Integration of the stochastic models of occupancy behavior schedules in BES tool for accurate prediction of energy consumption

Sihem Guernouti#1, Myriam Humbert#2

# Cerema – Center for Studies and Expertise on Risks, Environment, Mobility, Urban and Country planning, Nantes, France

1sihem.guernouti@cerema.fr

2myriam.humbert@cerema.fr

Abstract
A large number of simulation tools are available for the estimation of a building’s energy consumption (such as TRNSYS or EnergyPlus). These tools may produce yearly profiles of the energy needed by a building, for heating, cooling, lighting and ventilation as well as for hot water and electrical appliances. Nevertheless, several studies show that there are significant discrepancies between measured and calculated energy consumptions by these simulation tools. The stochastic variables linked to the climate and to the behaviour of the buildings’ occupants have important influences on those fluctuations. While climate data are available by measures with a correct accuracy, occupancy in all its stochastic activities are still lacking. The aim of this work is to evaluate how the existing occupancy stochastic models can contribute to a better accuracy in the prediction of energy consumption. The algorithms tested and presented concern the occupancy hours and occupancy rate. To achieve these objectives and to calibrate the tested models, the occupancy and energy are monitored on an existing office building. The occupancy monitoring results are presented and compared to the tested models. The impact of occupancy on the predicted energy consumption is then evaluated on the office building and compared with the monitored energy consumption.

Keywords – Occupancy behaviour; Markov chain; modelling; validation

1. Introduction
A large number of simulation tools are available for the estimation of a building’s energy consumption (such as TRNSYS [1] or EnergyPlus [2] These tools may produce yearly profiles of the energy needed by a building, for heating, cooling, lighting and ventilation as well as for hot water and electrical appliances. Nevertheless, several studies show that there are significant
discrepancies between measured and calculated energy consumptions by these simulation tools [3]. It is therefore necessary to reliably estimate the fluctuations in energy consumption. A study of 290 homes with the “same size” in [4] has showed that the heating consumption had a very large range of variation. The highest recorded consumption is twenty times greater than the lowest. This study shows the significant impact which have occupants on buildings energy consumption. The stochastic variables linked to the climate and to the behaviour of occupants have important influences on those discrepancies. A number of works [5], [6] show that occupancy is an important factor in estimating building energy consumption because the behaviours of occupants influence directly the use of systems. Moreover, several works [7], [8] have been done in developing occupancy-based control systems. So, climate data are available by measures with a correct accuracy, corresponding data linked to the occupant in all its stochastic activities are still lacking. The aim of this work is to adapt and to implement a method to realistically model the behaviour of the occupants of an office building. Thus, the aim would be to replace the standard patterns of occupants currently used in thermal simulation tools with those obtained using the proposed method of occupancy assessment. This, can help sizing associated systems (heating, cooling, ventilation, etc.) and to achieve better control strategies.

2. Modelling of occupants presence

Several advanced models have been proposed in the literature to model building occupancy ([9]; [10], [11]). In this work, we used the model of [11]. This model uses Markov chain with four transition functions: $T_{01}$ is the probability that the occupant being present knowing that he is absent. $T_{00}$ is the probability that the occupant stays absent. $T_{10}$ is the probability that the occupant being absent knowing that he is present and $T_{11}$ is the probability that the occupant stays present. For each time step, the Monte Carlo method will be applied to the probability of state change depending on occupant’s presence or not. A random number will be compared to this probability. Therefore, if it exceeds that number, we consider that the occupant will change it state: it will be absent if he was present and vice versa. This model allows treating the event that the occupant takes a long absence period. It means an absence more than 24 hours but that does not correspond to a weekend. This treatment of long absences assumed additional input data. In the presence model of
Page, there were four main input data: 1) the probability profile of weekly presence; 2) the mobility parameter: it means, the ratio between the sum of the probabilities of changing state and the sum of the probabilities of staying in the same condition; 3) the long absence number; 4) the distribution of durations of long absence periods. In this work, we deal with long absences in a different manner from the model of Page. We chose to use the distribution of intervals between long periods of absence, while Page used the occurrence of these periods. Otherwise, we have also changed the algorithm. Initially, at the beginning of each time step, we checked if the occupant decides to begin a long absence period, but by doing that these periods occurred too often. For that reason, we moved this checking in the time of departure of the occupant. Figure 1 presents the process of the developed presence model.

3. Application of occupancy model on monitoring building

The studied building

The Cerema building in Angers (France) was monitoring to qualify its energy performance. Specifically zone 1 was instrumented to test and validate occupancy models. This is an area of 114 m² and ceiling height of 2.6m. It has six offices (Figure 2). The parameters measured in offices are the state of the occupant (present / absent), the temperature (inside and outside), the state of the windows (closed / opened), and the interior lighting level. For each office, we have available one year (January 1st to December 31th 2013) of presence data with a time step of 5 minutes. After aggregated data every half hour (17 520 data values), we split each office data into two equal samples. Thus, the first data set is used for model calibration and the second for the comparison.

Profile of probability of presence

The profile of probability of presence was determined from the first data sets (8760 values) for the six offices. At each time step (1/2 hours) for each day of the week, the ratio of cases where the occupant was present on the considered interval is made. Figure 2 gives an overview of these profiles for a week. These profiles clearly show the diversity of presence behaviours that can be found in a tertiary building. The probabilities obtained for the studied offices are low, which may reflect a lot of travelling (external or internal). This observation confirms the working practices in Cerema where employees are often in travelling. The Z1B2 office has very low
presence rates in comparison with the other offices of zone 1. This is confirmed by the occupancy survey for this office. The occupant said he was often on the move. Conversely, the office Z1B4 has higher presence rates, which is confirmed by the survey, since the occupant of this office has less travel than the Z1B2 occupant. Finally, the office Z1B3 presents a different profile, as the presence rate is relatively the same even during the weekend. We found that the measurement data for this office included many missing data that can explain the resulting profile.

Figure 1: Generic flowchart of the presence model

**Mobility parameter**

The user sets the mobility parameter (μ). However, to have an order of magnitude of its value, we calculate μ using equation (1).

$$\mu = \frac{T_{01}(t) + T_{10}(t)}{T_{00}(t) + T_{11}(t)}$$ (1)

It is therefore necessary to extract from presence data, values of weekly probability of transition for each time step to preview the profile of mobility parameter. To facilitate the implementation of the presence model, we take a constant mobility parameter input data. Thus, we calculate, first, the average of μ for the offices at each time step. Then we calculate the average for the five working days (considering that during the weekend this parameter is null over two days would not reflect sufficiently a weekly profile). We calculated this parameter for each office from the first data sets of measurement of the six offices, (Table 1). Figure 3 shows the mean weekly profile
of the mobility parameter for five offices. The Z1B3 office was excluded from the average since it had an atypical behaviour.

Table 1: Average mobility parameter for each office and for 5 working days

<table>
<thead>
<tr>
<th>Office</th>
<th>Mean mobility parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z1B2</td>
<td>0.005</td>
</tr>
<tr>
<td>Z1B3</td>
<td>0.063</td>
</tr>
<tr>
<td>Z1B4</td>
<td>0.042</td>
</tr>
<tr>
<td>Z1B5</td>
<td>0.025</td>
</tr>
<tr>
<td>Z1B6</td>
<td>0.040</td>
</tr>
<tr>
<td>Z1B7</td>
<td>0.026</td>
</tr>
</tbody>
</table>

Figure 3: Weekly profile of the mean mobility parameter for studied offices and 5 working days (left) and Mean weekly profile of probability of presence for offices Z1B4, Z1B5, Z1B6, Z1B7 (right)

**Long absence periods**

To implement the presence model, it is necessary to analyse data related to long absence periods (i.e. absence periods upper than 24 hours, not corresponding to weekends). This study should enable to get a distribution for time intervals between each period and another for the length of these periods. The first step was an inventory of
these periods, calculation of their duration and time interval of their appearance for each of the studied offices. Then, we have selected a probability law, which is suitable to sampling data. Thus, the exponential law was chosen. This applies well to modelling execution time of a task and the life of a system. The density function is given by equation (2) where \( \lambda \) is the parameter of the distribution. We use the method of maximum-likelihood estimation (MLE) to evaluate \( \lambda \). Table 2 gives the values of the parameter \( \lambda \) for each studied office.

\[
f(x) = \frac{1}{\lambda} \exp\left(-\frac{x}{\lambda}\right) \text{ si } x > 0 \text{ et } 0 \text{ else}
\]

(2)

Table 2: Estimation of parameter \( \lambda \) for each office

<table>
<thead>
<tr>
<th>Office</th>
<th>( \lambda ) for long period absence duration</th>
<th>( \lambda ) for long period absence intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z1B2</td>
<td>121.33</td>
<td>124.19</td>
</tr>
<tr>
<td>Z1B3</td>
<td>213.67</td>
<td>159.50</td>
</tr>
<tr>
<td>Z1B4</td>
<td>85.56</td>
<td>83.46</td>
</tr>
<tr>
<td>Z1B5</td>
<td>94.94</td>
<td>96.35</td>
</tr>
<tr>
<td>Z1B6</td>
<td>89.04</td>
<td>92.00</td>
</tr>
<tr>
<td>Z1B7</td>
<td>72.24</td>
<td>74.04</td>
</tr>
</tbody>
</table>

**Considered data for the application of the model**

First, we tested the model of presence for each studied office using the parameters shown in Table 1, Table 2 and Figure 2 (bottom). We then carried simulations using means settings and excluding Z1B3 office from the calibration sample. We have calculated the averages of the remaining five offices:

- The presence probability profile shown in Figure 2.
- The parameter of the exponential distribution for the duration of periods of long absences (\( \lambda \)) equals to 92.62;
- The parameter of the exponential distribution for the intervals of long periods of absences (\( \lambda \)) equals to 94.0;
- The mobility parameter (\( \mu \)) equals to 0.028.

(Page, 2007) listed the statistics that should be calculated to estimate the success of the presence model. In order to restrict the number of figures in this paper, we have decided to present two criteria to analyze the results provided by the presence model: 1) the distribution of first arrivals; 2) the distribution of last departures.

Actually, the behaviour of occupants is usually very different at their first arrival and last departure from any intermediate arrivals and departures. The first arrival of the occupant generally corresponds to the setting by the occupant of his favourite configuration (state of the
blinds, set-point of the heating system, state of artificial lighting, etc.). These first settings might often stay unchanged until the last departure of occupant. The latter puts back the zone to its unoccupied state (switched off the lights and closed the windows, etc.) in anticipation of his absence until the next day or beyond.

**Results for first arrivals and last departures**

Indeed, it seems interesting to study the distributions of the first arrivals and last departures for daily presence of occupants. To do this, it was necessary to use a large statistical processing of data, both for measurements data and for profiles obtained by the presence model. Figure 5 presents comparisons of first arrivals and last departures profiles for each studied office using own calibration parameters. Moreover, Figure 6: Comparison of first arrivals and last departures achieved by the measurement (green) and the model with mean calibration parameters (red) for offices 2 and 3 ((same results for all offices)

shows comparisons of first arrivals and last departures profiles for each studied office using mean calibration parameters. Overall, we can see that the model used both with calibration parameters of each office and mean parameters over the Zone 1 represents well the first arrivals and last departures. However, the model used mean parameters, does not always represent peaks. We have also compared the cumulative frequency of presence for each office. We find that curves follow but with a lag. In addition, we have observed that the model underestimates the frequency of occurrence of presence. This may be due to a poor estimate of the absences less than 24 hours.

4. **Integration of occupancy model in BES tool**

To assess the contribution of the integration of occupancy model in a dynamic building energy simulation tool (BES), we used TRNSYS 17 to perform a sensitivity analysis on metabolic contributions of the six offices of the Zone 1. Modelling in TRNSYS includes Type 56 for the calculation of building heat exchange, Type 157 for aeralic exchanges and Type 34 for solar shadows. We have coupled to TRNSYS a stochastic model of presence (which generates the presence rate at each time step for the simulation) and a stochastic model of actions on windows that calculates the state of windows each time step. This model is not used for the energy simulations carried later. This integration was carried out through the type 155 used to link MATLAB to TRNSYS. Four occupancy scenarios were
tested: a reference scenario corresponding to the actual occupation of the offices, a scenario resulting from the presence stochastic model for each office (Stoc), an average stochastic scenario for all zone 1 (Stoc\_mean) and finally a conventional scenario (Conv) corresponding to the general practice for dynamic simulation in design offices in France (Figure 4). The latter scenario is used to characterize the occupation of the other offices of the building that are not part of zone 1. The other simulation parameters are identical between the four configurations. The comparison of the number of hours of presence per studied configuration is shown in Figure 4. Furthermore, for all scenarios, the set-point heating temperature is equal to the measured temperatures for offices of the zone1 and to 20°C for the others offices. The cooling set-point temperature is equal to 16°C in all offices.

Figure 4: Conventional daily presence schedule (left) and Number of hours of annual presence according to selected scenarios (right)

Figure 5: Comparison of first arrivals and last departures achieved by the measurement (green) and the model with office calibration parameters (red) for offices 2 and 3 (same results for all offices)

5. Conclusion and outlook
The first objective of this work was to test, to adapt and to validate the presence stochastic model developed by Page (Page, 2007). The second aim was to replace the standard (or conventional) profiles usually used in building energy simulations by those from a method for assessing the occupancy, in an attempt to reduce the discrepancies between calculated consumption and measured and to have more accurate estimation of loads for sizing HVAC systems. The carried tests of presence model have demonstrated the advantages and limitations of the use of the stochastic model. We can say that the developed stochastic presence model seems to reproduce reality and improves the prediction of the occupation over the conventional model. Nevertheless, there is still work to do to better estimate the absences lower than 24 hours. We then integrated stochastic model to predict the presence and opening windows in building energy simulation tool (BES): TRNSYS coupled to CONTAM. The results have shown the interest of integrating a stochastic presence model in a dynamic building energy simulation. It allows a more realistic estimation of internal gains related to the
occupants than standard scenarios that overestimate the presence time. Furthermore, a survey was conducted among the occupants of the six studied offices in the Zone1. It will be interesting to use this survey to get the input data of developed occupancy models. Indeed, until now, occupancy models were calibrated from large measurement data (weekly profile of presence, mobility parameter, long absences, etc.). However, as part of an audit, it is not always possible to measure the occupancy. Therefore, it is important to test the calibration of stochastic models from data survey or questionnaires made during an audit and to estimate the error in relation to models calibrated from the measured data.

Acknowledgment

This work was funded in part by the National Agency (ANR) Research through the Habitat intelligent solar photovoltaic program (MEMORY No. ANR-10-Habisol 006-project).

References