Indoor Climate Profiles in Danish Apartments based on High Temporal Resolution Measurements

Ricardo Forgiarini Rupp1, Jørn Toftum1, Gianluca Trotta2, Rune Korsholm Andersen1
1International Centre for Indoor Environment and Energy, Department of Civil Engineering, Technical University of Denmark, Kgs. Lyngby, Denmark, rrifo@byg.dtu.dk
2Department of the Built Environment, The Faculty of Engineering and Science, Aalborg University Copenhagen, Copenhagen, Denmark

Abstract
The aim of this work is to define profiles of the indoor climate in Danish multifamily residential buildings. Measurements of indoor temperature, relative humidity and CO2 concentration were conducted every five minutes for more than a year in 45 apartments using internet of things sensing devices. Cluster analysis was employed to derive the indoor climate profiles. The relationship between indoor and outdoor climate was analysed through piecewise linear regression. Three different profiles of indoor climate were identified, as well as heating schedule and setpoint temperatures, which may be considered when modelling occupant behaviour and performing building energy simulation.

Key Innovations
- Internet of things (IoT) sensing devices employed to acquire high temporal resolution measurements
- Piecewise linear regression analysis performed in clusters of the data to extract useful information for building simulation

Practical Implications
Use of hourly data to build indoor climate profiles. Avoid using only one profile of indoor climate, heating setpoint temperature or heating schedule.

Introduction
Building simulation relies on realistic input data of high quality in order to provide high quality and realistic results. No matter how good the simulation software and the modelling capabilities, the simulation results will not be realistic if the occupancy profiles, setpoints, loads etc. are not realistic. Understanding occupancy patterns and occupants’ control of the indoor climate is crucial to perform accurate and realistic building simulations (Yoshino et al., 2017; Yan et al., 2017; O’Brien et al., 2020). Testing different design scenarios in a building performance simulation software can significantly contribute to the design of buildings with good indoor environment and low energy use. Building simulation can be used in an optimisation process aimed at maximising comfort, health and productivity conditions and minimising energy use. But if the optimisation process is based on unrealistic conditions in terms of occupancy and heating setpoints, the building will end up being optimised for unrealistic conditions and may not work optimally, once built. As a consequence, there is a need for realistic occupancy patterns and realistic temperature setpoints that are easy to implement in existing simulation software. Some authors provided stochastic models of occupants’ heating setpoints based on measurements e.g. in 15 Danish dwellings (Fabi, Andersen, and Corgnati 2013; D’Oca et al. 2014). While these may be capable of replicating the diversity amongst occupants and the temporal variation in the occupants’ behaviour due to the stochastic nature, they are difficult to implement and require expert knowledge. Since the models are stochastic, the simulation results are harder to interpret and they require several simulations, resulting in longer simulation times.

In this work, data from a large field campaign in residential buildings in Denmark using IoT sensing devices was used in order to acquire information on indoor environmental quality. To our knowledge, the database is the most detailed Danish sample of indoor environment data in residential buildings. Other databases from commercial stakeholders exist, but the data come from sensors installed by the users themselves with risk of placing them in unsuitable places – e.g. in direct sunlight or close to a heating source. The aim of this work was to define profiles of the indoor climate in Danish multifamily residential buildings. However, in this work, we analysed data only from a residential building in Copenhagen. The high temporal resolution of the measurements may provide a better understanding of the link between indoor and outdoor temperatures at different seasons. Results may be used to better understand residents’ domestic temperature preferences and guide building simulation practitioners towards more realistic simulation results.

Method
Field data collection was performed in a multifamily residential building in Copenhagen, Denmark, totalling 45 apartment units. Cluster analysis was used for discovering distinct apartment groups characterized by different indoor profiles, encompassing air temperature, relative humidity, and CO2 concentration. In our empirical setup, the main advantage of cluster analysis lies in its ability to account for different occupancy patterns and bridge the gap between predicted and actual building energy performance. Furthermore, the derived
indoor profiles were used to investigate the relationship between indoor and outdoor temperature by employing piecewise linear regression. This approach allowed us to identify a threshold for temperature for each cluster based on the apartments’ operation mode (heating and non-heating periods).

The surveyed building
The 3-storey building chosen as a case study for this work is a typical multifamily residential building in Copenhagen with brick wall façade and ceramic roof tiles. Most of the 45 apartment units have a floor area of 66 m² with few larger units up to 93 m². All apartments are heated by radiators/convectors connected to a central water-based heating system and equipped with operable windows. The temperature is controlled using thermostatic radiator valves and by opening windows.

Data collection
Internet of things (IoT) sensing devices were installed in a central corridor of each apartment to record the indoor air temperature, relative humidity and CO₂ concentration. The standardization of the location of the devices in each apartment has the benefit of avoiding installing the equipment in unsuitable places, but it may not accurately represent the indoor climate of a specific room. Measurements were conducted every five minutes from April 2019 to November 2020. Hourly outdoor air temperature and relative humidity were taken from the Norwegian Meteorological Institute (2020).

Data analysis
Rigorous data cleansing and processing were carried out for the statistical analysis. Data analysis was conducted in R (R Core Team, 2021). The relation between indoor and outdoor temperature and humidity was investigated. Cluster analysis was employed in order to derive the indoor climate profiles. The idea was to group the different apartments based on the indoor temperature and estimate different indoor climate profiles for each cluster. The hierarchical cluster analysis technique was employed using the Euclidean distance as similarity measure and the Ward method as the linkage function. The hierarchy tree was represented by a dendrogram. The number of clusters was determined by analysing the growth of within-cluster variance according to the number of clusters – the so-called elbow method.

Daily profiles of indoor climate were derived for each cluster in different seasons. Profiles were built separately for weekdays and weekends. We also investigated the influence of using data at 5-min intervals and its hourly mean (average of 5-min data for each hour) on the determination of daily indoor climate profiles.

Piecewise regression analysis was conducted between the hourly indoor and outdoor temperature for each apartment. The data is composed of temperature measurements during the heating and non-heating periods, where the apartments were free running, i.e. the indoor temperatures are likely to be disconnected from outdoor temperatures during the heating operation; indoor temperatures tend to fluctuate according to the outdoor temperatures during the free-running mode. For this situation, fitting a linear regression model to the data (indoor vs. outdoor air temperature) would not describe the data very well. Piecewise regression allows fitting two linear models to the data for different ranges of the explanatory variable (outdoor air temperature). Piecewise regression has been applied by Rasmussen et al. (2020) to analyse the relationship between heating consumption and outdoor air temperature and by Nguyen et al. (2014) to analyse the relationship between the indoor and outdoor air temperature in homes located in Boston, USA. The point (i.e. the value of outdoor air temperature) where a change in the slope of the linear models occurs, is called a breakpoint. In this work, the breakpoint is an indication of the apartments’ operation mode: heating mode below the breakpoint (lower outdoor air temperatures) and free-running mode above the breakpoint.

Results
Almost 7 million measurements of environmental parameters were collected in the multifamily residential building in Copenhagen.

Outdoor and indoor climate
Table 1 presents a summary of the indoor thermal environment in the surveyed building. Overall, the environmental variables were normally distributed. However, considerable discrepancies in indoor environmental parameters were observed across the apartments, especially concerning the CO₂ concentration (high values above 1,000ppm).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Number of data points</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indoor air temperature (ºC)</td>
<td>6,845,774</td>
<td>22.7</td>
<td>1.9</td>
</tr>
<tr>
<td>Relative humidity (%)</td>
<td>6,845,774</td>
<td>45</td>
<td>9</td>
</tr>
<tr>
<td>CO₂ concentration (ppm)</td>
<td>6,819,208</td>
<td>838</td>
<td>453</td>
</tr>
</tbody>
</table>

Figure 1 shows the hourly outdoor and indoor air temperatures for the entire measurement period for the 45 apartments in the residential building. The hourly indoor air temperature ranged from 16.7 ºC in winter to 28.8 ºC in summer, while the hourly outdoor air temperature varied between -2.3 ºC and 30.0 ºC, respectively. The distribution of indoor air temperatures was marked by two periods: the heating period and the non-heating period. It is interesting to note the fluctuations of indoor air temperature according to the outdoor temperature, when the building was in the free-running mode, noticeably in summer (June to August).
The hourly outdoor and indoor relative air humidity in the surveyed building is shown in Figure 2. The hourly outdoor relative air humidity oscillated widely, ranging from 16.3% to 99.1%. The period from October 2019 to February 2020 was humid outdoors, but higher temperature indoors caused a lower stable indoor humidity. It is clear from Figures 1 and 2 that the building envelope and heating system played a strong role in maintaining both indoor air temperature and humidity within a narrower range than outdoors.

**Cluster analysis**

The outcome of the hierarchical cluster analysis can be seen in Figure 3. In Figure 3, the y-axis presents the dissimilarity between elements/clusters (i.e. a low value indicates similar elements). The x-axis in Figure 3 displays the 45 apartments, which were combined according to their similarities. Figure 3 shows the formation of three clusters. The dendrogram also shows that cluster #2 and cluster #3 were more similar to each other than to cluster #1.

**Indoor climate profiles**

Figures 4 to 9 present the daily indoor climate profiles for each cluster during weekdays in summer and winter. The differences in air temperature, relative humidity and CO₂ concentration between summer and winter were rather clear.

In summer, the air temperature followed the outdoor conditions, rising in the morning until midday (peak) and lowering afterwards (Figure 4). The curves also corroborate the findings from the cluster analysis pointing out that cluster #2 and cluster #3 were more similar compared with cluster #1. In winter, the indoor air temperature did not vary over time as much as in summer (Figure 7). It also diverged more between clusters than in summer, which indicates different occupant behaviour towards the setpoint of the heating system (the setpoint appeared to be around 20.5°C for cluster #1, 21.5°C for cluster #2 and ≈23.0°C for cluster #3). The distinct patterns of heating use may be triggered by several factors. Residential space heating demand is typically price and income inelastic and mainly associated with dwelling attributes, such as dwelling age, tenure, floor area, but also with socio-demographic characteristics such as household size and composition (e.g., Safari and Javid, 2016; Brounen et al., 2012; Alberini et al., 2011; Braun, 2010). Another strand of literature points to the role of everyday practices and perceived comfort in shaping space heating demand (e.g., Gram-Hanssen, 2010).

No major differences between the clusters were observed when analysing only the daily relative humidity profiles (Figures 5 and 8). Again, the thermal conditions (relative humidity) were more stable during winter than in summer partly due to the heating system, since raising the indoor temperature decreases the relative air humidity. The hourly mean relative humidity values were close to 40% in winter, while they were close to 50% in summer.
We suppose that occupant behaviour towards the opening of windows in summer, increasing the air exchange rate, provided lower levels of CO\textsubscript{2} concentration, in comparison with the winter, when windows were potentially kept closed. In winter, higher values of the CO\textsubscript{2} concentration could be seen (above 1,000 ppm, the threshold adopted by the Danish building regulations, 2018), which is concerning since such values are the mean of the hour. This indicates an inefficient removal of pollutants (and moisture) through ventilation.

The CO\textsubscript{2} concentration values provided a possibility to make an overall assessment of the occupancy. Occupants

### Table 2: Mean±S.D. of differences of indoor climate profiles between weekdays and weekends.

<table>
<thead>
<tr>
<th>Season</th>
<th>Cluster</th>
<th>Ta (ºC)</th>
<th>RH (%)</th>
<th>CO\textsubscript{2} (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summer</td>
<td>#1</td>
<td>-0.03±0.05</td>
<td>-2.5±1.0</td>
<td>-1.7±5.8</td>
</tr>
<tr>
<td></td>
<td>#2</td>
<td>-0.02±0.05</td>
<td>-2.0±0.8</td>
<td>-6.9±17.3</td>
</tr>
<tr>
<td></td>
<td>#3</td>
<td>-0.01±0.04</td>
<td>-1.8±0.7</td>
<td>-12.8±11.5</td>
</tr>
<tr>
<td>Winter</td>
<td>#1</td>
<td>0.04±0.05</td>
<td>0.0±0.2</td>
<td>-17.2±31.9</td>
</tr>
<tr>
<td></td>
<td>#2</td>
<td>0.06±0.02</td>
<td>0.1±0.3</td>
<td>-20.7±43.7</td>
</tr>
<tr>
<td></td>
<td>#3</td>
<td>-0.10±0.03</td>
<td>0.0±0.2</td>
<td>-37.1±41.4</td>
</tr>
</tbody>
</table>
typically left the home in the morning and returned in the late afternoon, but dissimilar profiles were identified in the database. Lower CO₂ concentration values were measured in cluster #1 than in the other groups for both seasons. We do not have available evidence to explain it; we can only speculate that, probably, this was caused by a higher ventilation/infiltration rate (the indoor air temperature was lower) or less occupants/occupancy (which would help to explain the flat pattern over time in winter).

We also built profiles of indoor climate for the weekends. Table 2 shows the differences between weekdays and weekends for each environment variable, season and cluster. Overall, the differences in indoor air temperature and relative humidity were not considerable. On the other hand, the CO₂ concentration values during weekends differed considerably from the weekdays – the only exception being cluster #1 in summer. The higher CO₂ levels during weekends could be caused by occupants spending more time at home, particularly in winter. These well-defined patterns of CO₂ levels with respect to seasons, weekdays, and weekends are similar to those found in residential electricity use (Trotta, 2020).

Finally, we assessed the influence of data resolution on the daily profiles of indoor climate. The overall mean±S.D. of the difference between the hourly and 5-min measurement values are -0.01±0.01 for air temperature, -0.01±0.03 for relative humidity and 1.72±1.81 for CO₂ concentration. This way indicating that it is not necessary to use data at 5-min intervals in order to build daily indoor climate profiles, i.e. the hourly mean data provided similar results. The required computation time did not change considerably between each other due to the size of the database used in this work (6,845,774 datasets, ≈ 900MB). However, when working with larger datasets it is recommended to use the hourly mean of the 5-min data to derive daily indoor climate profiles.

**The relationship between indoor and outdoor climate**

Figures 10 to 12 show the relationship between the hourly indoor air temperature and the outdoor air temperature for three apartments (one per cluster). Overall, our assumptions regarding the different relations between indoor and outdoor climate depending on the operation mode are confirmed by the results of the piecewise regression analysis.

During the heating period, the slopes of the (blue) linear models were close to zero (0), indicating a weak association between the indoor and the outdoor air temperature, as expressed by the low values of the correlation coefficient (r). In contrast, the slopes of the (orange) linear models were around 0.32-0.38 during the free-running mode and both variables were highly correlated (r > 0.7) - it is evident that indoor temperatures increased with increasing outdoor air temperature. Similar results were found for the piecewise regression analysis when considering the 45 apartments - Table 3 presents the mean ± S.D. of the slopes for the heating and the free-running mode.

![Figure 10: Piecewise regression between hourly indoor and outdoor air temperature for an apartment - Cluster #1. Model intercept=20.44, R²=0.65, N=13,992. Dashed line (y = x).](image1)

![Figure 11: Piecewise regression between hourly indoor and outdoor air temperature for an apartment - Cluster #2. Model intercept=22.28, R²=0.50, N=14,035.](image2)

![Figure 12: Piecewise regression between hourly indoor and outdoor air temperature for an apartment - Cluster #3. Model intercept=21.25, R²=0.55, N=11,937.](image3)
Our results are comparable to the ones found for the heating period (slope = 0.04, \( r = 0.40 \)) and the free-running mode (slope = 0.41, \( r = 0.91 \)) by Nguyen et al. (2014) in homes located in Greater Boston (USA). However, they used the daily means of indoor and outdoor air temperature and obtained a higher correlation between the indoor and outdoor air temperature than in our study.

The piecewise regression analysis identified the mean breakpoints as 11.1\(^\circ\)C, 13.6\(^\circ\)C and 12.5\(^\circ\)C for cluster #1, #2 and #3, respectively (Table 3). This might illustrate the different behaviours taken regarding the heating system: some people use heating earlier (e.g. cluster #2, Figure 11, breakpoint = 13.4\(^\circ\)C) than others (e.g. cluster #1, Figure 10, breakpoint = 10.6\(^\circ\)C). Such behaviours could be related to their individual thermal preferences, with cluster #2 occupants preferring a higher indoor temperature than the others – the mean ± S.D. of indoor air temperatures was 20.6\(^\circ\)C ± 1.3 (cluster #1), 22.3 \(^\circ\)C ± 1.2 (cluster #2) and 21.5 \(^\circ\)C ± 1.1 (cluster #3), when considering the data below the breakpoint in Figures 10-12. The different indoor temperatures could also be a consequence of the temperatures in the adjacent apartments (Cali et al., 2016), the thermal inertia of the building, the window orientation, and other internal loads. However, we do not have this information about the apartments. These findings are consistent with the daily profiles of indoor air temperature, as already shown in Figures 4 and 7. The mean values of indoor air temperature (i.e. 20.6-22.3 \(^\circ\)C) may be used as the heating setpoint for building energy simulations. The breakpoints may be used as a reference for the modelling of the operation (schedule) of heating systems within the context of residential buildings in Denmark, i.e. the apartments operate in heating mode when the outdoor temperature is below the breakpoint (range between 11.1 and 13.6 \(^\circ\)C).

For the free-running mode, the dashed line (\( y = x \)) in Figures 10-12 shows that most indoor temperature values were higher than the outdoor temperature. If we consider 26\(^\circ\)C of indoor air temperature as the overheating threshold, the measured indoor thermal conditions were above the threshold 7.6% (cluster #1), 15.5% (cluster #2) and 6.5% (cluster #3) of the time (Figures 10-12). Again, cluster #2 occupants tended to experience higher indoor temperatures than the others.

**Conclusion**

A comprehensive database on indoor environmental quality was built and analysed in this work, which provided a better understanding of the indoor climate in multifamily residential buildings. Some degree of overheating and high CO\(_2\) concentration (above 1,000ppm in winter) were observed in the apartments.

Cluster analysis indicated three different profiles of indoor climate. One should consider different profiles when modelling occupant behaviour. No considerable differences in the profiles were found between weekdays and weekends considering the indoor air temperature and the relative humidity. However, CO\(_2\) concentration values during weekends were higher than during weekdays. We found that it was sufficient to build daily indoor climate profiles based on hourly data rather than 5-minute values. Further studies should be conducted to analyse the need for high temporal resolution measurements for other applications, such as to estimate occupancy and/or air change rates.

In the apartment, the transition between the heating operation and the free-running mode happened between 11.1 to 13.6 \(^\circ\)C of outdoor air temperature. The heating setpoints were around 20.6-22.3 \(^\circ\)C. These values are useful to feed building simulations in order to achieve more realistic predictions of energy use and indoor environmental conditions. Examination of these transition values and relationships between indoor and outdoor temperature using a larger and representative sample could allow the generalization of the results.

**Acknowledgement**

The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement no. 713683 (COFUNDfellowsDTU).

**References**


---

**Table 3: Summary of the piecewise regression analysis for the 45 apartment units.**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Cluster</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heating slope</td>
<td>#1</td>
<td>0.033</td>
<td>0.107</td>
</tr>
<tr>
<td></td>
<td>#2</td>
<td>0.001</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>#3</td>
<td>0.043</td>
<td>0.070</td>
</tr>
<tr>
<td>Free-running slope</td>
<td>#1</td>
<td>0.354</td>
<td>0.108</td>
</tr>
<tr>
<td></td>
<td>#2</td>
<td>0.324</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td>#3</td>
<td>0.343</td>
<td>0.048</td>
</tr>
<tr>
<td>Breakpoint</td>
<td>#1</td>
<td>11.1</td>
<td>2.8</td>
</tr>
<tr>
<td></td>
<td>#2</td>
<td>13.6</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>#3</td>
<td>12.5</td>
<td>2.3</td>
</tr>
</tbody>
</table>


