

# What influences occupants' behavior in residential buildings: An experimental study on window operation in the KTH Live-In Lab

Mahsa Farjadnia<sup>1,\*</sup>, Angela Fontan<sup>2</sup>, Alessio Russo<sup>2</sup>, Karl Henrik Johansson<sup>2</sup>, Marco Molinari<sup>1</sup>

**Abstract**—Window-opening and window-closing behaviors play an important role in indoor environmental conditions and therefore have an impact on building energy efficiency. On the other hand, the same environmental conditions drive occupants to interact with windows. Understanding this mutual relationship of interaction between occupants and the residential building is thus crucial to improve energy efficiency without disregarding occupants' comfort. This paper investigates the influence of physical environmental variables (i.e., indoor and outside climate parameters) and categorical variables (i.e., time of the day) on occupants' behavior patterns related to window operation, utilizing a multivariate logistic regression analysis. The data considered in this study are collected during winter months, when the effect on the energy consumption of the window operation is the highest, at a Swedish residential building, the KTH Live-In Lab, accommodating four occupants in separate studio apartments. Although all the occupants seem to share a sensitivity to some common factors, such as air quality and time of the day, we can also observe individual variability with respect to the most significant drivers influencing window operation behaviors.

## I. INTRODUCTION

Energy demand in the residential and commercial buildings sector in the European Union accounts for approximately 40% of total end-use energy [1] and 76% of this energy is used in comfort control in buildings using HVAC (heating, ventilation, and air conditioning) systems [2]. The behavior of building occupants, specifically window-opening actions, plays a fundamental role in the amount of energy used in buildings [3], particularly during the heating season. Previous research indicates that building energy consumption can be reduced if the control strategies of HVAC systems are optimized considering how individuals behave [4]. Therefore, efforts have been devoted over the years to developing accurate modeling of occupant behavior [5], [6], [7]. In this work, we argue that a better understanding of the drivers

that motivate occupants to interact with windows in smart buildings is significant in order to improve the efficiency of recently developed control strategies, in particular, data-driven predictive control techniques [8], [9], [10], [11].

### A. Related Work

Existing literature on window-opening behavior can be classified based on the applied theoretical methods, e.g., logistic regression [12], [13], [14], Markov processes [12], [15], deep learning methods [16], and the type of building under study, residential [17], [18] or office buildings [16], [19], [20]. A detailed review is provided in [21], where various modeling techniques are categorized based on their complexity, data requirements, and level of implementation. Logistic regression was reported among the widely implemented and relatively simple models. A remarkable contribution in modeling window behavior can be found in [12], where the authors applied logistic regression, a Markov process, and a combination of these approaches to model the relationship between the window behavior of occupants and environmental drivers in office buildings, using seven years of data. The authors concluded that indoor temperature was a significant driver in the window-opening action, and outdoor conditions best described the window-closing action.

Although most of the existing literature is focused on office buildings (for a review [3]), the interest in analyzing occupants' behavior in residential buildings is increasing. In [18] the authors assessed the window-opening action of occupants of a residential building in the UK using one-year data considering environmental drivers, time of the day, and season. The authors of [14] utilized a multilevel logistic regression model to evaluate the effect of household features, such as area and type of residents (i.e., smoker and age), in the window-opening behavior of occupants in 10 apartments. Features leading the occupants to interact with windows in 60 residential buildings are studied in [6]. Time of the day, carbon dioxide concentration, and average outside temperature were reported as the most common drivers in opening and closing the windows.

Although some of the key challenges have been identified in previous studies, it is important to emphasize that there is still no consensus on what types of drivers (physical environmental, physiological, social, and psychological) lead to an interaction between occupants and buildings. Moreover, existing literature primarily focuses on analyzing drivers of actions in multi-occupancy residential buildings, which only indicates the general behavior of the occupants and does not reveal individual differences between each occupant [22].

This work was supported by the Swedish Energy Authority and IQ Samhällsbyggnad, under the E2B2 program, grant agreement n. 47859-1 (Cost-and Energy-Efficient Control Systems for Buildings), by SSF (Swedish Foundation for Strategic Research), under grant agreement n. RIT17-0046 (CLAS—Cybersäkra lärande regelsystem), by the Swedish Research Council Distinguished Professor Grant 2017-01078, and by the Knut and Alice Wallenberg Foundation Wallenberg Scholar Grant. The KTH Live-In Lab was made possible by a donation from the Einar Mattsson-Group, whose support is kindly acknowledged. Donations from Akademiska Hus and Schneider Electric are also kindly acknowledged.

<sup>1</sup>M. Farjadnia and M. Molinari are with the Division of Applied Thermodynamics and Refrigeration, KTH Royal Institute of Technology, 100 44 Stockholm, Sweden, E-mail: {mahsa,marco}@kth.se.

<sup>2</sup>A. Fontan, A. Russo, and K. H. Johansson are with the Division of Decision and Control Systems, KTH Royal Institute of Technology, 100 44 Stockholm, Sweden, E-mail: {angfon,alessio,kallej}@kth.se.

The authors are also affiliated with Digital Futures.

\*Corresponding author.

## B. Contributions

This work aims at modeling occupants' behavior in the context of smart homes, by determining the most statistically significant drivers that can be used to identify case-specific behaviors. In particular, we focus on the interaction between occupants and windows during the winter months, when the effects of behavior patterns are most noticeable in energy consumption. To this end, we consider two models distinguishing between window-opening and window-closing actions. By building separate models, we can capture the specific characteristics associated with each action more effectively. We investigate the influence that drivers such as indoor temperature, carbon dioxide (CO<sub>2</sub>) consumption, indoor relative humidity (RH), outside temperature, and time of the day have on the window-opening and window-closing behavior of the occupants at the KTH Live-In Lab.

The KTH Live-In Lab infrastructure is designed to offer a comprehensive infrastructure (to collect, store, and visualize data) and dedicated experimental setup with state-of-the-art monitoring capability for residential buildings [23], previously available mainly (if not only) in office spaces. The results obtained in this research then contribute to the growing literature focusing on household behavior and on the dynamic interaction between occupants and the residential building. In addition, differently from the aforementioned studies (e.g., [6]) where apartments typically have not only a larger floor surface but also more than one occupant per apartment, in the experimental setup of the KTH Live-In Lab the studios have (on average) a floor surface of 20m<sup>2</sup> and are leased to a single occupant. This particular configuration allows us to study not only influential drivers in *residential* buildings, but also *heterogeneity* among tenants. The motivation is that, as tenants do not have direct control over the air quality settings, insights gained in this study regarding the individual occupant characteristics and preferences will serve as a foundation for developing customized controller designs capable to capture individual needs and comfort levels.

To model the occupant's window operation behavior we employ a logistic regression technique, following a stepwise regression procedure similar to the one illustrated in [6]. We consider the aforementioned explanatory variables, i.e., indoor and outside temperatures, RH, and CO<sub>2</sub>. In addition, we add a categorical variable to represent different day segments where different probabilities of interaction with windows (opening or closing) were observed. Our results show that the most common driver for tenants to open a window during the winter season is the day segment, followed by CO<sub>2</sub>. For window-closing behavior, the most common drivers are the day segment, and indoor and outside temperatures. The classification models are evaluated using different performance indicators, such as the ROC (receiver operating characteristic) curves and the precision-recall curves.

During data analysis and processing, we recognized typical challenges of collecting and working with real data from a testbed, which might have influenced the predicting

capabilities of the computed models. The two most common were missing data (lost in the recording process), which means that the available number of measurements in the same apartment differed between each sensor, and imbalanced data, which means that the number of times tenants interacted with windows was low (w.r.t. summer months).

Apart from these practical issues, difficulties to model the tenants' actions may also be explained by unexpected behavioral patterns that we discovered after the experimental study ended, through a discussion with the participants. For example, it was revealed that there were instances where a tenant would keep a window open, but at the same time would leave an electrical radiator on. This is consistent with representing the KTH Live-In Lab as a complex cyber-physical-human system [24], [25]. It is not the scope of this paper, but future directions will consider further challenges that arise from the interplay between the human component (i.e., the tenants) and the cyber-physical systems (i.e., the building). For instance, we plan to investigate the role and effect of social influence in environmentally significant behaviors [26].

## C. Outline

The paper is organized as follows: Section II introduces the experimental setup and gives a brief overview of the KTH Live-In Lab building, data infrastructure, and tenants; Section III describes the modeling and data preprocessing procedure; Section IV presents the two obtained models for window-closing and window-opening behavior; finally, Section V offers conclusive remarks.

## II. THE KTH LIVE-IN LAB

The KTH Live-In Lab includes several building testbeds that range from students' accommodations to lecture buildings [27]. Data collected from the sensors installed in the testbeds is stored and shared through the KTH Live-In Lab datapool (Fig. 1). The KTH Live-In Lab is an example of cyber-physical-human system, featuring a redesignable testbed layout (Section II-A), an extended sensor network (Section II-B), and an advanced interaction capability with testbed occupants (Section II-C).

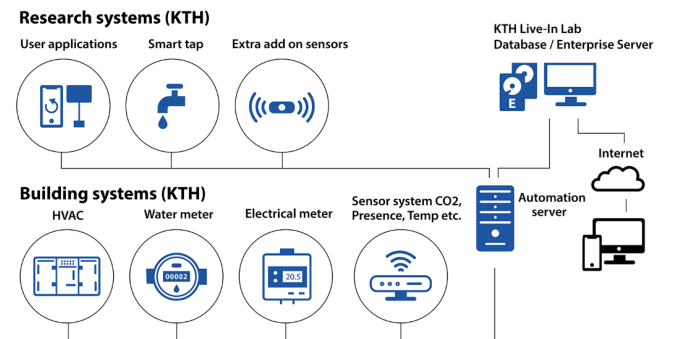


Fig. 1. Simplified illustration of the KTH Live-In Lab data infrastructure. Image source: [liveinlab.kth.se/en/infrastruktur/datapool/data-infrastructure](https://liveinlab.kth.se/en/infrastruktur/datapool/data-infrastructure).

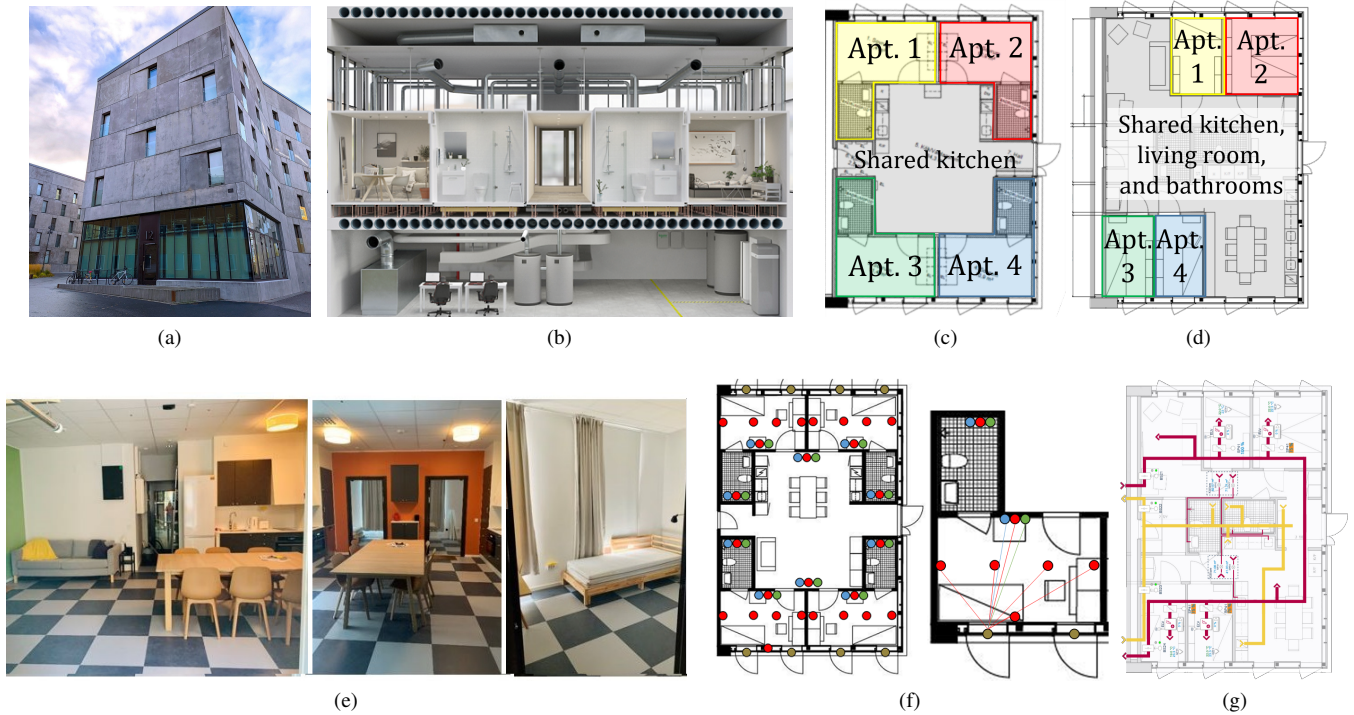


Fig. 2. (a): Outdoor view of the KTH Live-In Lab. (b): Sectional view of the KTH Live-In Lab. (c): The layout of the KTH Live-In Lab applied from August 2020 up to June 2021. (d): The current layout of the KTH Live-In Lab applied since August 2021. (e): Inside view (common living space, kitchen, bedroom) of the KTH Live-In Lab before the summer of 2021. (f) Sensors network in the whole KTH Live-In Lab with the layout as in (c) (left) and in a single apartment (right). The depicted sensors are magnetic sensors used to detect window-opening (light brown), temperature sensors (red), RH sensors (blue), and CO<sub>2</sub> sensors (green). Light, motion, and VOC sensors are not depicted. (g): Ventilation system of the current layout of KTH Live-In Lab. Red and yellow colors indicate supply and extraction air ducts, respectively. Images source: [liveinlab.kth.se/en/infrastruktur/testbed-infrastruktur](https://liveinlab.kth.se/en/infrastruktur/testbed-infrastruktur).

### A. Building Infrastructure

The overall floor area of the KTH Live-In Lab is 300m<sup>2</sup>, see Figs. 2a and 2b for an outdoor view and a sectional view of the KTH Live-In Lab, respectively, and its layout can be redesigned (Figs. 2c and 2d). The layout of the KTH Live-In Lab before the summer of 2021 comprised four independent apartments, each consisting of a living room/bedroom and a bathroom, and a shared kitchen (Figs. 2c and 2e). In the current design, the concept of shared spaces was implemented by incorporating shared living room and bathrooms in addition to the shared kitchen, while still keeping private bedrooms (Fig. 2d).

It is worth noting that the heating is provided through a ventilation system which uses an air-handling unit, with dedicated distribution outlets for each apartment (Fig. 2g). This setup allows building managers to customize control over heating and air quality settings, while tenants have limited control of the system, constrained to the heating setpoints.

### B. Data Infrastructure

The KTH Live-In Lab is equipped with advanced sensing technologies used to monitor indoor environmental parameters, such as indoor temperature, RH, CO<sub>2</sub>, and volatile organic compounds. Additional sensors include contact sensors deployed to detect, for instance, windows and doors status (open or closed), occupancy, and light sensors. These sensors

are used to study solutions that optimize the use of energy for heating and ventilation, as well as maximize the use of daylight, and improve light comfort. All measurements are taken from sensors placed at apartment level (Fig. 2f). A detailed description can be found in [23].

### C. Tenants

The tenants of the KTH Live-In Lab are undergraduate students attending KTH Royal Institute of Technology that reside at the KTH Live-In Lab full-time. Independently from the layout (Fig. 2), each apartment is leased to a single tenant.

## III. MODELING OCCUPANTS' BEHAVIOR

### A. Development and evaluation of the models for opening-window and closing-window behavior

Multivariate logistic regression was used to infer the probability of opening or closing a window based on a certain set of independent explanatory variables (e.g., indoor temperature or RH). In this paper, we use interchangeably the notation “drivers” and “explanatory variables”. The probability  $p$  of a window-opening or window-closing action is defined in terms of the logit function as:

$$\ln \left( \frac{p}{1-p} \right) = \alpha + \beta_0 x_0 + \beta_1 x_1 + \dots + \beta_n x_n, \quad (1)$$

given  $n$  explanatory variables denoted by  $x_i$ ,  $i = 1, \dots, n$ . In eq. (1),  $\alpha$  is the intercept, and  $\beta_i$  is the coefficient of the explanatory variable  $x_i$ .



For the analysis we followed a procedure similar to the one illustrated in [6]; however, differently from [6], we include an additional validation step (Fig. 3). We report the details herein for completeness. Note that the following procedure was applied twice, first using window-opening data and then using window-closing data.

First, data was partitioned: 80% of the total data set for the selection of significant explanatory variables and model training, and 20% for model validation.

Then, the training data was partitioned into 10 sub-samples for cross-validation: 9 sub-samples to be used for model training through forward and backward selection, while 1 sub-sample to be used for model validation. The procedure was repeated 10 times to utilize each sub-sample once as a test sub-sample. The Akaike Information Criterion (AIC) was used for selecting the explanatory variables in the regression model, to take into account the risk of overfitting and the risk of underfitting [6]. In practice, the selection of the most suitable model given  $n$  explanatory variables was executed by the `stepwiseglm` function in MATLAB software, which operates the following steps. First, each individual variable is fitted by a regression model and, after calculating the AIC for each fit, the single-variable model with the lowest AIC is selected. Then, the remaining variables are fitted by a bivariate regression model, and the one with the lowest AIC is chosen and compared to the selected single-variable model. If the single-variable model has a lower AIC, it is chosen as the final model. Otherwise, three-variable models are considered. While a forward-selection procedure continues to check if adding variables may lead to a model with a lower AIC, in parallel a backward-selection procedure checks if a lower AIC can be obtained by removing variables from the models. The outcome of the forward and backward selection is a model with the lowest AIC. In our case, this method was used to determine the most significant explanatory variables for window-opening or window-closing actions.

Once the significant drivers were singled out, all training data were used to fit a logistic regression model with determined explanatory variables.

Finally, the validation data was used to evaluate the goodness of fit of the estimated model, defined in terms of area under the curve (AUC) of the ROC curve and AUC of the precision-recall curve. The value of AUC is between 0 and 1, with 1 indicating perfect classification.

### B. Collection and Preprocessing of sensors data

The data used in this study was collected from October to December 2020, corresponding to the layout of the KTH Live-In Lab applied before summer 2021 (Fig. 2c). An illustration of the measured data is presented in Fig. 4. The sensors used in this experimental study are standard commercial sensors typically employed in building applications. These sensors have accuracy that is common for such purposes [6], e.g., the accuracy of the indoor and outdoor temperature sensors is within  $\pm 0.5^\circ\text{C}$  and  $\pm 0.6^\circ\text{C}$ , respectively. The data was preprocessed before utilization due to missing values, incorrect measurements, and unsynchronized data resulting

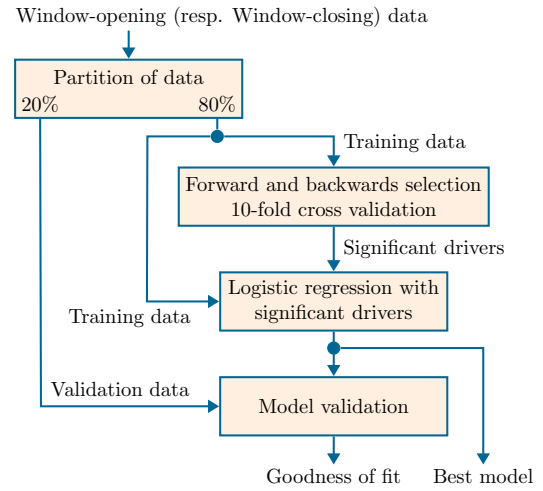


Fig. 3. Diagram of the process for selecting the best fitting window-opening (resp. window-closing) model based on monitored window-opening (resp. window-closing) data, i.e., measurements collected when the window was closed (resp. open).

from differences in the sampling rate of different sensors. After preprocessing the data, the following variables with a sample time of one minute were considered:

- Window status (open/closed) magnetic sensors [-];
- Indoor temperature ( $T_{\text{indoor}}$ ) [ $^\circ\text{C}$ ];
- Carbon dioxide concentration ( $\text{CO}_2$ ) [ppm];  $\log(\text{CO}_2)$  was used to have a better distribution for logistic regression analysis;
- Relative humidity (RH) [%];
- Outside temperature ( $T_{\text{outside}}$ ) [ $^\circ\text{C}$ ].

Based on observations of the window-opening activity in different apartments (Fig. 5), we decided to consider a categorical variable, denoted by day segment, which groups specific time periods throughout the day where difference in occupants behavior was most noticeable:

- Day segment 1 (DS1): between 06:00 and 13:00;
- Day segment 2 (DS2): between 13:00 and 22:00;
- Day segment 3 (DS3): between 22:00 and 06:00.

The daily average outside temperature was used to avoid any correlation with the day segment variable.

## IV. EXPERIMENTAL RESULTS

Following the procedure explained in Section III, we obtained a window-opening and a window-closing model for each apartment of the KTH Live-In Lab. For ease of exposition, we first illustrate the window-opening and window-closing models for apartment 2 (Section IV-A), and then we present the main results of the logistic regression analysis for all apartments (Section IV-B).

### A. Illustrative example of window-opening and window-closing models

The models for window-opening and window-closing behavior for apartment 2 are reported in eq. (2) and eq. (3),

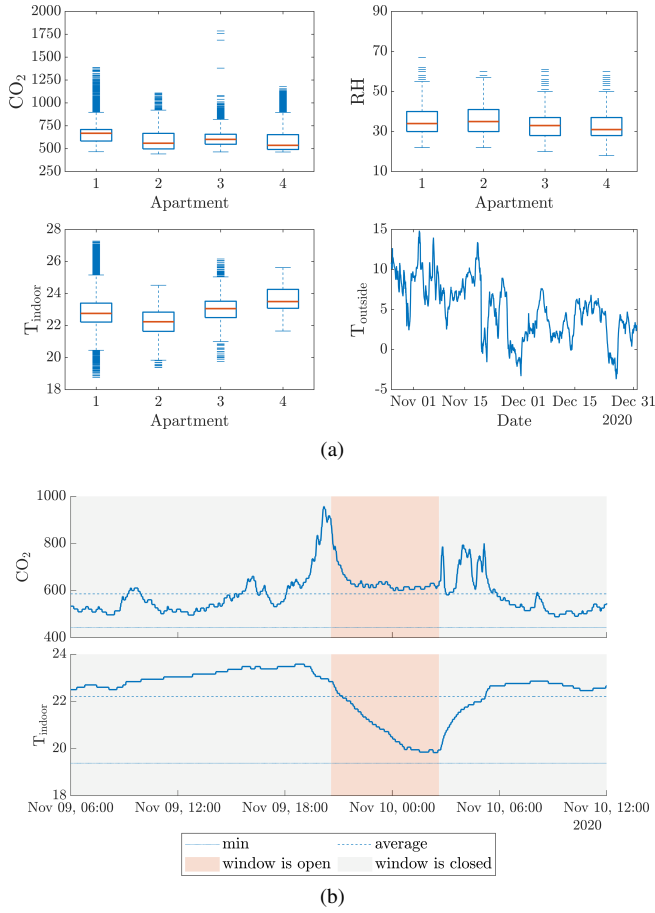


Fig. 4. (a): Monitored data for the KTH Live-In Lab apartments during the period Nov-Dec 2020: CO<sub>2</sub>, RH, indoor temperature, and outside temperature. The box is plotted from the 1st to the 3rd quartile, the red line in the boxes represents the median, and the blue horizontal lines the outliers. The outside temperature can be seen in the bottom right panel. (b): Zoomed-in detail (Nov 9-10, 2020, apartment 2) to compare CO<sub>2</sub> (top panel) and indoor temperature T<sub>indoor</sub> (bottom panel) levels with opening- (red area) and closing- (gray area) window behavior.

where the subscripts  $o$  and  $c$  stand, respectively, for opening and closing window action.

$$\ln\left(\frac{p_o}{1-p_o}\right) = \alpha_{o,DS} + \beta_{o,T_{indoor}}x_{o,T_{indoor}} + \beta_{o,\log(CO_2)}x_{o,\log(CO_2)} + \beta_{o,RH}x_{o,RH} + \beta_{o,T_{outside}}x_{o,T_{outside}}, \quad (2)$$

$$\ln\left(\frac{p_c}{1-p_c}\right) = \alpha_{c,DS} + \beta_{c,T_{indoor}}x_{c,T_{indoor}} + \beta_{c,RH}x_{c,RH} + \beta_{c,T_{outside}}x_{c,T_{outside}}. \quad (3)$$

The explanatory variables (Section III-B) are denoted by  $x_{.,T_{indoor}}$ ,  $x_{.,T_{outside}}$ ,  $x_{.,RH}$ , and  $x_{.,\log(CO_2)}$ , and the intercept  $\alpha_{.,DS}$  is selected depending on the day segment being analyzed (Fig. 5 and Section III-B for details). In this study, we decided to omit the interaction between variables to focus on the simplicity and ease of interpretation of the results.

The values of intercept and coefficients of eq. (2) and eq. (3) are listed in Table I. The data was standardized before employing logistic regression to enable a comparison

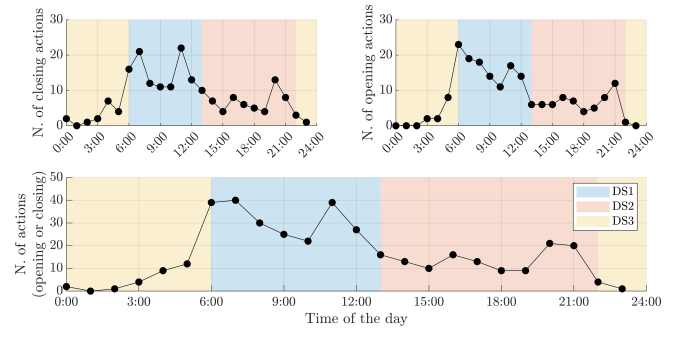


Fig. 5. Observed total N. of times the windows were closed (top left panel), opened (top right panel), and opened or closed (bottom panel) in all apartments of the KTH Live-In Lab in Nov-Dec 2020, with respect to the time of the day. The three day segments (DS) are defined based on the probability of occupants interacting with a window (opening or closing action): DS1 = [6:00-13:00], DS2 = [13:00-22:00], and DS3 = [22:00-6:00] (Section III-B). The probability of occupants interacting with a window is the highest during DS1, and the lowest during DS3.

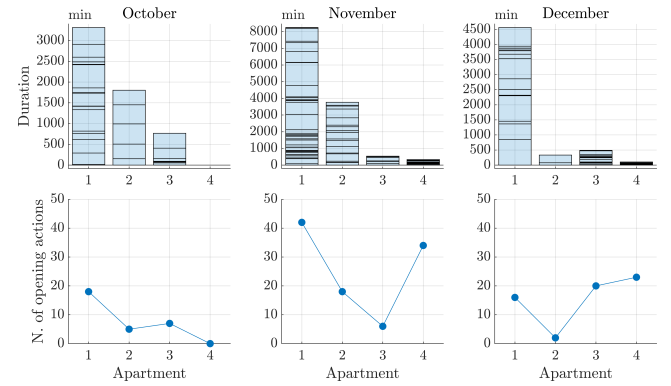


Fig. 6. Observed total duration of open windows in number of minutes (top panels) and total N. of times the windows were open (bottom panels) in all apartments of the KTH Live-In Lab in Oct-Dec 2020, with respect to the month. In the top panels, horizontal black lines in each bar highlight the single durations of open window for each apartment across one month.

between the explanatory variables and draw conclusions on their influence on the occupant's behavior. A positive (resp. negative) coefficient indicates a positive (resp. negative) correlation between the associated driver and the opening or closing window action. Figure 7 illustrates the goodness of fit of the models in eqs. (2) and (3). The main key driver characterizing the behavior of the resident of apartment 2 is the time of the day, implying unlikeliness to open the windows during the night segment (DS3, top panel of Fig. 8), and indicating a preference to close the windows in the morning (DS1, bottom panel of Fig. 8).

The developed model for window-opening action demonstrates a high sensitivity of the occupant to CO<sub>2</sub> (poor air quality drives the occupant towards window-opening behavior) and susceptibility to T<sub>outside</sub> (a warmer outside temperature drives the occupant towards window-opening action and hinders towards window-closing action). The negative correlation between indoor temperature and window-opening action could be explained by the occupant's attitude of keeping the indoor temperature roughly constant and below a potential discomfort threshold. Indeed, the range

TABLE I

ESTIMATED INTERCEPTS AND COEFFICIENTS OF THE WINDOW-OPENING (EQ. (2)) AND WINDOW-CLOSING (EQ. (3)) MODELS OF APARTMENT 2.

Drivers	Opening action			Closing action		
	Coeff.	Conf. interval 2.5%	Conf. interval 97.5%	Coeff.	Conf. interval 2.5%	Conf. interval 97.5%
$\alpha_{DS1}$	-1.729	-1.776	-1.682	1.184	0.936	1.432
$\alpha_{DS2}$	-0.242	-0.271	-0.212	0.023	-0.099	0.143
$\alpha_{DS3}$	-3.720	-3.793	-3.646	-0.728	-0.873	-0.583
$\beta_{T_{\text{indoor}}}$	-0.627	-0.664	-0.591	-0.244	-0.305	-0.184
$\beta_{\log(\text{CO}_2)}$	1.095	1.073	1.117	-	-	-
$\beta_{RH}$	0.299	0.249	0.345	0.564	0.4394	0.690
$\beta_{T_{\text{outside}}}$	0.980	0.926	1.030	-0.554	-0.686	-0.424

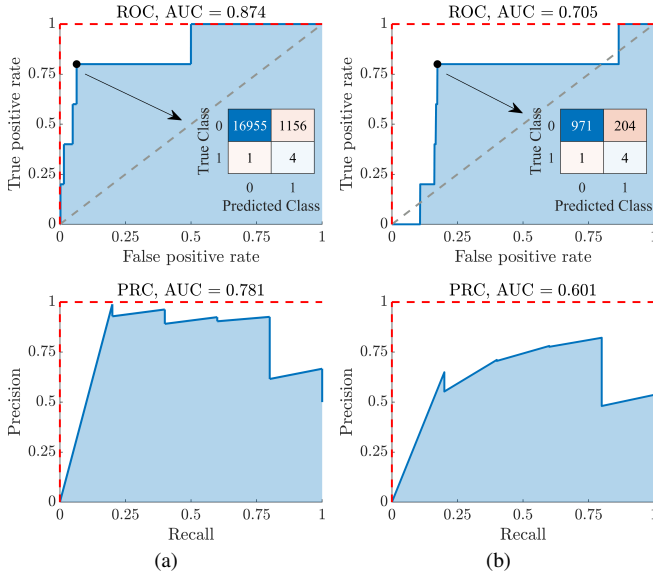


Fig. 7. Goodness of fit for the models of apartment 2 (Section IV-A). The top panels show the ROC curves and confusion charts (inset panel), presented for the optimal operating point indicated by the black symbol. The bottom panels show the precision-recall curves. Red dashed lines indicate the perfect classifier, and gray dashed lines the baseline classifier. (a): Opening-window model. (b): Closing-window model.

$T_{\text{indoor}}$  belongs to in apartment 2 is lower compared to other occupants in KTH Live-In Lab (Fig. 4a).

### B. Window-opening and window-closing models of all apartments in the KTH Live-In Lab

This section provides an overview of the drivers influencing the behavioral patterns of the occupants at the KTH Live-In Lab, focusing in particular on window-opening actions as clearer trends can be observed across the models. Our findings support the results found in [6], [13]; for readability, we first discuss the results corresponding to the window-opening models and then to the window-closing models.

The magnitude of the effect of physical environmental drivers on occupants' window-opening pattern varies in different apartments and can be interpreted in terms of different habits and personal preferences. For instance, occupants of apartment 1 and apartment 3 seem to be more conscious of the indoor temperature while the occupants of apartment 2 and apartment 4 appear to be more sensitive to the  $\text{CO}_2$

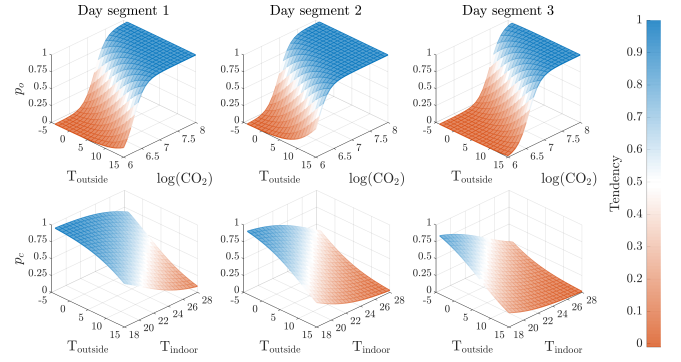


Fig. 8. Probability of window-opening action w.r.t. outside temperature and  $\log(\text{CO}_2)$  (top panels), and probability of window-closing action w.r.t. indoor and outside temperatures (bottom panels), during different day segments, i.e., morning (DS1), afternoon (DS2), and night (DS3). The surfaces are color-coded based on the tendency of the occupant for apartment 2 towards window-opening or window-closing actions.

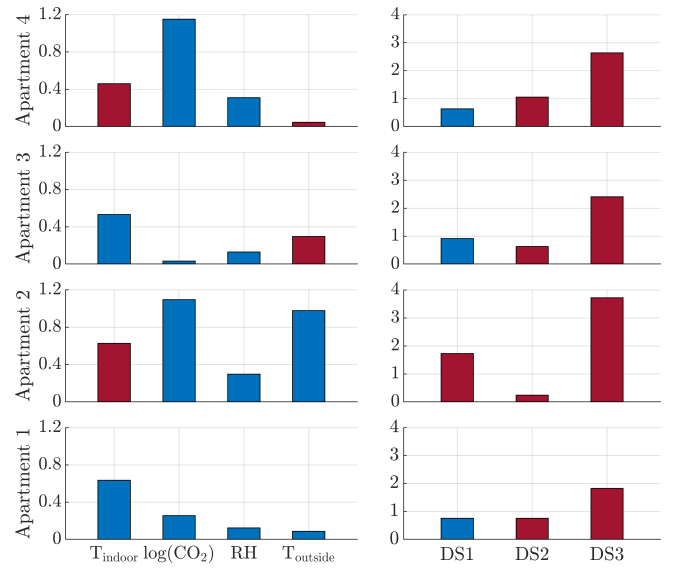


Fig. 9. Coefficients and intercept of drivers for window-opening action of all apartments. The blue (resp. red) color indicates a positive (resp. negative) correlation with the window-opening action.

concentration levels. However, all the developed models suggest that the occupants are sensitive to air quality (i.e.,  $\text{CO}_2$  and RH). Further, none of the occupants seems prone to open windows during the night, preferring instead (in general) to open windows in the morning (Fig. 9).

Similar to the window-opening case, occupants' window-closing behavior varies w.r.t. different physical environmental stimuli, while a common factor influencing all residents' behavior is the time of the day. Further, the occupants of apartments 2 and 3 seem both sensitive to indoor and outside temperatures; in particular, the lower the indoor and outside temperatures, the higher the tendency towards a window-closing action. However, for the apartments 1 and 4, the AUC of the ROC curve and precision-recall curve are lower than those obtained in the window-opening case, indicating that for these apartments the models are less intuitive, and no significant driver is shown apart from the time of the day.

*Remark 1:* A practical challenge is that our data is highly imbalanced, primarily due to the data period under consideration (winter months), which makes the analysis more complex. Although we considered a 10-fold cross-validation procedure and a balanced weight approach to improve the performance of the logistic regression method, we need to be cautious when evaluating the reliability of the obtained models, even if we obtain high values of AUC of ROC and precision-recall curves (e.g., over 60% for the models in eqs. (2) and (3) for apartment 2). In addition, a possible cause for the encountered modeling challenges is the specific layout of the KTH Live-In Lab considered, where the presence of a common area affects the occupancy of apartments.

## V. CONCLUSION

In this work, we focused on analyzing occupants' behavior patterns related to window operation using real data from a Swedish residential facility, the KTH Live-In Lab. The influence of physical environmental (i.e., indoor and outside temperatures, air quality) and categorical (i.e., time of the day) drivers is captured particularly in the study of the window-opening actions. The findings suggest that, each individual occupant reacts differently to environmental stimuli, while all of the occupants are sensitive to the air quality. Data analysis also indicates a strong correlation between the time of the day and the occupants' tendency to interact with windows (probably related to daily activities or habits).

This study excluded certain environmental drivers, such as seasonality and solar irradiation, as well as, physiological or psychological drivers, which we plan to take into account in future works, in combination with an extended data collection period. Additionally, future work includes using more advanced statistical methods to deal with heavily imbalanced datasets, as well as analyzing the occupants' behaviors with other techniques (e.g., bayesian methods, and conformal prediction methods).

The analysis in this paper is based on data collected during the winter months when window-opening actions highly affect energy consumption. Hence, this work will serve as the foundation to include the dynamical behavior of occupants in designing a data-driven predictive controller [10] for HVAC systems in smart buildings to improve energy efficiency while maintaining individual occupants' thermal comfort.

## REFERENCES

- [1] European Commission, Directorate General for Energy, "Clean Energy for All Europeans," [data.europa.eu/doi/10.2833/21366](https://data.europa.eu/doi/10.2833/21366).
- [2] L. Pérez-Lombard, J. Ortiz, and C. Pout, "A review on buildings energy consumption information," *Energy and buildings*, vol. 40, no. 3, pp. 394–398, 2008.
- [3] V. Fabi *et al.*, "Occupants' window opening behaviour: A literature review of factors influencing occupant behaviour and models," *Building and Environment*, vol. 58, pp. 188–198, 2012.
- [4] X. Dai, J. Liu, and X. Zhang, "A review of studies applying machine learning models to predict occupancy and window-opening behaviours in smart buildings," *Energy and Buildings*, vol. 223, p. 110159, 2020.
- [5] V. M. Barthelmes *et al.*, "Exploration of the bayesian network framework for modelling window control behaviour," *Building and Environment*, vol. 126, pp. 318–330, 2017.
- [6] D. Calì *et al.*, "Analysis of occupants' behavior related to the use of windows in German households," *Building and Environment*, vol. 103, pp. 54–69, 2016.
- [7] T. Hong *et al.*, "Advances in research and applications of energy-related occupant behavior in buildings," *Energy and buildings*, vol. 116, pp. 694–702, 2016.
- [8] M. Khosravi *et al.*, "Data-driven predictive control of buildings; a regression based approach," in *IEEE Conference on Control Technology and Applications (CCTA)*, 2019, pp. 605–610.
- [9] L. Di Natale *et al.*, "Lessons learned from data-driven building control experiments: Contrasting gaussian process-based mpc, bilevel deepc, and deep reinforcement learning," in *2022 IEEE 61st Conference on Decision and Control (CDC)*. IEEE, 2022, pp. 1111–1117.
- [10] M. Farjadnia *et al.*, "Robust data-driven predictive control of unknown nonlinear systems using reachability analysis," *European Journal of Control*, accepted, 2023.
- [11] Y. Yang, G. Hu, and C. J. Spanos, "HVAC energy cost optimization for a multizone building via a decentralized approach," *IEEE Transactions on Automation Science and Engineering*, vol. 17, no. 4, pp. 1950–1960, 2020.
- [12] F. Haldi and D. Robinson, "Interactions with window openings by office occupants," *Building and Environment*, vol. 44, no. 12, pp. 2378–2395, 2009.
- [13] R. Andersen *et al.*, "Window opening behaviour modelled from measurements in danish dwellings," *Building and environment*, vol. 69, pp. 101–113, 2013.
- [14] S. Shi *et al.*, "Effects of household features on residential window opening behaviors: A multilevel logistic regression study," *Building and environment*, vol. 170, p. 106610, 2020.
- [15] M. Schweiker *et al.*, "Verification of stochastic models of window opening behaviour for residential buildings," *Journal of Building Performance Simulation*, vol. 5, no. 1, pp. 55–74, 2012.
- [16] R. Markovic *et al.*, "Window opening model using deep learning methods," *Building and Environment*, vol. 145, pp. 319–329, 2018.
- [17] J. Rouleau and L. Gosselin, "Probabilistic window opening model considering occupant behavior diversity: A data-driven case study of canadian residential buildings," *Energy*, vol. 195, p. 116981, 2020.
- [18] R. V. Jones *et al.*, "Stochastic behavioural models of occupants' main bedroom window operation for uk residential buildings," *Building and Environment*, vol. 118, pp. 144–158, 2017.
- [19] R. Markovic *et al.*, "Comparison of different classification algorithms for the detection of user's interaction with windows in office buildings," *Energy Procedia*, vol. 122, pp. 337–342, 2017.
- [20] N. Li *et al.*, "Probability of occupant operation of windows during transition seasons in office buildings," *Renewable Energy*, vol. 73, pp. 84–91, 2015.
- [21] B. Dong *et al.*, "Modeling occupancy and behavior for better building design and operation—a critical review," in *Building Simulation*, vol. 11. Springer, 2018, pp. 899–921.
- [22] Y. Jian *et al.*, "Individual-based insight into occupants' interaction with windows in apartments in beijing," *Journal of Building Engineering*, vol. 56, p. 104795, 2022.
- [23] D. Rolando, W. Mazzotti Pallard, and M. Molinari, "Long-term evaluation of comfort, indoor air quality and energy performance in buildings: The case of the KTH Live-In Lab Testbeds," *Energies*, vol. 15, no. 14, p. 4955, jul 2022.
- [24] A. M. Annaswamy *et al.*, Eds., *Cyber-Physical-Human Systems: Fundamentals and Applications*. UK: Wiley, in Press, 2022.
- [25] F. Lamnabhi-Lagarrigue *et al.*, "Systems & Control for the future of humanity, research agenda: Current and future roles, impact and grand challenges," *Annual Reviews in Control*, vol. 43, pp. 1–64, 2017.
- [26] A. Fontan *et al.*, "Social interactions for a sustainable lifestyle: The design of an experimental case study," in *22nd IFAC World Congress*, Yokohama, Japan, 2023.
- [27] M. Molinari *et al.*, "Using living labs to tackle innovation bottlenecks: the KTH Live-In Lab case study," *Applied Energy*, vol. 338, no. 120877, 2023.