A Hidden Markov-switching Building Occupancy Model

Sebastian Wolf*, Magnus Bitsch, Henrik Madsen
Technical University of Denmark, Department of Applied Mathematics and Computer Science, Kgs. Lyngby, Denmark
* Corresponding author, Email: sewo@dtu.dk

1 Introduction
Heating and ventilation strategies in buildings can be improved significantly if information about the current presence and activity status of the occupants is taken into account. Therefore, there is a high demand for inexpensive sensor-based methods to detect the occupancy status. This study suggests a new occupancy model based on the use of CO\textsubscript{2} trajectories, trained on measurements in class rooms of two schools in Denmark. A Hidden Markov-switching Model with autoregressive observations using a normal state dependent distribution was employed to identify the occupancy states.

2 Methods
The model consists of an underlying inhomogeneous Markov chain of states X\textsubscript{t} and a series of observations Y\textsubscript{t} = CO\textsubscript{2}. For each state \( i \) that \( X\textsubscript{t} \) takes at time \( t \), \( Y\textsubscript{t} \) follows a normal distribution. Furthermore, the observations follow a auto-regressive process.

\[ Y\textsubscript{t} \sim N(c\textsubscript{i} + \phi \cdot Y\textsubscript{t-1} + \epsilon\textsubscript{t}, \sigma^2) \]

The model can be expressed by

\[ p(X\textsubscript{t} = i | X\textsubscript{t-1} = j) \sim (\Gamma\textsubscript{ij}) \]

where \( \Gamma\textsubscript{ij} \) is a transition probability matrix, \( c\textsubscript{i} \) are the state means, \( \phi \) the auto-regressive parameters and \( \epsilon\textsubscript{t} \sim N(0, \sigma^2) \). This modelling approach is a generalization both of Hidden Markov Models and Autoregressive models. In contrast to ordinary Hidden Markov Models, the suggested method takes into account that the current \( CO\textsubscript{2} \) level is not only dependent on the occupancy status but also heavily dependent on its own past values.

3 Results
Table 1 shows Akaike’s information criterion (AIC) and the Bayesian information criterion (BIC) for different numbers of states in the model. One can see that both criteria are lowest for \( m = 5 \) states. The table also displays the number of parameters which is \( m^2 + 6m \).

<table>
<thead>
<tr>
<th>states</th>
<th>parameters</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>16</td>
<td>-19187</td>
<td>-19080</td>
</tr>
<tr>
<td>3</td>
<td>27</td>
<td>-20326</td>
<td>-20146</td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>-20771</td>
<td>-20503</td>
</tr>
<tr>
<td>5</td>
<td>55</td>
<td>-21392</td>
<td>-21023</td>
</tr>
</tbody>
</table>

The analysis of residuals shows that the suggested method inherits the dynamics of the \( CO\textsubscript{2} \) curves much better than an ordinary Hidden Markov Model. Figure 2 shows that the residuals are fairly mutually independent and normally distributed. This indicates that the model describes sufficiently the variation in the data.

4 Discussion
The auto-regressive part of the model accounts well for the high temporal dependency in a \( CO\textsubscript{2} \) curve. Therefore, the suggested model can be considered as a promising candidate for extracting information about the different states of building occupants’ activity.