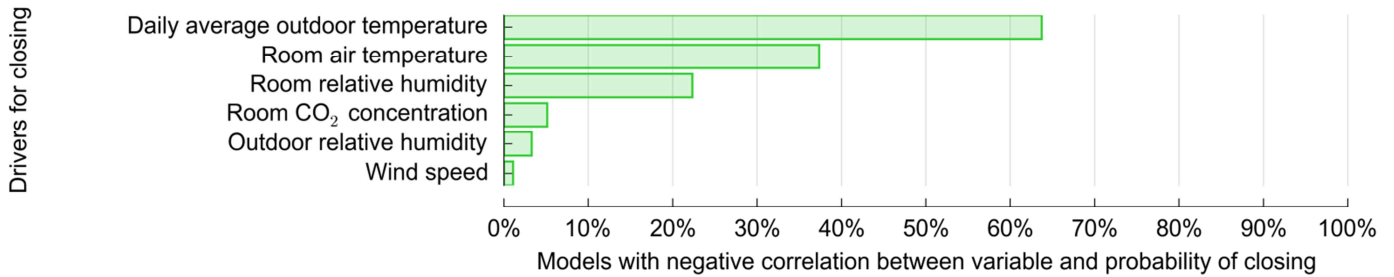
Figure 15. Drivers for closing and number of rooms/models with a negative correlation between the driver and the probability of closing.

4.4. Differences in behavior related to room typology

In this section, the explanatory variables of the logistic regression models for the opening and the closing action are illustrated, depending on room typology. The names "living room", "bedroom" and "children room" only refer to the nomenclature used in the floor space of Figure 3; the real use of those rooms is unknown. In contrast to the other rooms, the living room has a balcony-door. The most common drivers for the opening and the closing action, depending on room typology, are listed in Table 4 and Table 5 respectively.

The daily average outdoor temperature influenced the opening models of kitchen windows less than the window models of the other rooms. In addition, the relative humidity was present in two thirds of the models of bathroom windows, and in more than half of the models of kitchen windows. Finally, wind speed was found in 13 % of the opening models of the living room windows, against an average value over the entire field test equal to 4 %; for the closing models, wind speed was found in 55.6 % of the models, against an average value over the entire field test of 20.9 % This could be due to the fact that the living room windows are much bigger than the windows in the other rooms; therefore wind may have a stronger impact on the comfort of occupants.

4.5. Differences in behavior related to HVAC system

Within this section, the share of explanatory variables within the models is compared depending on the retrofit solution of the buildings (as illustrated in Table 6 and Table 7). The CO_2 concentration was a driver for 53.5 % of the windows, varying between 41.3 % for B2E2 and 70.8 % for B3E3: however, these two entrances have the same ventilation system. Similar considerations can be applied to the other variables. In conclusion, there were no clear trends showing differences in the drivers for window opening behavior between the buildings' different ventilation systems..

4.6. Differences in behavior related to level of the apartments

In this section, the share of explanatory variables within the models is compared depending on the level of the apartments (as illustrated in Table 8 and Table 9). The variables varied strongly depending on the level of the apartments. However, it was not possible to find systematic patterns in the variation of the behavior. The change in behavior may have been a result of distribution of occupants within the different levels.

Table 4 Share of the drivers for the opening action, of all the models and of the models of each room typology (e.g. Time 73.7 % for the kitchen indicates that 73.7 % of the 60 models obtained by analyzing the windows in the 60 kitchens, located in the demonstration buildings, used the explanatory variable, time).

	All	Kitchen	Bathroom	Livingroom	Bedroom	Childrenroom
Time	72.2%	73.7%	76.7%	83.3%	57.7%	68.0%
Room CO_2 concentration	53.5%	47.4%	56.7%	68.5%	40.4%	54.0%
Room air temperature	41.8%	47.4%	51.7%	50.0%	19.2%	38.0%
Room relative humidity	39.9%	52.6%	66.7%	33.3%	17.3%	24.0%
Daily average outdoor temperature	38.8%	24.6%	31.7%	51.9%	42.3%	46.0%
Outdoor relative humidity	35.9%	45.6%	36.7%	40.7%	26.9%	28.0%
Time & Room air temperature	5.5%	3.5%	8.3%	7.4%	0.0%	8.0%
Time & Room CO_2 concentration	4.8%	1.8%	1.7%	3.7%	7.7%	10.0%
Wind speed	4.0%	3.5%	1.7%	13.0%	0.0%	2.0%
Time & Daily average outdoor temperature	2.9%	0.0%	1.7%	3.7%	5.8%	4.0%
Time & Room relative humidity	2.6%	1.8%	6.7%	1.9%	0.0%	2.0%
None	1.8%	1.8%	0.0%	0.0%	3.8%	4.0%
Time & Outdoor relative humidity	0.7%	1.8%	0.0%	1.9%	0.0%	0.0%

1 Table 5 Share of the drivers for the closing action of all the models and of the models of each room typology.

	All	Kitchen	Bathroom	Livingroom	Bedroom	Childrenroom
Daily average outdoor temperature	67.8%	64.9%	75.0%	74.1%	57.7%	66.0%
Time	49.1%	40.4%	36.7%	72.2%	46.2%	52.0%
Room CO_2 concentration	42.5%	26.3%	43.3%	72.2%	36.5%	34.0%
Room air temperature	40.7%	42.1%	35.0%	44.4%	40.4%	42.0%
Outdoor relative humidity	33.0%	22.8%	25.0%	55.6%	30.8%	32.0%
Room relative humidity	29.3%	19.3%	23.3%	50.0%	21.2%	34.0%
Wind speed	20.9%	8.8%	3.3%	55.6%	15.4%	24.0%
Time & Daily average outdoor temperature	7.3%	7.0%	5.0%	13.0%	7.7%	4.0%
Time & Room CO_2 concentration	5.9%	0.0%	0.0%	20.4%	1.9%	8.0%
None	3.3%	5.3%	1.7%	1.9%	3.8%	4.0%
Time & Room air temperature	2.9%	1.8%	0.0%	9.3%	0.0%	4.0%
Time & Room relative humidity	1.1%	0.0%	0.0%	1.9%	0.0%	4.0%
Time & Outdoor relative humidity	0.7%	0.0%	0.0%	3.7%	0.0%	0.0%
Time & Wind speed	0.7%	0.0%	0.0%	1.9%	0.0%	2.0%

2

3

4 Table 6 Share of the drivers for the opening action, of all the models and of the models of each retrofit layout.

	All	B2E1	B2E2	B2E3	B3E1	B3E2	B3E3
Time	72.2%	77.8%	76.1%	84.6%	70.8%	66.0%	60.4%
Room CO_2 concentration	53.5%	57.8%	41.3%	64.1%	45.8%	42.6%	70.8%
Room air temperature	41.8%	46.7%	37.0%	41.0%	35.4%	36.2%	54.2%
Room relative humidity	39.9%	42.2%	47.8%	61.5%	43.8%	29.8%	18.8%
Daily average outdoor temperature	38.8%	37.8%	41.3%	48.7%	35.4%	42.6%	29.2%
Outdoor relative humidity	35.9%	28.9%	47.8%	41.0%	43.8%	23.4%	31.3%
Time & Room air temperature	5.5%	8.9%	8.7%	5.1%	4.2%	4.3%	2.1%
Time & Room CO_2 concentration	4.8%	4.4%	6.5%	7.7%	8.3%	2.1%	0.0%
Wind speed	4.0%	2.2%	4.3%	2.6%	2.1%	4.3%	8.3%
Time & Daily average outdoor temperature	2.9%	0.0%	0.0%	7.7%	4.2%	4.3%	2.1%
Time & Room relative humidity	2.6%	2.2%	2.2%	5.1%	4.2%	2.1%	0.0%
None	1.8%	2.2%	0.0%	0.0%	4.2%	2.1%	2.1%
Time & Outdoor relative humidity	0.7%	0.0%	2.2%	0.0%	2.1%	0.0%	0.0%

5

6 Table 7 Share of the drivers for the closing action, of all the models and of the models of each retrofit layout.

	All	B2E1	B2E2	B2E3	B3E1	B3E2	B3E3
Daily average outdoor temperature	67.8%	77.8%	80.4%	74.4%	58.3%	61.7%	56.3%
Time	49.1%	53.3%	43.5%	64.1%	47.9%	57.4%	31.3%
Room CO_2 concentration	42.5%	51.1%	32.6%	51.3%	39.6%	51.1%	31.3%
Room air temperature	40.7%	35.6%	41.3%	35.9%	43.8%	44.7%	41.7%
Outdoor relative humidity	33.0%	37.8%	32.6%	48.7%	31.3%	31.9%	18.8%
Room relative humidity	29.3%	17.8%	13.0%	35.9%	41.7%	36.2%	31.3%
Wind speed	20.9%	17.8%	21.7%	23.1%	25.0%	23.4%	14.6%
Time & Daily average outdoor temperature	7.3%	2.2%	8.7%	10.3%	12.5%	2.1%	8.3%
Time & Room CO_2 concentration	5.9%	8.9%	6.5%	2.6%	6.3%	6.4%	4.2%
None	3.3%	2.2%	4.3%	0.0%	4.2%	6.4%	2.1%
Time & Room air temperature	2.9%	2.2%	0.0%	2.6%	2.1%	10.6%	0.0%
Time & Room relative humidity	1.1%	2.2%	0.0%	0.0%	2.1%	2.1%	0.0%
Time & Outdoor relative humidity	0.7%	0.0%	0.0%	0.0%	2.1%	0.0%	2.1%
Time & Wind speed	0.7%	0.0%	0.0%	0.0%	2.1%	2.1%	0.0%

7

8

Table 8 Share of the drivers for the opening action, of all the models and of the models of each level.

	All	1st Floor	2nd Floor	3rd Floor	4th Floor	5th Floor
Time	72.2%	73.6%	68.5%	81.5%	74.1%	63.8%
Room CO_2 concentration	53.5%	47.2%	51.9%	51.9%	50.0%	65.5%
Room air temperature	41.8%	54.7%	20.4%	46.3%	29.6%	56.9%
Room relative humidity	39.9%	37.7%	42.6%	44.4%	37.0%	37.9%
Daily average outdoor temperature	38.8%	34.0%	37.0%	48.1%	38.9%	36.2%
Outdoor relative humidity	35.9%	28.3%	37.0%	51.9%	37.0%	25.9%
Time & Room air temperature	5.5%	9.4%	5.6%	3.7%	1.9%	6.9%
Time & Room CO_2 concentration	4.8%	0.0%	7.4%	0.0%	3.7%	12.1%
Wind speed	4.0%	9.4%	1.9%	1.9%	0.0%	6.9%
Time & Daily average outdoor temperature	2.9%	1.9%	0.0%	3.7%	5.6%	3.4%
Time & Room relative humidity	2.6%	1.9%	1.9%	1.9%	3.7%	3.4%
None	1.8%	0.0%	1.9%	0.0%	3.7%	3.4%
Time & Outdoor relative humidity	0.7%	0.0%	0.0%	1.9%	1.9%	0.0%

Table 9 Share of the drivers for the closing action, of all the models and of the models of each level.

	All	1st Floor	2nd Floor	3rd Floor	4th Floor	5th Floor
Daily average outdoor temperature	67.8%	62.3%	66.7%	70.4%	68.5%	70.7%
Time	49.1%	52.8%	46.3%	59.3%	37.0%	50.0%
Room CO_2 concentration	42.5%	43.4%	46.3%	50.0%	29.6%	43.1%
Room air temperature	40.7%	43.4%	31.5%	48.1%	35.2%	44.8%
Outdoor relative humidity	33.0%	28.3%	35.2%	37.0%	25.9%	37.9%
Room relative humidity	29.3%	24.5%	29.6%	42.6%	27.8%	22.4%
Wind speed	20.9%	20.8%	16.7%	25.9%	20.4%	20.7%
Time & Daily average outdoor temperature	7.3%	11.3%	3.7%	11.1%	3.7%	6.9%
Time & Room CO_2 concentration	5.9%	11.3%	5.6%	7.4%	1.9%	3.4%
None	3.3%	0.0%	1.9%	3.7%	5.6%	5.2%
Time & Room air temperature	2.9%	3.8%	0.0%	3.7%	3.7%	3.4%
Time & Room relative humidity	1.1%	1.9%	1.9%	0.0%	1.9%	0.0%
Time & Outdoor relative humidity	0.7%	0.0%	0.0%	3.7%	0.0%	0.0%
Time & Wind speed	0.7%	0.0%	0.0%	0.0%	1.9%	1.7%

5. Discussion

In the analysis the daily average outdoor temperature was used instead of the actual outdoor temperature measured each minute to infer the probability of opening and closing of windows. The wide outdoor temperature variation within a day may have different impact on the probability of opening and closing the windows. However, the choice was made to avoid correlation issues of this variable with other variables such as room temperature, outdoor humidity and time of day. Additionally, this can facilitate the use of the models for future applications when, for example, only daily average values of the outdoor temperature are available.

The models that were used to investigate the drivers were all developed using the AIC as a means to select variables. It turned out that all variables in the models had a statistically significant impact ($p<0.05$) on the window opening/closing probability. This implies that the observed correlations are not likely to be random. In the analyses, we investigated the sign of the correlation (positive or negative) but we did not investigate the size of the coefficients, which is a measure of the size of the effect. As such, the results in this paper only state how the variables affected the window opening/closing probability (positively or negatively). It does not state how big the effects were.

Furthermore, in order to keep the interpretation of the results at a manageable level, interactions between continuous variables were not included. Interaction terms between the categorical variable "time" with the continuous variables were included, but were used by less than 10 % of the models.

The main scope of the work was the investigation of the drivers for opening and closing windows, as discussed in section 4.2 and 4.3 respectively. Time of day was the most frequent driver for opening windows and the second most frequent driver for closing windows. This could be an indication that the observed window opening behavior was linked to specific activities which occur at specific times. E.g. a state change is less likely to occur when the occupants are asleep than when they are awake. And most occupants tend to sleep during the night. As such, the lower probability of opening and closing windows during the night was probably linked to the occupants' sleeping patterns rather than the time itself. Likewise, the peaks in state changes during morning and evening in Figure 8 could be a consequence of typical activities at those times of the day (e.g. showering, cooking, or simply moving from one room to another). However, since we did not gather information about the occupants' activities, we cannot point to a direct explanation.

In general, the time of day, the CO_2 concentration and the average outdoor temperature appeared most frequently in the models of window opening and closing. These results support the tendency of the results in [27] and underline the importance of including indoor air quality indicators in any attempts of modeling occupants' window opening behavior in residential buildings.

The most frequently occurring driver for closing of windows was the daily outdoor average temperature and in the majority of the models the association was negative, i.e. the probability of closing a window increased with decreasing daily outdoor average temperature. This is in agreement with Fabi et al. [41], who found that outdoor temperature was negatively correlated with window closing in two of the three best performing of window opening behavior. From this, a general picture emerges: Occupants tend to open the windows at specific times of day (probably associated to activities) and when the CO_2 concentration and relative humidity is elevated. They tend to close windows when it is cold outside and at specific times of day (probably associated with their activities).

The differences in drivers for the behavior related to the room typology, to the retrofit layout and to the level of the apartment were investigated. The room typology (kitchen, bathroom, other) and the window type (balcony-door, normal window) clearly lead occupants to a diversified behavior. The time of day remained as the single most frequent driver for window opening across all room typologies and room relative humidity was the second most frequent driver in the kitchen and bathroom. In the living room and children room, the second most frequent driver was the CO2 concentration. This indicates that the occupants' window opening behavior was driven by the activities in the home and aimed at removing pollutants from the indoor air (in this case, moisture from showering and cooking and bioeffluents from the rest of the rooms). No systematic changes in drivers for the observed behavior were recognizable, when comparing the retrofit layout and the level of the apartments. Interestingly, no pattern was recognized when comparing B2 and B3E3 (equipped with exhaust ventilation) to B3E1 and B3E2 (equipped with balanced ventilation with heat recovery). Even though the drivers for the opening and closing of windows were not different between the retrofit layouts, the average time with window open was different, both between the apartments, and between the entrances. For example, in 2012, B2E3 had open windows for 29 % of the time, B3E2 for 39 % of the time. In order to investigate whether those differences were statistically significant or not (e.g. due to the random distribution of tenants within the buildings), the distributions of the average positions of the windows of each apartment were used for a double comparison (each comparison for two entrances) in the Wilcoxon-Mann-Whitney test: the results indicated statistical significance of the differences between B2E1 and B3E2, and between B2E3 and B3E2, but only for the year 2012.

Each window was modeled through a logistic function. Those models can be implemented in building energy performance simulation software. Since the occupants, and thus the models, react to variables that are influenced by the presence of occupants (e.g. carbon dioxide concentration in the air), those models can only be used within a simulation, when occupants' presence profiles are available. Only if presence patterns are provided, the CO_2 emissions caused by occupants could be reproduced in the dynamic simulation, and therefore the model could render the occupants' reaction to this air quality parameter correctly. However, this has the disadvantage that presence profiles and window models would not be congruent, since those would come from different observations or only from general assumptions. Furthermore, in the literature, a model that reproduces the presence of occupants in the residential sector at room level resolution, and based on observations, does not exist at the time of writing. To test the models within a building energy simulation software, presence patterns simulating the movements of occupants between the rooms of an apartment could be modeled using an agent based approach through inhomogeneous Markov Chains based on assumptions about the presence of each occupant and the probability of their transition between rooms at given timestamps. The occupants of 34 of the 60 apartments were interviewed, and disclosed the number of occupants in their own apartment, as well as the average time per day when the apartment is not occupied, distinguishing between weekdays and weekends. Such information, together with the collected data in the rooms, could be used to generate the necessary presence patterns, making use of the Bayes theorem and the Markov Chains technique [42].

The LRA procedure was effective when focusing on the evaluation of the drivers but had the disadvantage of generating two models for each window, one for opening and one for closing. In order to model the occupants' behavior, it would probably be more meaningful to make one model per apartment and use the room as an explanatory variable [27]; alternatively, it would be possible to make one general model for opening and one for closing, using the mixed effect modeling technique, and setting the apartment and the room as further explanatory variables [43].

To generate congruent models of presence and behavior, with a holistic approach, it would be beneficial to follow the procedure described within this paper for buildings with known occupancy patterns. Those patterns could help on the one hand, to better and more precisely fit the models, and on the other hand, to use the models under the real occupants' presence conditions.

Logistic regression was confirmed to be a strong and robust analysis methodology for investigating the drivers for occupants to interact with the built environment. The strengths of the models are related to the high time resolution of the data, the use of an indoor air quality parameter as an explanatory variable, and the distinction of the room use typology. The weaknesses are mostly due to the lack of data related to the presence of occupants. Through the logistic regression analysis, each window was modeled singularly. Thus, 546 models were generated. In order to reduce the number of models while including the occupants' diversity, Haldi [28] suggested the use of a generalized linear mixed effect logistic regression model. This approach could be investigated in future modeling activities, based on the observed data from the demonstration buildings or of new field tests, where the presence of occupants could be monitored.

6. Conclusions

The results of the logistic regression identified the drivers of the opening and closing of windows. The most common drivers for opening windows were time of 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 closing windows were the daily average outdoor temperature (for almost 70 % of the modeled windows) and the time of day (for more than 50 % of the modeled windows).

The room typology (kitchen, bathroom, other) and the window type (balcony-door/normal window) affected the frequency with which the drivers appeared in the models of both window openings and closings. The time of day and average outdoor temperature remained the most frequent drivers for opening and closing, respectively, across all room typologies and window types. For the opening case, the room relative humidity was the second most frequent driver (after the time of day) in the

kitchen and bathroom, while in the living room and children's room, the second most frequent driver was the CO_2 concentration. This indicates that the occupants' window opening behavior was driven by the activities in the home (like showering and cooking activities, which produce moisture) and by pollutants in the indoor air (windows were opened to remove moisture from kitchen and bathroom and to remove bioeffluents from the rest of the rooms).

Finally, no systematic changes in drivers for the observed behavior were recognizable, when comparing the ventilation principles and the level of the apartments.

Acknowledgment

References

[1]   R. Haas, H. Auer, P. Biermayr, The impact of consumer behavior on residential energy demand for space heating, Energy Policy (1998) 195–205.

[2]   R. Haas, P. Biermayr, The rebound effect for space heating Empirical evidence from Austria, Energy Policy (2000) 403–410.

[3]   H. Erhorn, Bedarf - Verbrauch: Ein Reizthema ohne Ende oder die Chance für sachliche Energieberatung?, available at http://www.buildup.eu/publications/1810 (accessed on May 16, 2015).

[4]   L. Tronchin, K. Fabbri, Energy performance building evaluation in Mediterranean countries: Comparison between software simulations and operating rating simulation, Energy and buildings 40 (2008) 1176–1187.

[5]   H. Hens, W. Parijs, M. Deurinck, Energy consumption for heating and rebound effects, Energy and buildings 42 (2010) 105–110.

[6]   A.C. Menezes, A. Cripps, D. Bouchlaghem, R. Buswell, Predicted vs. actual energy performance of non-domestic buildings: Using post-occupancy evaluation data to reduce the performance gap, Applied Energy 97 (2012) 355–364.

[7]   G. Dall'O', L. Sarto, A. Galante, G. Pasetti, Comparison between predicted and actual energy performance for winter heating in high-performance residential buildings in the Lombardy region (Italy), Energy and buildings 47 (2012) 247–253.

[8]   P. de Wilde, The gap between predicted and measured energy performance of buildings: A framework for investigation, Automation in Construction 41 (2014) 40–49.

[9]   A. Al-Mumin, O. Khattab, G. Sridhar, Occupants' behavior and activity patterns influencing the energy consumption in the Kuwaiti residences, Energy and buildings 35 (2003) 549–559.

[10]   A.L. Pisello, C. Piselli, F. Cotana, Influence of human behavior on cool roof effect for summer cooling, Building and Environment (2014).

[11]   O. Guerra Santin, C.A. Tweed, In-use monitoring of buildings: An overview of data collection methods, Energy and buildings 93 (2015) 189–207.

[12]   T. de Meester, A.-F. Marique, A. de Herde, S. Reiter, Impacts of occupant behaviours on residential heating consumption for detached houses in a temperate climate in the northern part of Europe, Energy and buildings 57 (2013) 313–323.

[13]   D. Müller, T. Osterhage, D. Calì, Quartierskonzept Energieeffizientes Rintheim - wissenschaftliche Begleitung. Lehrstuhl für Gebäude- und Raumklimatechnik, RWTH Aachen University. Final Report, 2012.

[14]   D. Calì, S. Rita, D. Müller, T. Osterhage, Holistic Renovation and Monitoring of Residential Buildings, ECEEE (2013).

[15]   J. Virote, R. Neves-Silva, Stochastic models for building energy prediction based on occupant behavior assessment, Energy and buildings 53 (2012) 183–193.

[16]   D. Calì, P. Matthes, K. Huchtemann, R. Streblow, D. Müller, CO2 based occupancy detection algorithm: Experimental analysis and validation for office and residential buildings, Building and Environment (2015) 39–49.

[17]   I. Richardson, M. Thomson, D. Infield, A high-resolution domestic building occupancy model for energy demand simulations, Energy and buildings 40 (2008) 1560–1566.

[18]   U. Wilke, F. Haldi, J.-L. Scartezzini, D. Robinson, A bottom-up stochastic model to predict building occupants' time-dependent activities, Building and Environment 60 (2013) 254–264.

[19]   Energy in Buildings and Communities Programme, IEA-EBC Annex 66 - Definition and Simulation of Occupant Behavior in Buildings, available at http://www.annex66.org/ (accessed on January 19, 2016).

[20]   J. Page, D. Robinson, N. Morel, J.-L. Scartezzini, A generalised stochastic model for the simulation of occupant presence, Energy and buildings 40 (2008) 83–98.

[21]   J. Page, Simulating occupant presence and behaviour in buildings. Ph.D. Thesis, Laboratoire d'Energie Solaire et de Physique du Bâtiment, École polytechnique fédérale de Lausanne, 2007.

[22]   H. Polinder, M. Schweiker, van der Aa, Ad, K. Schakib-Ekbatan, V. Fabi, R. Andersen, N. Morishita, C. Wang, S. Corgnati, P. Hiselberg, D. Yan, B. Olesen, T. Bednar, A. Wagner, Occupant behavior and modeling: Separate Document Volume II - Total energy use in buildings - analysis and evaluation methods - Final Report Annex 53, International Energy Agency - Institute for Building Environment and Energy Conservation, 2013.

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

[24] European Environment Agency, Household energy consumption per dwelling by end-use: Chart — Household energy consumption per dwelling by end-use, available at http://www.eea.europa.eu/data-and-maps/daviz/energy-consumption-by-end-uses-1#tab-chart_1 (accessed on February 11, 2016).

[25] V. Fabi, V.A. Rune, S. Corgnati, B.W. Olesen, Occupants' window opening behaviour: A literature review of factors influencing occupant behaviour and models, Building and Environment 58 (2012) 188–198.

[26] R. Fritsch, A. Kohler, M. Nygård-Ferguson, J.-L. Scartezzini, A stochastic model of user behaviour regarding ventilation, Building and Environment 25 (1990) 173–181.

[27] R. Andersen, V. Fabi, J. Toftum, S.P. Corgnati, B.W. Olesen, Window opening behaviour modelled from measurements in Danish dwellings, Building and Environment 69 (2013) 101–113.

[28] F. Haldi, D. Robinson, Interactions with window openings by office occupants, Building and Environment 44 (2009) 2378–2395.

[29] R.V. Andersen, Occupant behaviour with regard to control of the indoor environment: Ph.D. Thesis (2009).

[30] R.V. Andersen, B.W. Olesen, J. Toftum, Modelling window opening behaviour in Danish dwellings: International Conference on indoor air quality and climate, Austin, Texas. (2011).

[31] S. D'Oca, T. Hong, A data-mining approach to discover patterns of window opening and closing behavior in offices, Building and Environment 82 (2014) 726–739.

[32] T.J. Hastie, R.J. Tibshirani, Generalized additive models, Chapman and Hall/CRC, London, 1990.

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

[34] R. Rovers, J. Kimman, C. Ravesloot, Towards 0-impact buildings and built environments, Techne Press, Amsterdam, 2010.

[35] D. Calì, T. Osterhage, D. Müller, Field study of different retrofit solutions for residential housing. Contribution to a conference proceedings, Euregional conference Sustainable Building, SB 10 (2010).

[36] D. Calì, T. Osterhage, D. Müller, Rebound effect related to retrofit solutions for residential housing monitoring data from a field test, SB11 Helsinki World Sustainable Building Conference (2011).

[37] D. Calì, T. Osterhage, D. Müller, Retrofit Solutions for Residential Buildings, International Journal of Sustainable Building Technology and Urban Development 2 (2011) 131–136.

[38] D. Calì, T. Osterhage, R. Streblow, D. Müller, Advanced Renovation of three Residential Buildings – Evaluation of a Field Test. Contribution to a conference proceedings, Energy Efficiency in Domestic Appliances and Lighting proceedings of the 7th International Conference (2013).

[39] K. Wolfrum, A. Bloch, Messsystem für vielkanaliges Monitoring von Energieströmen und Raumklimadaten in energetisch sanierten Wohngebäuden in Karlsruhe Rintheim, Forschung aktuell 2012, S. 18-20, Hochschule Karlsruhe - Technik und Wirtschaft (2012).

[40] CEN, Ventilation for buildings – Design criteria for the indoor environment, 1998.

[41] V. Fabi, R.K. Andersen, S. Corgnati, Verification of stochastic behavioural models of occupants' interactions with windows in residential buildings, Building and Environment 94 (2015) 371–383.

[42] H. Han, C. Shin, Bayesian Analysis of Passenger Occupancy Based on the CO2 Concentration Measured in a Subway Station, Clima.

[43] F. Haldi, A probabilitsic model to predict building occupants' diversity toards their interaction with the building envelope., Conference of International Building Performance Simulation Association (2013).

	B2E1, models of a living room's window					
	Opening action			Closing action		
	Coef.	Conf. interval		Coef.	Conf. interval	
Expl. variable		2.50%	97,5%		2.50%	97,5%
α_{Night}	-10.089	-11.262	-8.916	2.539	0.559	4.519
$\alpha_{Morning}$	-8.214	-9.431	-6.998	3.317	1.309	5.325
$\alpha_{Rest\ of\ the\ day}$	-7.795	-9.002	-6.587	3.955	1.956	5.954
$\beta_{f(CO2)}$	-551.15	-696.32	-405.98	-785.70	-934.88	-636.52
β_{RT}	0.134	0.082	0.187	-0.268	-0.348	-0.188
β_{RRH}	n.p.	n.p.	n.p.	-0.058	-0.071	-0.044
β_{DAT}	n.p.	n.p.	n.p.	-0.089	-0.105	-0.074
β_{ARH}	n.p.	n.p.	n.p.	0.027	0.022	0.031

	All	Kitchen	Bathroom	Livingroom	Sleepingroom	Childrenroom
Time	72.2%	73.7%	76.7%	83.3%	57.7%	68.0%
Room CO_2 concentration	53.5%	47.4%	56.7%	68.5%	40.4%	54.0%
Room air temperature	41.8%	47.4%	51.7%	50.0%	19.2%	38.0%
Room relative humidity	39.9%	52.6%	66.7%	33.3%	17.3%	24.0%
Daily average outdoor temperature	38.8%	24.6%	31.7%	51.9%	42.3%	46.0%
Outdoor relative humidity	35.9%	45.6%	36.7%	40.7%	26.9%	28.0%
Time & Room air temperature	5.5%	3.5%	8.3%	7.4%	0.0%	8.0%
Time & Room CO_2 concentration	4.8%	1.8%	1.7%	3.7%	7.7%	10.0%
Wind speed	4.0%	3.5%	1.7%	13.0%	0.0%	2.0%
Time & Daily average outdoor temperature	2.9%	0.0%	1.7%	3.7%	5.8%	4.0%
Time & Room relative humidity	2.6%	1.8%	6.7%	1.9%	0.0%	2.0%
None	1.8%	1.8%	0.0%	0.0%	3.8%	4.0%
Time & Outdoor relative humidity	0.7%	1.8%	0.0%	1.9%	0.0%	0.0%

	All	Kitchen	Bathroom	Livingroom	Sleepingroom	Childrenroom
Daily average outdoor temperature	67.8%	64.9%	75.0%	74.1%	57.7%	66.0%
Time	49.1%	40.4%	36.7%	72.2%	46.2%	52.0%
Room CO_2 concentration	42.5%	26.3%	43.3%	72.2%	36.5%	34.0%
Room air temperature	40.7%	42.1%	35.0%	44.4%	40.4%	42.0%
Outdoor relative humidity	33.0%	22.8%	25.0%	55.6%	30.8%	32.0%
Room relative humidity	29.3%	19.3%	23.3%	50.0%	21.2%	34.0%
Wind speed	20.9%	8.8%	3.3%	55.6%	15.4%	24.0%
Time & Daily average outdoor temperature	7.3%	7.0%	5.0%	13.0%	7.7%	4.0%
Time & Room CO_2 concentration	5.9%	0.0%	0.0%	20.4%	1.9%	8.0%
None	3.3%	5.3%	1.7%	1.9%	3.8%	4.0%
Time & Room air temperature	2.9%	1.8%	0.0%	9.3%	0.0%	4.0%
Time & Room relative humidity	1.1%	0.0%	0.0%	1.9%	0.0%	4.0%
Time & Outdoor relative humidity	0.7%	0.0%	0.0%	3.7%	0.0%	0.0%
Time & Wind speed	0.7%	0.0%	0.0%	1.9%	0.0%	2.0%

	All	B2E1	B2E2	B2E3	B3E1	B3E2	B3E3
Time	72.2%	77.8%	76.1%	84.6%	70.8%	66.0%	60.4%
Room CO_2 concentration	53.5%	57.8%	41.3%	64.1%	45.8%	42.6%	70.8%
Room air temperature	41.8%	46.7%	37.0%	41.0%	35.4%	36.2%	54.2%
Room relative humidity	39.9%	42.2%	47.8%	61.5%	43.8%	29.8%	18.8%
Daily average outdoor temperature	38.8%	37.8%	41.3%	48.7%	35.4%	42.6%	29.2%
Outdoor relative humidity	35.9%	28.9%	47.8%	41.0%	43.8%	23.4%	31.3%
Time & Room air temperature	5.5%	8.9%	8.7%	5.1%	4.2%	4.3%	2.1%
Time & Room CO_2 concentration	4.8%	4.4%	6.5%	7.7%	8.3%	2.1%	0.0%
Wind speed	4.0%	2.2%	4.3%	2.6%	2.1%	4.3%	8.3%
Time & Daily average outdoor temperature	2.9%	0.0%	0.0%	7.7%	4.2%	4.3%	2.1%
Time & Room relative humidity	2.6%	2.2%	2.2%	5.1%	4.2%	2.1%	0.0%
None	1.8%	2.2%	0.0%	0.0%	4.2%	2.1%	2.1%
Time & Outdoor relative humidity	0.7%	0.0%	2.2%	0.0%	2.1%	0.0%	0.0%

	All	B2E1	B2E2	B2E3	B3E1	B3E2	B3E3
Daily average outdoor temperature	67.8%	77.8%	80.4%	74.4%	58.3%	61.7%	56.3%
Time	49.1%	53.3%	43.5%	64.1%	47.9%	57.4%	31.3%
Room CO_2 concentration	42.5%	51.1%	32.6%	51.3%	39.6%	51.1%	31.3%
Room air temperature	40.7%	35.6%	41.3%	35.9%	43.8%	44.7%	41.7%
Outdoor relative humidity	33.0%	37.8%	32.6%	48.7%	31.3%	31.9%	18.8%
Room relative humidity	29.3%	17.8%	13.0%	35.9%	41.7%	36.2%	31.3%
Wind speed	20.9%	17.8%	21.7%	23.1%	25.0%	23.4%	14.6%
Time & Daily average outdoor temperature	7.3%	2.2%	8.7%	10.3%	12.5%	2.1%	8.3%
Time & Room CO_2 concentration	5.9%	8.9%	6.5%	2.6%	6.3%	6.4%	4.2%
None	3.3%	2.2%	4.3%	0.0%	4.2%	6.4%	2.1%
Time & Room air temperature	2.9%	2.2%	0.0%	2.6%	2.1%	10.6%	0.0%
Time & Room relative humidity	1.1%	2.2%	0.0%	0.0%	2.1%	2.1%	0.0%
Time & Outdoor relative humidity	0.7%	0.0%	0.0%	0.0%	2.1%	0.0%	2.1%
Time & Wind speed	0.7%	0.0%	0.0%	0.0%	2.1%	2.1%	0.0%

	All	1st Floor	2nd Floor	3rd Floor	4th Floor	5th Floor
Time	72.2%	73.6%	68.5%	81.5%	74.1%	63.8%
Room CO_2 concentration	53.5%	47.2%	51.9%	51.9%	50.0%	65.5%
Room air temperature	41.8%	54.7%	20.4%	46.3%	29.6%	56.9%
Room relative humidity	39.9%	37.7%	42.6%	44.4%	37.0%	37.9%
Daily average outdoor temperature	38.8%	34.0%	37.0%	48.1%	38.9%	36.2%
Outdoor relative humidity	35.9%	28.3%	37.0%	51.9%	37.0%	25.9%
Time & Room air temperature	5.5%	9.4%	5.6%	3.7%	1.9%	6.9%
Time & Room CO_2 concentration	4.8%	0.0%	7.4%	0.0%	3.7%	12.1%
Wind speed	4.0%	9.4%	1.9%	1.9%	0.0%	6.9%
Time & Daily average outdoor temperature	2.9%	1.9%	0.0%	3.7%	5.6%	3.4%
Time & Room relative humidity	2.6%	1.9%	1.9%	1.9%	3.7%	3.4%
None	1.8%	0.0%	1.9%	0.0%	3.7%	3.4%
Time & Outdoor relative humidity	0.7%	0.0%	0.0%	1.9%	1.9%	0.0%

Daily average outdoor temperature	67.8%	62.3%	66.7%	70.4%	68.5%	70.7%
Time	49.1%	52.8%	46.3%	59.3%	37.0%	50.0%
Room CO_2 concentration	42.5%	43.4%	46.3%	50.0%	29.6%	43.1%
Room air temperature	40.7%	43.4%	31.5%	48.1%	35.2%	44.8%
Outdoor relative humidity	33.0%	28.3%	35.2%	37.0%	25.9%	37.9%
Room relative humidity	29.3%	24.5%	29.6%	42.6%	27.8%	22.4%
Wind speed	20.9%	20.8%	16.7%	25.9%	20.4%	20.7%
Time & Daily average outdoor temperature	7.3%	11.3%	3.7%	11.1%	3.7%	6.9%
Time & Room CO_2 concentration	5.9%	11.3%	5.6%	7.4%	1.9%	3.4%
None	3.3%	0.0%	1.9%	3.7%	5.6%	5.2%
Time & Room air temperature	2.9%	3.8%	0.0%	3.7%	3.7%	3.4%
Time & Room relative humidity	1.1%	1.9%	1.9%	0.0%	1.9%	0.0%
Time & Outdoor relative humidity	0.7%	0.0%	0.0%	3.7%	0.0%	0.0%
Time & Wind speed	0.7%	0.0%	0.0%	0.0%	1.9%	1.7%

Daily average outdoor temperature	67.8%	62.3%	66.7%	70.4%	68.5%	70.7%
Time	49.1%	52.8%	46.3%	59.3%	37.0%	50.0%
Room CO_2 concentration	42.5%	43.4%	46.3%	50.0%	29.6%	43.1%
Room air temperature	40.7%	43.4%	31.5%	48.1%	35.2%	44.8%

- Occupants open and close windows depending on drivers leading them to take action.

- Drivers can be identified through logistic regression analysis.

- The most common drivers for opening action are: time of the day and CO_2 concentration.

- The most common drivers for closing action are: outdoor temperature and time of the day.

- Thermal comfort and AIQ play a role for occupant behavior in regard to natural ventilation.