Development of occupancy-based multi-scale building archetypes

Divyanshu Sood [a], *Ibrahim Alhindawi* [b], *Usman Ali* [a], *James McGrath* [b], *Miriam Byrne* [b], *James O'Donnell* [a]

[a] School of Mechanical and Materials Engineering and Energy Institute, University College Dublin, Dublin, Ireland, divyanshu.sood@ucdconnect.ie, usman.ali@ucd.ie, james.odonnell@ucd.ie

[b] School of Physics & Ryan Institute's Centre for Climate and Air Pollution Studies, National University of Ireland Galway, Ireland, i.alhindawi1@nuigalway.ie, james.a.mcgrath@nuigalway.ie, miriam.byrne@nuigalway.ie

Abstract. In the context of the European building stock, more than 50 % of buildings were built before 1960, and it has been estimated that 75 % of the current building stock will still exist in 2050. A typical approach to estimate energy consumption at multiple scales is by using archetypes which are the cohort of representative buildings with similar characteristics. Typically, archetypes are classified based on year of construction, type of dwelling and type of heating system. Since, this classification does not account for the stochastic nature of occupancy, a typical occupant presence pattern from the literature is considered. This study develops a methodology to generate stochastic occupancy profiles using the UK Time Use Survey (TUS) 2014-15 data. The occupancy profiles take into account the affect of the day of the week, the month of the year, the number of residents in the household and the type of dwelling. To test the methodology, we used the Irish residential building archetypes and 5-8% variations in energy use intensity are observed using the developed occupancy profiles for an apartment archetype having one and two occupants. The generated occupancy profiles facilitate the pathway to develop robust archetypes for reliable energy prediction at an urban scale. Furthermore, robust archetypes allow policymakers and urban planners to recommend appropriate energy efficiency measures for the sustainable development of residential building sector.

Keywords. Urban building energy modelling, Stochastic occupant behaviour, Building archetypes
DOI: https://doi.org/10.34641/clima.2022.243

1. Introduction

Urbanization has put more stress on the environment, and therefore, more than ever, sustainability needs to be considered in the construction sector for energy-efficient urban planning. Energy in Buildings is a trending topic as part of the wider climate change agenda. Buildings as a sector are the largest energy consumer with 36 % and 39 % of CO_2 emissions in the European Union and the United States, respectively. The European Union (EU) has set the target to reduce greenhouse gas emissions by 40 % before 2030 when compared to 1990 levels [1].

Cities contain thousands of buildings, and energy estimation of each building at such a larger scale is a tedious task. However, energy estimation at district, city, regional or national scale is required to allow policymakers and building energy experts to formulate energy policies for sustainable urban planning. In addition, the energy estimation helps evaluate the scope for energy retrofitting to make the existing buildings energy efficient and reduce associated carbon emissions. Hence, archetypes are developed to simplify the analysis and represent a more extensive building stock. Reinhart et al. [2] defines archetypes as a representative set of buildings that share similar geometric and non-geometric characteristics. Normally, archetype characterisation uses national average values of occupancy schedules and U-values. However, the national average values are invalid while estimating energy at district or city, or regional scale [3].

Instead of using average values to fill the gap in data, researchers use data-driven approaches to improve the quality of the data to develop an effective methodology for archetype development [4]. Machine learning techniques such as segmentation can be used on large datasets including EPC and TABULA to identify archetypes and corresponding input parameters. Ali et al. [5] developed archetypes at multiple scales to predict the energy demand at various levels by segmenting archetypes based on construction year and type of dwelling. The values during characterisation were the average values according to the scale of analysis. Fewer studies focused on integrating occupancy behaviour into archetypes. Buttitta et al. [6] developed occupancy integrated archetypes, segmented based on various occupancy schedules developed using Time Use Survey (TUS) data and considered occupancy pattern as active or non-active. Although, this does not capture the logic behind occupants' active and non-active status in the buildings. Moreover, using archetypes with fixed occupancy patterns have led to unrealistic peak heat demand at high temporal resolutions [7]. The deviation in energy estimation could be up to 30 % when fixed occupancy profiles are considered with archetypes rather than occupancy integrated archetypes[6].

The energy consumption patterns for dwellings are highly stochastic and vary considerably between different customers. Yao et al. [8] categorised energy consumption determinants in two categories: physical and behavioural, to generate load profiles. Widen et al. [9] used time-use data to interpret occupants' usage patterns along with appliance ratings to produce the energy consumption profiles. The load profiles generated using meter data can be hourly/weekly/monthly or 15 minutes, depending on the data availability. Sensors help capture the load variations in a building and based on these variations, occupancy schedules can be developed by using a Markov Chain model [10], [11]. These deterministic models are easy to implement and can be represented as fractions with values in the range [0,1]. Chong et al. [12] used spatial occupancy data and applied the Bayesian calibration approach to reduce occupancy prediction errors from 37% to 24%. The IEA-EBC Annex 66 project focused on setting up a platform to standardise a simulation methodology to incorporate occupancy and occupant behaviour into building energy modelling. Several publications were produced under Annex 66 [13], [14]. The majority focused on developing occupancy models based on data collected from sensors and energy meters to track the presence and absence of occupants based on load profiles. The aforementioned studies do not explain the behavioural changes among occupants and their variations in daily activities due to the type of dwelling they live in and the number of occupants residing in the dwelling. It is crucial to realistically develop the profiles of the occupants in a household to understand the logic behind variations in energy use.

Previous studies link occupancy to building energy modelling and minimal information is available on developing occupancy models for UBEM and evaluating their effect on energy consumption. The major issue is the use of typical occupant presence profiles from the literature or is based on priorknowledge that leads to large discrepancies in the predicted energy. As occupants significantly influence the energy consumption in buildings, energy usage is highly dependent on the number of occupants, their behaviour and type of dwelling they livein. Due to highly stochastic occupant behaviour, a fixed pattern could lead to substantial differencesin energy estimation of similar dwellings. It is crucial to consider the realistic interaction of occupants with the building while quantifying archetypes to estimate energy consumption at an urban level. Therefore, archetypes classification should also consider the number of occupants so that their behavioural patterns can be linked within archetypes.

The novelty of this paper lies in formulating a methodology to develop stochastic occupancy profiles and integrating them into building archetypes to reliably predict the energy consumption of building stock. This study classifies the building archetypes based on the number of occupants along with considered year of construction and type of dwelling. This classification allows us to investigate the influence of occupants and their activities on energy consumption at different geographical scales such as postcode, county and national.

The paper consists of the following sections: Section 2 defines the methodology used to develop occupancy profiles and occupancy-based archetypes. Section 3 introduces a case study to implement the devised methodology. This section also includes the analysis of results and discussion. Section 4 describes the conclusions from the study.

2. Methodology

This paper assesses the impact of stochastic occupancy behaviour on energy consumption of multi-scale archetypes. To develop the occupancy profiles, UK Time Use Survey (TUS) 2014-15 data are used. The data represents various activities performed by occupants on a daily basis.

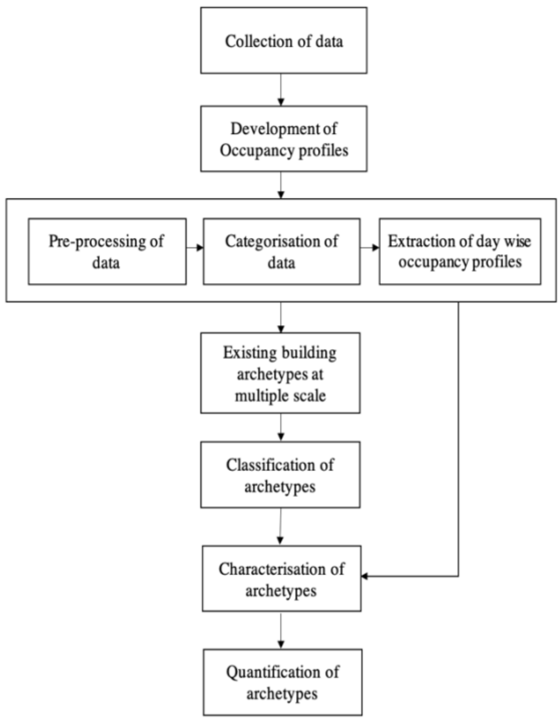

Fig. 1 - Overarching methodology to study the impact of occupancy on multi-scale archetypes

Fig 1 illustrates the overarching methodology used in this study to develop and evaluate occupancy enabled archetypes. This study used the mixed-method approach as the data used to develop occupancy profiles are qualitative data which is further used to quantify energy consumption. The mixed-method approach uses both qualitative and quantitative techniques for data collection and analysis [15]. The TUS data are qualitative and were recorded through interviews and by updating activity diaries.

2.1 Collection of data

The TUS data can be directly linked with the user's behaviour to generate occupancy profiles. The data consists of an activity diary that records occupants' activities such as sleeping, cooking, eating, bathing and so on at 10-minute intervals. For 4460 house-holds data, 11460 interviews were recorded in a diary for two consecutive days by a different house-hold member. The diary also consists of other information related to the type of housing, the number of adults & children in the house, the month of the year and the day of the week on which the diary was filled. The data were recorded in such a way that for day of the week, 1 represents Sunday, 2 represents Monday followed by other days of the week in a similar pattern. The diary assigned a unique serial number to every house that helped identify the total number of respondents from the household.

2.2 Pre-processing of data

Originally, the data were recorded in a numeric format and every number represents the activity performed by the occupants. To start with, the numeric entries were replaced with the corresponding activities as given in the dictionary provided with the data. The occupancy data are usually collected through surveys that are prone to data irregularities such as incomplete data and missing data. It is important to remove these inconsistencies in order to make the data usable. TUS data contains few entries that do not have all the information related to the occupancy activities. Therefore, rows with incomplete or non-applicable entries were removed from the dataset for data consistency throughout the analysis. The data has more than 110 various activities performed by different individuals. To reduce the number of activities and make data useful for further processing, activities that fall under similar category were combined. For instance, admin work, office work and shopping were named as "Outdoor Activities (OA)" to understand the data better and to remove the unnecessary variables. The data were recorded in 2014-15, therefore office work is considered as an "Outdoor Activities (OA)". More details are provided in Tab 1.

Tab. 1: Details on grouping of activities

Activity group name	Activities
Fitness (FT)	Exercise Gym Fitness
Outdoor Activities (OA)	Office work Admin work Travel Shopping Office break
Socialising (Soc)	Cinema Movie Concert

2.3 Categorisation of data

Firstly, the data are grouped based on the type of accommodation, and we assumed that apartments shared similar characteristics as flats or maisonettes. Therefore, the occupancy profiles developed for flat or maisonette is also used for the various apartment types (see Fig. 2). Similarly, occupancy profiles for other dwellings such as semi-detached houses, detached houses, and terraced houses used profiles developed for the houses or bungalows (see Fig. 2).

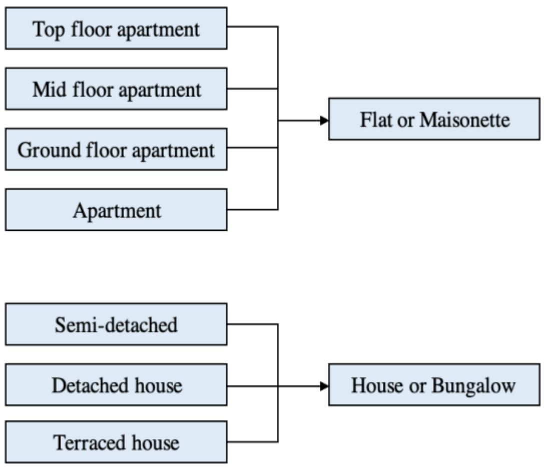

Fig. 2 - Assumptions to use the occupancy profiles for various building archetypes

The occupancy profiles are based on the day of the week, the month of the year, the number of occupants and the type of dwelling, see Algorithm 1. The total number of occupants in a household is determined by adding the total number of adults and children. This enables us to group the data based on the number of occupants living in a household. This allows us to visibly understand the variations in occupancy behaviour. To further capture the effect of monthly and daily variations in occupancy activities, data were sorted into different months followed by days of the week. The above-mentioned steps for categorisation enable us to capture the effect of various parameters on occupancy behaviour.

Algorithm 1 Categorisation of TUS data to develop occupancy profiles

$i = HouseorBungalow, FlatorMaisonette$
for $TypeOfAccomodation = i$ **do**
  for $OccuNum = 1, 2$ **do**
    for $Month = 1, 2, 3 \ldots 12$ **do**
      for $DayOfWeek = 1, 2, 3 \ldots 7$ **do**
        DayOfWeek.mode()
      end for
    end for
  end for
end for

2.4 Extraction of occupancy profiles

Once the data categorisation is completed, the next step is to extract the profiles. For categorical data or qualitative data, a mode is the best way to identify the frequently occurring data point in a given dataset[16]. Statistics.mode() was used to measure the central tendency in each time slot to extract the most frequently performed activities. Mode provides the activity that appears most often in each time slot, leading to a full day activity profile. The data were divided into 10-minute intervals; therefore, we have 144-time slots per day. The developed methodology also supports statistics.multimode(), i.e. bimodal or multimodal, to consider the second, third or further modes, if they exist. However, this study only used the profiles based on the first mode to balance the computational load, time and complexity.

2.5 Modelling and quantification of archetypes

Archetypes are helpful to estimate the energy demand at various geographical scales when insufficient building stock data is available. This study utilised the previously developed archetypes and further classified them based on the number of occupants in a dwelling to incorporate the stochastic nature of occupants. The archetypes are modelled in Design Builder and the 7/12 activity schedule is considered to incorporate the developed day wise occupancy schedule and simulated using EnergyPlus software, which is the most used software for BEPS. EnergyPlus is considered a reliable software for simulating residential buildings by carefully selecting and defining input parameters. The equipment schedule has also been modified as per the occupancy presence or absence and the activity schedule. The input data such as multi-scale U-values required for the energy simulations were taken from Ali et al. [5] except for the occupancy related data. The occupancy profiles developed in this study are linked with the archetypes and using jEPlus Macros, parametric simulations were performed.

3. Results and discussion

There was an 85 % growth in apartments between 2002 and 2016 in Ireland [17]; therefore this study focuses on evaluating the impact of occupancy on the energy consumption of an apartment built after 2006 using multi-scale archetypes. A case study of the Irish apartment archetype is chosen to establish the relationship between occupancy and multi-scale building archetypes. In this study, we used archetypes at three geographical scales namely, postcode (Dublin 1), county (Dublin county), and national (Ireland).

3.1 Overview of the TUS data

The initial impression of the TUS data provides an insight into the diverse range of activities performed by occupants in a household. In TUS data, 16,550 diary days were completed by respondents selected for interviews. Fig. 3 represents the value count of the various activities recorded in the data by the respondents. It is evident from the Fig. 3 that occupants undertake various activities on a daily basis and sleeping has the highest value count in the data followed by TV Watching, Office work and so on.

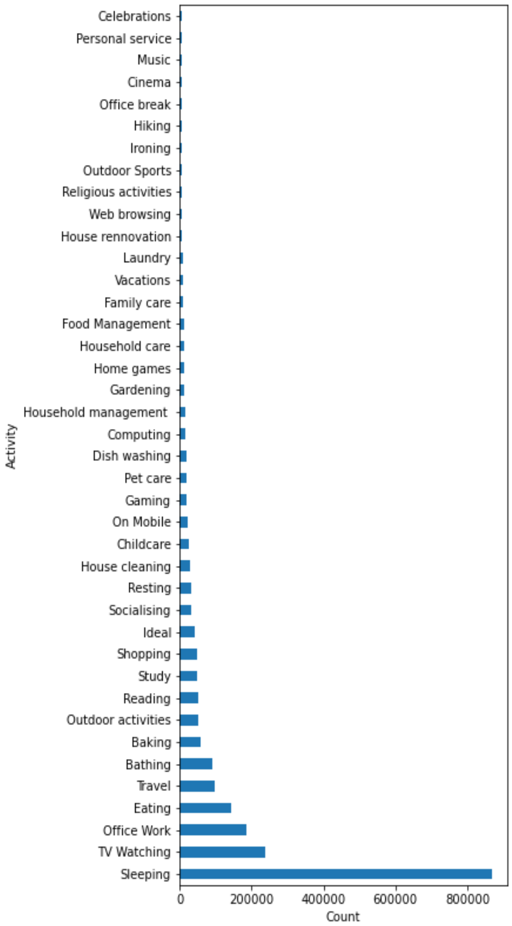

Fig. 3 - Different activities and their value count in the TUS data at 10-minute intervals

There are approximately 40 various activities mentioned in the TUS data recorded by the occupants. The TUS data are an effective way to record occupant activities that enable us to understand on how people spend their time on a daily basis.

3.2 Stochastic occupancy profiles

In this study, we developed stochastic occupancy profiles using activity diary data that consists of actual occupancy patterns updated at 10-minutes intervals. The profiles represent the various activities performed by occupants on a daily basis.

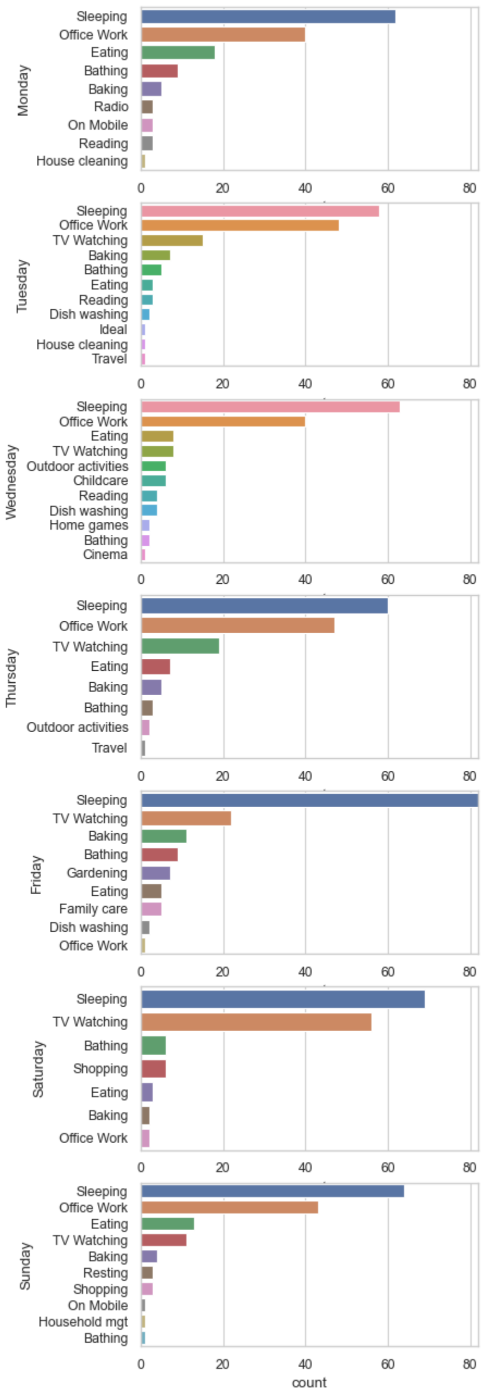

Fig. 4 - Value count of activities performed in a week in the month of January to represent the variations in occupancy behaviour in a one occupant household

A month has seven profiles representing seven days of the week. For instance, for January, the occupancy profile for Monday is representative of all the Mondays in January. In total, 168 profiles were developed for an apartment with one occupant and two occupants. Fig 4 represents the value count of activities performed in a one occupant apartment. From Fig 4, we can see that each day is different in terms of activities performed and duration of the activities. For instance, the sleeping count for Monday is over 60, and for Tuesday, the count falls below 60. Similarly, for Friday, the sleeping count is more than 80, 10-15 counts more than Saturday and Sunday, respectively. Similarly, TV watching time, bathing time, eating and cooking time varies significantly each day. This shows the merits of having realistic profiles for building energy modelling as these enable us to closely understand the occupancy behaviour within the building. These variations are usually ignored and assumed constant to reduce the complexity of analysis.

Fig 5 represents the variation in activities performed by occupants on Monday in a one and two occupant apartment. Apart from the value count, the kind of activities performed by occupants is different in both cases. Based on the value count and kind of activities performed, it can be comprehended that using fixed profiles for energy simulations can lead to significant discrepancies in predicted energy.

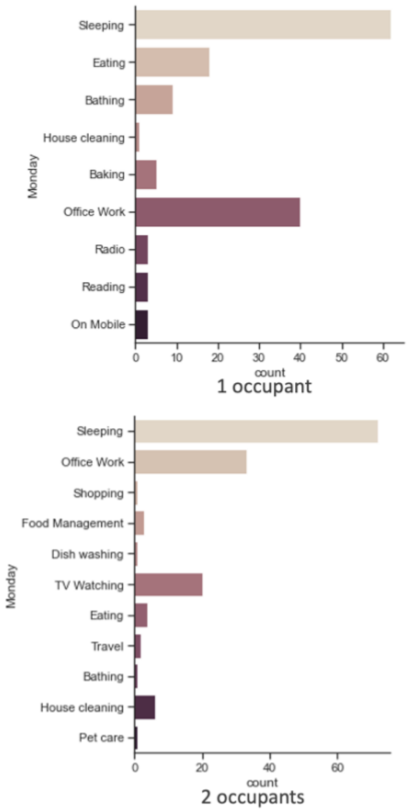

Fig. 5 - Variations in activities on Monday in a one and two occupant apartment

3.3 Multi-scale U-Values

The multi-scale U-values for apartment archetypes are taken from [5]. It is evident from Fig 6 that U-values for different building elements such as wall, roof, window, floor, and door varies at different geographical scales.

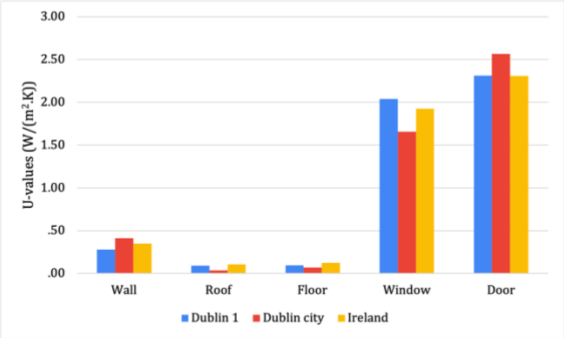

Fig. 6 - U-values for an Apartment built after 2006 over multiple scales (W/(m^2-K))

For instance, U-values for apartment windows are 2.04, 1.66, and 1.93 ($W/(m^2K)$) for Dublin 1, Dublin county and Ireland, respectively. U-values significantly affect the building energy estimation and should be carefully chosen for the multi-scale analysis. For multi scale archetypes, assuming national average U-values for other geographical scales such as postcode and county remain invalid as these values lead to substantial discrepancies in energy output.

3.4 Building energy performance simulation

This section represents the effect of the number of occupants and their behaviour on multi-scale apartment archetypes. The multi-scale apartment archetypes with typical occupancy profiles and state-of-the-art developed stochastic occupancy profiles are compared in terms of EUI. We used the EUI results from a study carried out by Ali et al. [5] as a base case that used typical occupancy profiles provided by the Design Builder software irrespective of the number of occupants and type of dwelling.

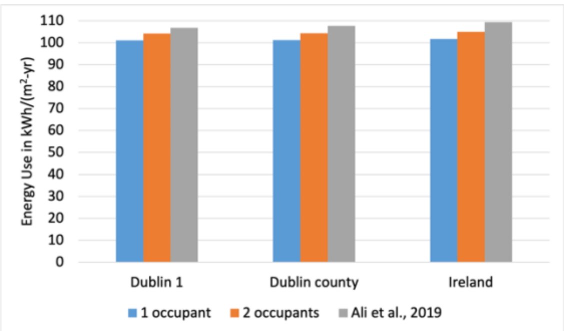

Fig. 7- Comparison of apartment archetype energy use with 1 occupant profile and the base case (kWh/(m^2-yr)

Fig 7 compared the base case with one occupancy and two occupancy household. It is evident from the Fig 7 that energy consumption using conventional profiles is different from the results obtained using profiles based on one and two occupants. There is approximately 5-8 % variation in EUI when the base case is compared with different occupancy households using archetypes at multiple scales. The absolute differences in EUI were found to be 3.14 $kWh/(m^2yr)$ when one occupancy household is compared with two occupancy households at the national level. As can be observed from the Fig 7, fixed occupancy profiles used in the base case overestimate the EUI which increases the pre-bound effect, i.e. actual energy consumption is lower than the calculated energy. For instance, occupants' activity schedules can be completely different when a comparison is made based on the number of occupants. This concludes the importance of using dynamic occupancy profiles to reliably estimate the energy consumption of archetypes.

4. Conclusions and Future work

Using typical compact occupancy profiles from the literature leads to over or under estimation of energy demand. The variations in occupancy behaviour are inherently uncertain, and assuming typical profiles does nothing but increase the uncertainty. For UBEM, archetypes are used to estimate energy consumption at multiple scales. Conventionally, archetypes are classified in such a way that they do not incorporate occupancy. This study focuses on understanding how the number of occupants and their stochastic nature impacts energy assessment using archetypes. The methodology evaluated the effect of occupancy on multi-scale archetypes. The developed occupancy profiles are based on the type of dwelling, the number of occupants, months of the year and days of the week. Based on Fig 4, it is evident that no day is same in terms of activities performed by occupants. Therefore, using fixed occupancy profiles lead to large discrepancies in predicted energy consumption. The results showed that there is approximately 5-8 % variations in EUI when base case is compared with one and two persons household for an apartment archetype. The variations are due to the number of occupants and their distinct activities. The relative difference of 5-8% is observed at an archetype level when fixed & compact occupancy schedule is substituted with stochastic & yearly schedule. This variation does not seem significant at an archetype level. However, when archetypes' results are extrapolated to various geographical scales, this 5-8% variation can be significant compared to measured energy values at a particular scale. A conventional means of archetypes' classification does not consider the number of occupants and their associated stochastic nature, which is essential as archetypes' result is extrapolated to various geographical scales by using the total number of existing buildings corresponding to that particular archetype. This extrapolation is usually based on determining a factor of multiplication which is typically the number of different types of dwellings. This energy difference can result in inappropriate identification of the inefficient energy areas and lead to unsatisfactory policy recommendations.

The developed occupancy profiles enable us to understand the stochastic nature of occupants and push the researchers to focus on collecting more occupancy data for better energy modelling results. The methodology will help in developing various occupancy profiles and integrate them within archetypes to improve the building energy modelling results. Occupancy integrated archetypes allow us to reduce the gap between estimated and measured values at the building stock level. Measured values refer to the cumulative energy consumption of building stock at various geographical scales such as postcode, county and national. These archetypes improve the reliability of the results at the building stock level and help policy makers, local authoritiesand urban planners to modify, update and implement energy recommendations for the better and sustainable growth of residential building sector.

The present work is limited to an apartment archetype having up to 2 occupants. This study only used the first mode to develop occupancy profiles, however, the research could be extended to take advantage of the second or third mode and develop more detailed occupancy profiles that further enhance the results by providing a range of values. Furthermore, the work could be extended for different types of archetypes for a comprehensive evaluation of the effect of occupancy on multi-scale residential building archetypes. Moreover, IEQ is greatly influenced by the number of occupants and significantly affect energy consumption. Apart from the energy analysis, these realistic occupancy schedules allow us to determine the overheating scenario and various indoor air quality parameters to appropriately evaluate the performance of energy efficiency measures.

5. Acknowledgement

This publication has emanated from research supported in part by a research grant from Sustainable Energy Authority of Ireland (SEAI) Grant number 19/RDD/537. The opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Sustainable Energy Authority of Ireland.

6. Data access statement

The datasets generated during and/or analysed during the current study are not publicly available due to ethical, legal or commercial restrictions.

7. References

[1] A. Malhotra *et al.*, “Information modelling for urban building energy simulation—A taxonomic review,” *Build. Environ.*, p. 108552, Nov. 2021, doi: 10.1016/J.BUILDENV.2021.108552.

[2] C. F. Reinhart and C. C. Davila, “Urban building energy modeling--A review of a nascent field,” *Build. Environ.*, vol. 97, pp. 196–202, 2016.

[3] C. Cerezo, J. Sokol, S. AlKhaled, C. Reinhart, A. Al-Mumin, and A. Hajiah, “Comparison of four building archetype characterization methods in urban building energy modeling (UBEM): A residential case study in Kuwait City,” *Energy Build.*, vol. 154, pp. 321–334, 2017.

[4] G. Tardioli, A. Narayan, R. Kerrigan, M. Oates, J. O’Donnell, and D. P. Finn, “A methodology for calibration of building energy models at district scale using clustering and surrogate techniques,” *Energy Build.*, vol. 226, p. 110309, 2020.

[5] U. Ali, M. Haris, C. Hoare, E. Mangina, and J. O. Donnell, “Energy & Buildings A data-driven approach for multi-scale building archetypes development,” vol. 202, 2019, doi: 10.1016/j.enbuild.2019.109364.

[6] G. Buttitta, W. J. N. Turner, O. Neu, and D. P. Finn, “Development of occupancy-integrated archetypes: Use of data mining clustering techniques to embed occupant behaviour profiles in archetypes,” *Energy Build.*, vol. 198, pp. 84–99, Sep. 2019, doi: 10.1016/j.enbuild.2019.05.056.

[7] L. Diao, Y. Sun, Z. Chen, and J. Chen, “Modeling energy consumption in residential buildings: A bottom-up analysis based on occupant behavior pattern clustering and stochastic simulation,” *Energy Build.*, vol. 147, pp. 47–66, 2017.

[8] R. Yao and K. Steemers, “A method of formulating energy load profile for domestic buildings in the UK,” *Energy Build.*, vol. 37, no. 6, pp. 663–671, 2005.

[9] J. Widén and E. Wäckelgård, "A high-resolution stochastic model of domestic activity patterns and electricity demand," *Appl. Energy*, vol. 87, no. 6, pp. 1880–1892, 2010.

[10] K. U. Ahn, D. W. Kim, C. S. Park, and P. de Wilde, "Predictability of occupant presence and performance gap in building energy simulation," *Appl. Energy*, vol. 208, pp. 1639–1652, Dec. 2017, doi: 10.1016/J.APENERGY.2017.04.083.

[11] J. Parker, A. Hardy, D. Glew, and C. Gorse, "A methodology for creating building energy model occupancy schedules using personal location metadata," *Energy Build.*, vol. 150, pp. 211–223, Sep. 2017, doi: 10.1016/J.ENBUILD.2017.06.014.

[12] A. Chong, G. Augenbroe, and D. Yan, "Occupancy data at different spatial resolutions: Building energy performance and model calibration," *Appl. Energy*, vol. 286, p. 116492, Mar. 2021, doi: 10.1016/J.APENERGY.2021.116492.

[13] S. Pan *et al.*, "Cluster analysis for occupant-behavior based electricity load patterns in buildings: A case study in Shanghai residences," *Build. Simul. 2017 106*, vol. 10, no. 6, pp. 889–898, May 2017, doi: 10.1007/S12273-017-0377-9.

[14] S. Yilmaz, S. K. Firth, and D. Allinson, "Occupant behaviour modelling in domestic buildings: the case of household electrical appliances," *https://doi.org/10.1080/19401493.2017.1287775*, vol. 10, no. 5–6, pp. 582–600, Nov. 2017, doi: 10.1080/19401493.2017.1287775.

[15] J. W. Creswell, "Mixed-method research: Introduction and application," in *Handbook of educational policy*, Elsevier, 1999, pp. 455–472.

[16] M. Allen, "The SAGE Encyclopedia of Communication Research Methods," *SAGE Encycl. Commun. Res. Methods*, Jul. 2017, doi: 10.4135/9781483381411.

[17] CSO, "'Occupied Dwellings - CSO - Central Statistics Office.'" 2016, [Online]. Available: https://www.cso.ie/en/releasesandpublications/ep/p-cp1hii/cp1hii/od/.