A methodology for creating building energy model occupancy schedules using personal location metadata

James Parker[*], *Adam Hardy, David Glew, Christopher Gorse*

Centre for the Built Environment, Leeds Sustainability Institute, Leeds Beckett University,

Broadcasting Place, Woodhouse Lane, Leeds, LS2 9EN, United Kingdom

[*]Corresponding author: j.m.parker@leedsbeckett.ac.uk

Abstract

Occupants affect energy consumption in buildings by contributing internal heat gains, increasing internal carbon dioxide levels and adapting their behaviour. Estimated occupancy schedules are used in building energy models for regulatory compliance purposes and when empirical data are not available. Metadata, such as personal location data, is now collected and visualised on a global scale and can be used to create more realistic occupancy schedules for non-domestic facilities, such as large retail outlets. This paper describes a protocol for extracting and using freely available metadata to create occupancy schedules that are used as inputs for dynamic simulation models. A sample set of twenty supermarket building models are used to demonstrate the impact metadata schedules have when compared with models using the estimated schedules from regulatory compliance. Metadata can be used to create bespoke occupancy profiles for specific buildings, groups of buildings and building archetypes. This method could also help reduce the gap between predicted and actual performance. In the example models, those using the regulatory compliance schedules underestimated heating demand by approximately 10% and overestimated cooling demand by over 50% when compared to models using the metadata schedules. Although this work focuses on UK facilities, this methodology has scope for global application.

Keywords

Occupancy schedules, dynamic simulation modelling, metadata, non-domestic buildings, commercial buildings

Highlights

- A methodology for collecting building occupancy metadata is defined.

- Metadata has been collated to create occupancy schedules for supermarket buildings.

- Metadata occupancy schedules have been applied to building energy models.

- The metadata schedules have a significant impact on predicted energy and thermal performance.

1. Introduction

Amid the global need to reduce greenhouse gas emissions, the built environment offers substantial scope for increasing energy efficiency and reducing the amount of anthropogenic carbon dioxide (CO_2) released into the atmosphere [1]. Non-domestic and commercial buildings account for a significant proportion of CO_2 emissions and represent approximately 18% of the UK's total carbon footprint [2]. Building occupants form part of a complex system within non-domestic facilities and have a direct impact on energy performance. They therefore represent an intrinsic data input within building energy models [3].

As part of the ambition to reduce energy consumption within the built environment, building energy modelling is now an established and cost-effective tool used in the design process for both new-build and major retrofit projects [4, 5]. The design stage offers the first opportunity to reduce building energy consumption and building energy modelling plays a pivotal role in this process [6]. The importance of occupancy patterns and occupant behaviour has been emphasised in recent years through

the IEA Annex 66: Definition and Simulation of Occupant Behaviour in Buildings [7]. Occupancy patterns and occupant behaviour become even more critical to energy and thermal performance when modelling buildings that include occupancy-based controls [8].

Occupancy schedules are used in building energy models to dynamically control occupant density and internal heat gains within the thermal envelope. Ideally, designers will have access to empirical data when creating these occupancy patterns. When producing models for non-domestic regulatory compliance in the UK, estimated occupancy schedules are used to control the temporal distribution of the internal heat gains from people within these models; the National Calculation Method (NCM) provides estimated occupancy schedules for a range of non-domestic buildings that can be used in dynamic simulation modelling (DSM) [9]. The NCM and associated DSM calculations are used, along with other forms of analysis, to underpin the development of Government policy in relation to energy consumption and carbon emissions from the non-domestic building sector [10]. The NCM schedules (and similar default schedules in other countries) can also be used when no empirical occupancy data is available to building energy modellers. As well as controlling the passive internal heat gains associated with building occupants, they can also be used to control other variables such as ventilation rates or lighting operations.

This paper describes a methodology for the use of metadata to create more realistic occupancy schedules that can be used in DSM models of commercial buildings and, potentially, other non-domestic facilities. The term 'metadata' is used to describe a diverse range of data in the modern world but can be succinctly described as "data about data" [11]. It has also been described as data-driven 'knowledge' and it is this definition that is most practical in the context of this work [12]. Metadata are now collected on a global scale and personal location data is used to illustrate the temporal popularity of various locations, including buildings that are open to the general public. It is estimated that over 80% of people within the UK now use smartphones on a daily basis and it is this type of device that is the main source of the location data used here [13].

This work uses data collated and visualised by Google as part of their business analytics services. A method for collecting the published data and converting them into DSM occupancy schedules is described. A sample set of 80 UK supermarkets are used as a case study to demonstrate how this method can be applied. The sample set is used to create schedules that describe average occupancy within this building type and used to create personal location metadata schedules (referred to in the remainder of this paper as 'metadata schedules' for convenience). The metadata schedules are then used within a smaller sample of 20 supermarket DSM calculations to quantify the impact that they have on estimated energy performance.

It is common for there to be significant gaps between the predicted energy consumption of a building and the metered energy consumption of the operational facility [6, 14]. There are multiple reasons for this gap, which has been recorded as being over twice that predicted for some non-domestic buildings in the UK [14, 15]. It is, however, possible to create accurate building energy models by improving the quality and accuracy of input data [5, 16-21]. In many cases, the gap between modelled and measured energy performance is a result of comparing metered data with outputs from models created to achieve regulatory compliance [14, 15]. The work presented in this paper aids the creation of more realistic occupancy schedules for this type of compliance modelling which can, in turn, play a role in improving the accuracy and pragmatism of the model energy performance estimates. The methodology and resultant occupancy schedules are also of practical use to the wider community of building designers, engineers and researchers in the field of building energy performance for both new build and retrofit projects. As energy efficient building technologies mature, their theoretical limits are becoming narrower and, especially in the future, this will place greater emphasis on producing robust design. It will also place more emphasis on energy conservation measures [22]. Producing more accurate occupancy schedules and models of user behaviour can help to maximise the savings from energy efficient technologies and to quantify the potential impact of behavioural change.

In the following section, literature concerned with the simulation of occupancy in building energy models is reviewed. Although there is a broad range of work published on this topic, the review demonstrates the original nature of the approach defined in this work. The methodology section describes how the published data can be collected and formed into occupancy schedules that are suitable for use within DSM software. Results illustrate the occupancy schedules informed by the metadata and compare them with those included in existing regulatory compliance guidance. The results section also compares predicted energy and thermal performance from DSM models using the metadata schedules with models using the default occupancy estimates.

2. Occupancy schedule data for simulation

As mentioned in the introduction, the role of building occupants in building energy simulations is the subject of an IEA Annex [7]. As noted by Yan et al. [4], the importance of occupants was previously highlighted by Annex 53: Total Energy Use in Buildings, which identified occupant behaviour as one of the key influences on building energy performance along with climate, building envelope, building services, indoor design criteria, and operation and maintenance [4]. The aim of Annex 66 is to define occupant behaviour in buildings and to account for this behaviour in building energy simulation models [7]. Of particular relevance to this work, Yan et al, have published a review of the current state-of-the art in this field, as well as identifying areas for further work [4]. Occupants can interact with a number of building elements and artefacts, including opening windows, lighting systems, shading devices, HVAC controls and use of appliances [4]. These interactions can also be defined as resulting in either passive or active effects [23]. Occupants often have very little opportunity to interact with control systems in commercial and non-domestic buildings. Their presence within buildings is, however, a driver for energy performance due to occupants passively adding sensible and latent internal heat gains and influencing the need to condition the spaces that they occupy [3]. Occupants

also increase internal CO_2 levels during occupied periods. Occupant impact on building performance can usefully be divided in to two main influences as defined by Wang et al. [8]. The number of people present in a building or zone at any given time can be defined as *occupancy* [8]; this definition is also consistent with that defined by Liao et al. [24]. The interaction of occupants with building systems can be defined as *occupant interaction*.

For occupancy modelling to be useful to designers and engineers, it must "…balance practicality with accuracy" [4]. Yan et al. cite work published by Melfi et al, which identifies the different resolutions that relate to building occupancy [25]. These are divided into temporal resolution, spatial resolution and occupancy resolution. Following the Melfi et al. model, the occupancy patterns created in this work are at an hourly temporal resolution, at a building-level spatial resolution (in this case, supermarkets) and at an occupied state/number of people occupancy resolution. In the context of the work presented in this paper, it is also important to note that Yan et al. describe models with 'excessive inputs' and that are 'excessively complex' as being 'over-fitted' [4]. This is relevant to the occupancy schedules described here as the data are used to create average weekly profiles, rather than complex profiles that differ from day to day and from week to week in an attempt to capture constantly dynamic and stochastic behaviours. The schedules created in this work are intended for use in models where there is no empirical data available, in early stage design calculations and regulatory compliance calculations. It is therefore practical to use simple weekly average profiles when creating models for representative facilities within a particular classification of building type.

There can be a significant difference between estimated occupancy patterns for building archetypes and those found in operational facilities; this variance is illustrated in section 4 of this paper. In a study conducted by Sun & Hong, it is noted that observed occupied hours in a case study office building are significantly lower (under 50%) than the normalised occupancy schedules published by the Department of Energy (DOE) in the USA [22]. The same study also found there to be significant differences in both the frequency and density of occupancy in various zone types when compared with

the DOE schedules [22]. This will have an impact on the thermal conditions in each space and the energy performance of the building as a whole.

Most commonly, the presence of occupants in a space is represented in building energy simulations using deterministic schedules which define the temporal variation of occupancy [8]. These are usually based upon hourly time-steps for each day of the week, with average weekly profiles used for annual simulations. It is this type of occupancy schedule that is used in the NCM [9]. Currently in the UK, the NCM underpins the Standard Building Energy Model (SBEM) which forms a central part of the Government's strategy concerning energy consumption and resultant carbon emissions from non-domestic buildings [10]. Various DSM software is approved for the creation of SBEM calculations using the estimated occupancy schedules, in conjunction with various other data inputs. These calculations are used to demonstrate compliance with Part L of the UK Building Regulations and in the creation of Energy Performance Certificates for non-domestic buildings [10]. The SBEM, along with other models related to domestic energy consumption, are used to develop and administer Government policies concerned with energy consumption in the built environment and, as such, represent an important part of the Government's overall strategy to combat global climate change. Therefore, there is, theoretically, a politically relevant incentive to ensure that these modelling methodologies are robust and provide a realistic estimate of energy performance.

2.1. Collection of occupancy schedule data

Yan et al. define three different approaches to monitor occupant schedules and behaviours; these include: observational studies (covering occupancy/equipment use monitoring and adaptive occupant behaviour monitoring); occupant surveys and interviews; and laboratory studies [4]. Although questionnaires and interviews can be an effective means of gathering a wider range of behaviours, they

are relatively time consuming, especially when collecting data related solely with occupancy numbers and schedules [26].

A similar set of data collection methods are noted by Yang et al. in a review of occupancy sensing systems [26]. The method described in this work can be defined as an 'occupancy and equipment monitoring' approach as it uses metadata collected through the use of mobile devices to create occupancy schedules. Interaction with mobile devices is cited as a technology that will help to expand understanding and data collection opportunities in this field of research [4, 26, 27].

There are examples within the literature of methods that utilise building performance data to infer occupancy schedules and investigate the relationship between occupancy and energy consumption. Monitored data relating to plug-loads in particular have been used to simulate the impact of occupancy and occupant behaviour on building energy performance [28-31]. A range of researchers have identified sources of readily available monitored data that can be used to infer occupancy schedules. It has been demonstrated that electricity consumption recorded by main meters can be used to calculate occupancy rates within a building and improve the accuracy of building energy models [32, 33]. Other work has extended analysis to multiple data sources and introduced machine learning and artificial neural network (ANN) analysis to calculate and predict occupancy patterns. Wang et al. used a combination of relative humidity sensors, PIR, cameras, RFID and CO_2 concentration sensors as inputs for ANN analysis [34]. Monitored noise levels have also been used as an indication of occupied periods and used to train ANN models [35]. Other work in this field has used inter-related variables such as internal and external temperatures, HVAC demand and lighting consumption as source training data for ANN models [36-38].

There are also examples of dedicated sensor systems designed to collect occupancy data in buildings. Yang et al. divide these into two groups: the first using radio frequency signal technologies; and the second using infrared, ultrasound or video technologies to capture state changes in specific spaces [26]. Occupancy sensors using bi-directional infra-red beams can be used to count the number

of people entering and exiting specific buildings, zones or rooms [39]. There are numerous examples of Passive Infrared (PIR) sensors being used to measure occupancy in specific spaces and this is a relatively mature technology [31, 40, 41]. Wang et al have used data collected through internal CO_2 sensor networks to monitor occupancy. Internal CO_2 concentration has been used as a proxy for occupancy levels due to the natural increase in levels when the number of people in a space rises [42]. There is however a proposed need for advanced modelling of occupancy, as identified by Liao et al. [24]. As well as there being a need to capture the stochastic and dynamic nature of occupancy, Liao et al. also reference work that has found significant measurement errors from occupancy sensing technologies [43, 44].

Researchers have used wireless systems as data collection networks for occupancy levels within specific buildings [25, 45, 46]. A paper published by Martani et al. describes a method for using the number of wireless connections within a University building as a proxy for occupant numbers [45]. Due to the distribution of wireless hubs, occupant numbers can be measured at a room level resolution. This provides more detail than the method proposed in this paper but requires access to secure building-specific systems for data to be collected. It does however use a similar source of data, namely personal devices such as smart phones, tablets and laptops [45]. Melfi et al. describe a similar approach to collecting occupancy data using existing IT infrastructure. As well as using mobile devices, they also use personal desktop computer use and explore the potential to add discrete sensors to the existing infrastructure [25]. Sensors connected to existing Ethernet networks were used by Pitt et al. to collect occupancy and behaviour data and to feed this into building control systems to help improve energy efficiency [46].

All of the existing approaches to collecting and collating occupancy data that are described in the literature can provide the basis for occupancy schedules for building simulation programmes. They are however principally limited to application for specific buildings and do not use publically available data to create the occupancy schedules. In turn, this means that large sets of data are not collated to

inform the creation of more realistic average occupancy schedules for specific building types and zones. Where there are no specific occupancy data for a building it may not therefore be appropriate to use a profile for a similar building. In these instances, these publically available data sets could be used as an alternative by building modellers to predict the performance of facilities for which specific occupancy patterns are unknown.

The papers published by Yan et al and Yang et al provide a comprehensive review of the extensive work in the field of occupant behaviour monitoring [4, 26]. However, there are no specific examples cited in their reviews that use location data from mobile devices (e.g. smart phones) to define occupancy patterns in any type of domestic or non-domestic buildings. The main advantage of location metadata from mobile devices is its use of existing, wide-spread, infrastructure to monitor occupancy schedules, and this type of metadata therefore forms the focus of this work.

2.2. Modelling occupancy

There are a range of methods through which occupancy can be included within building energy models. These have been summarised in a number of recent publications [3, 4, 8, 26, 27, 47]. Much of the published work in this field aims to simulate the stochastic behaviour of occupants within buildings where there are multiple opportunities for adaptive behaviour through interaction with specific control mechanisms. As mentioned previously, this is not the aim of the research presented here, as its focus is upon commercial buildings where a large majority of occupants have very limited opportunity for adaptive behaviour. Models used to capture the complex stochastic behaviour of occupants include Bernoulli processes, Markov chains and Survival analysis [4]. The coupling of two or more simulation tools, referred to as co-simulation, has been proposed as a robust means of integrating this level of complexity into DSM tools [47]. This level of sophistication is not necessary in the methods proposed in this paper. There are however examples of modelling techniques that can accurately account for the

stochastic and dynamic nature of real occupancy schedules [8, 24, 48]. In the work presented by Wang et al., there is some regularity to occupancy schedules in main office spaces but it undoubtedly remains stochastic in nature. There is little consistency in occupancy within transitional spaces such as bathrooms and circulation areas which emphasises the value of the approach proposed by Wang et al. [8]. There may be some benefit in linking the metadata schedules defined in this work with this type of stochastic model as is discussed further in section 5 of this paper.

Of direct relevance to this work is a review published by Feng et al. that defines four types of occupancy models: (1) building level (number of occupants in a building at a given time); (2) space level and occupied status (presence of occupants at a particular time); (3) space level and number of occupants (number of occupants in a space at a given time); and (4) occupant level (detailed tracking of specific individuals) [3]. The work presented in this paper uses deterministic profiles that control the density of occupants within a space using an hourly factor, relating most closely to the second type of model noted above. Sun & Hong note four different approaches to simulating occupancy in building energy model software, namely: direct input or control; built-in occupancy behaviour models; user function or custom code; and co-simulation [22]. It is the first of these that is relevant to the schedules identified in this work as it is the occupancy patterns that are being refined. Direct inputs rely on the on the semantics of the particular software [22]. They are therefore more accessible to most building energy modellers, especially those outside of academic research, many of whom use proprietary software. The three remaining approaches are more effective when accounting for stochastic behaviours, particularly in buildings where occupants have greater control of systems, for example, opening windows or turning off lights.

In deterministic schedules, like those commonly used in DSM software, a factor between 0-1 specifies occupancy rates of between 0-100% throughout the day. Feng et al. reference work published by Duarte et al. which demonstrates large temporal variations between occupancy in office buildings [49]. Recent work published by Liang et al. focuses on the impact of occupancy on baseline energy

models in commercial buildings [50]. A large office building is used in their study and although detailed occupancy is found to be stochastic in nature, the number of occupants within each weekday hourly time slot had relatively low uncertainty. This is relevant to the temporal resolution (hourly) of the schedules described in this paper. Building energy models are considered to be 'dynamic' when they run an annual simulation with at least 8,760 time-steps, equal to the total number of hours in a standard calendar year. The same study also found a high correlation between occupancy and total building energy consumption [50].

It is important to note that much of the work described in this literature review discusses the use of relatively detailed empirical data to refine the practice of simulating occupancy in building energy models. The schedules presented in this paper are intended for use in the modelling of buildings where this type of detailed, empirical data is unavailable. The method described in the next section can be used to create occupancy profiles for specific buildings at an hourly resolution. It can also be used to create average profiles for a particular building type, intended for wider use.

3. Methodology

The methodology for creating occupancy schedules required the metadata to be collected for one facility at a time, in an iterative process; it is envisaged that this procedure can become more efficient with further development. Although the method is relatively time consuming in its current format, it does represent a comparatively efficient means of amassing a broad range of data for a large number of facilities when compared with the existing methods described in section 2.1.

3.1. Metadata collection

The first stage in obtaining the metadata is to enter the name and location of a specific facility as a search term into the Google search engine using web browser software. The metadata indicating popular times is described as "…aggregated and anonymized data from users who have chosen to opt in to Google Location History" [51]. It is therefore indicative of the temporal proportional intensity and frequency of visits to a specific location.

These type of data are only displayed if a facility has a 'sufficient' amount of visit data logged. It is however unknown what number of data is deemed sufficient in this instance. The search return for a sufficiently busy facility will display a chart similar to that shown in Fig. 1; this is an example for one of the supermarket buildings used in the case study that forms the basis of this work. The population data (y axis) are presented at an hourly resolution on the x axis. The value for each hourly time-step of each day is displayed as a proportion of the peak hour of the full week. It is important to note that the peak hour does not provide an absolute value for the number of people in a facility, it indicates the hour with the proportionately highest rate of occupancy for the average weekly profile. In a building simulation occupancy schedule, this would represent a factor of 1.0 (100%) occupancy for the relevant most popular hour of the week. Data presented are an average for each day of the week for the particular facility over the preceding several weeks; the scope of 'several' is not specified and this represents an area for further work, which is addressed in the discussion section of this paper.
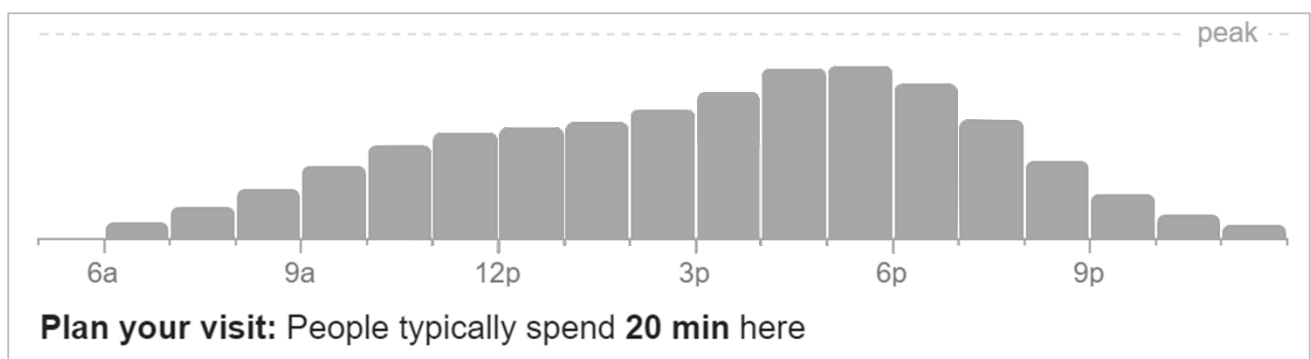
Fig. 1. Example Google chart of popular times for a specific supermarket.

The source code for the chart can be viewed by opening the context menu for the webpage content and selecting the 'Inspect' option. When using the Windows operating system, this is achieved by placing the mouse cursor over a particular bar in the chart and using the right hand mouse button (commonly referred to as 'right-click') which opens the context menu. This action opens a side window which can be used to inspect the webpage code elements. A value in pixels (px) is shown for each of the bars. As noted above, the value for each hour is shown as a proportion of the peak hourly value for that week. In all cases, the peak value is shown as 75 pixels. A value of 75 pixels was therefore used as the maximum value (i.e. 100% occupancy) when calculating proportionate values for all other hours for each day of the week. An hourly resolution for the occupancy data is consistent with the resolution used in existing regulatory compliance guidance. It is common practice to use occupancy schedules with an hourly time-step in DSM. A simple calculation was completed to define a percentage occupancy value for each hour of each day, using 75 pixels as the maximum value to calculate the relevant percentage for any given hour.

At the time this technique was first used, the protocol for extracting data could be followed when using either a standard Google web-search and when searching in Google Maps. However, during the time that this paper developed, when using the described approach to extract data in Google Maps, a different set of source code was revealed. The alternative code discovered in the updated version of Google Maps, defined the height of bars in the displayed charts by the number of pixels measured from the top of the chart; an example of this would be the bar for the peak occupancy hour being zero pixels from the top of the chart. The source code in the updated version of Google Maps also included a percentage value for each bar at each hourly time step. Although this percentage value provides a more immediate answer, the process of extracting the values for each time step was much slower due to the format of the source code.

As mentioned previously, the work presented in this paper, at its current stage of development, is focused on conceptual development and all data was extracted manually for each of the example stores. As mentioned at the beginning of this section, this is however resource intensive and it is envisaged that a more automated method of harvesting this data will be developed in further work.

3.2. Average occupancy schedules

So far, the method has described how a weekly occupancy schedule can be created for a specific facility. In isolation, results from this exercise can be useful for designers and engineers involved in the evaluation of retrofit options for a single facility. In the absence of empirical data, this single occupancy pattern can be used to help improve the accuracy of their baseline DSM model by helping to calibrate outputs against measured data. It is also useful to define indicative, average occupancy schedules for specific building types so that they can be used in early stage design analysis, across larger groups of buildings and for regulatory compliance purposes. A simple method was employed to achieve this. In the work reported here, supermarket buildings were used as a case study to demonstrate the application of this methodology.

A group of 80 supermarket buildings were selected at random from the 6 leading supermarket brands; 10 from each brand with daily opening hours and an additional 10 stores from the 2 brands that operate 24 hour stores. Metadata were collected for each store using the method described above and average occupancy schedules were created following interrogation of the data. It is again important to reiterate that the values for each store do not represent an absolute number of occupants, rather a proportionate value for each hour relative to the most popular hour of the week. The average schedules for building types are therefore based upon proportionate hourly values for each store and not the absolute number of occupants at any stage. In the building energy models, the absolute number of occupants is controlled using a m^2/person default value as described in the next section. The results of

this exercise and example occupancy schedules are also presented in the section 4. Annual schedules were created using the compiled weekly profiles, using Sunday opening for all bank holidays in the calendar year. The empirical schedules created using the metadata have been compared to those defined in the regulatory guidance for supermarket buildings (large food retail stores) [9].

3.3. Building energy models

A group of DSM models were used to demonstrate the impact that the metadata occupancy schedules have on predicted energy performance. This is quantified by comparing outputs from the models using the metadata schedules with those from models using the default NCM occupancy schedules. The models used in this work are simplified DSM models of supermarket buildings and include a rationalised group of zone types which have been described in previous work [52]. The group of models were calibrated against sub-metered monthly energy consumption from 2014 to within accepted error ranges for both the Mean Biased Error and Cumulative Variation of Root Mean Square Error [53]. All models use a Test Reference Year (TRY) simulation weather file relevant to each building's geographic location. Model geometries are based upon plan drawings and site surveys, as are the construction types. Inputs for lighting and equipment internal heat gains and operating schedules were defined using sub-metered data.

In the baseline models, occupancy schedules are based upon the NCM guidance. It is important to note that the density of occupancy in all models (both NCM and metadata versions) is based upon the NCM input and is set at $4.28m^2$ per person, with each person representing $140W/m^2$ of internal heat gain [9]. The metadata schedules control the density of occupants during a given hour by using a proportional factor (between 0-1). Although the metadata can be used to define occupancy schedules, it cannot be used to define the absolute number of people within a specific facility. The absolute number of occupants is calculated in the model based upon the $4.28m^2$ value; for example, a room with

a floor area of 428m^2 would contain 100 people during the hour of peak occupancy. For the schedules to be of general use for building types, an assumed value of this nature is necessary, especially when models are being created for new facilities with no historic or projected occupancy data. However, in absolute terms, this does represent a limitation of the method which is discussed later in section 5. It should also be noted that the schedules created here have been used for all zones in each model. The majority of occupants in these buildings are visitors (shoppers) but staff occupancy in non-public areas is assumed to follow a similar pattern. In particular, the metadata schedules may not be suitable for transitional spaces such as bathrooms and non-public circulation areas as highlighted in the literature [8]. The majority of the research found in the literature focuses on office buildings so further work is required to establish whether large publically accessible facilities have similarly stochastic patterns of occupancy in transitional zones. It is also possible that some of the location data comes from mobile devices owned by staff; all of these limitations are reviewed in the discussion section 5.

Infiltration rates vary across the sample models but are all active for 24 hours per day as a background air exchange. As discussed in the literature review, occupants of buildings often have a wide range of adaptive opportunities and other behaviours that have an impact on energy and thermal performance. However, in many commercial buildings, this is not the case. In the real buildings that these models are based on, building operations are closely controlled by facility managers and occupants have a very limited adaptive range (mainly limited to adjusting clothing levels).

There are two other model inputs that are directly affected by the metadata schedules. The mechanical ventilation rates are the same in all models and are fixed at 10 litres per second per person. As this ventilation rate is linked to the density of occupants, this model input is affected by the introduction of the metadata schedules. Natural ventilation rates in the models are also linked to occupancy schedules as air is exchanged through the main entrance to the building. The amount of time that these doors are open is directly linked to the occupancy schedules (i.e. customers entering and exiting the buildings). Both of these model inputs will have an impact of energy and thermal

performance as they account for a large proportion of convective heat exchange. A summary of the main model inputs that are controlled by operating schedules is presented in Table 1; it is included for indicative purposes only. Some inputs are noted as having a variable value, this is due to the characteristic differences between each of the case study buildings used in this work. The diurnal cycles relate to opening hours, with the maximum value occurring during standard opening hours and lower set back values outside of these times.

Table 1. Summary of inputs controlled by operating schedules in case study building energy models.

| Input	Operating profile	Unit	Value
Occupancy	Metadata schedules	m^2/person	140 W/person
Lighting	Diurnal cycle	W/m^2/100 lux	Variable
Equipment	Diurnal cycle	W/m^2	Variable
Natural ventilation	Metadata schedules	Air changes/hour	Variable
Mechanical ventilation	Diurnal cycle	l/s/person	10 l/s/person
Heating	Diurnal cycle	°C set point	18°C (opening hours)
Cooling	Diurnal cycle	°C set point	24°C (opening hours)

4. Metadata occupancy schedules

Two alternative opening patterns were evident for supermarket buildings in the UK; the first is for stores that operate over 24 hours and the second is for stores that open on a daily basis (referred to as 'daily' in the remainder of this paper). It is proposed that these are treated as two different building operating types for energy modelling purposes. This is necessary due to the direct impact occupancy

and operations have on energy consumption. In all cases, Sunday opening hours are restricted to a 6-hour period as is dictated by UK law.

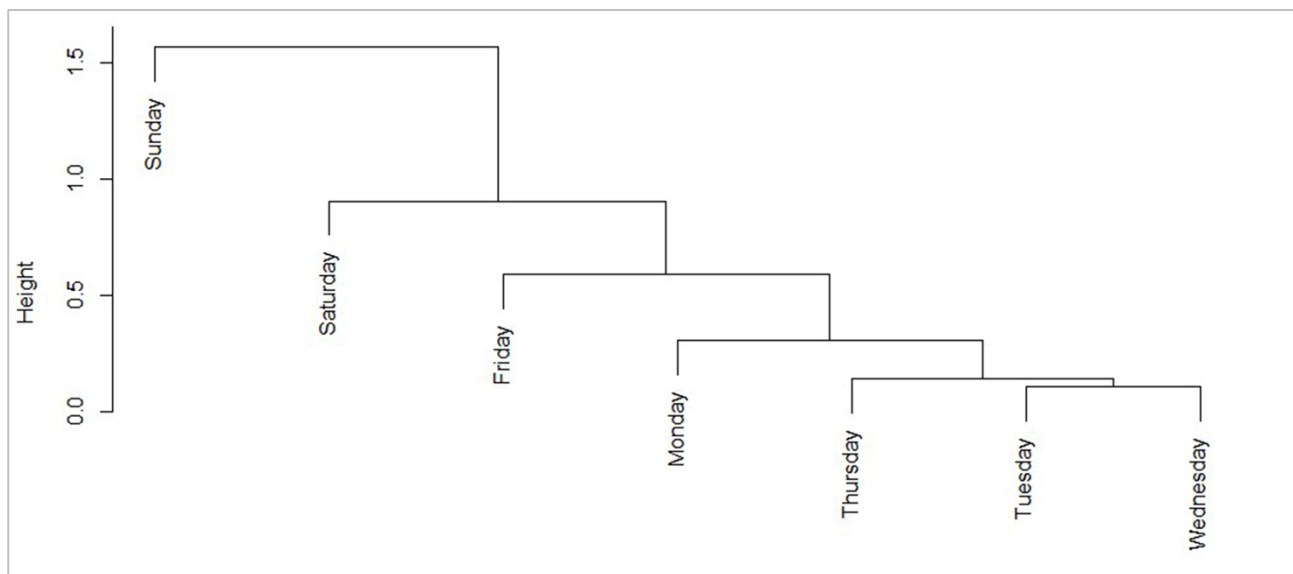## 4.1. Stores with 24 hour opening patterns

At the time this research was completed, there were two supermarket brands operating 24 hour stores in the UK. The brands are referred to here as brands A and B to preserve anonymity. Metadata was collected for 10 different stores from each brand portfolio. The average daily profiles for both brands are illustrated in Fig. 2.

(a) Brand A *(b) Brand B*
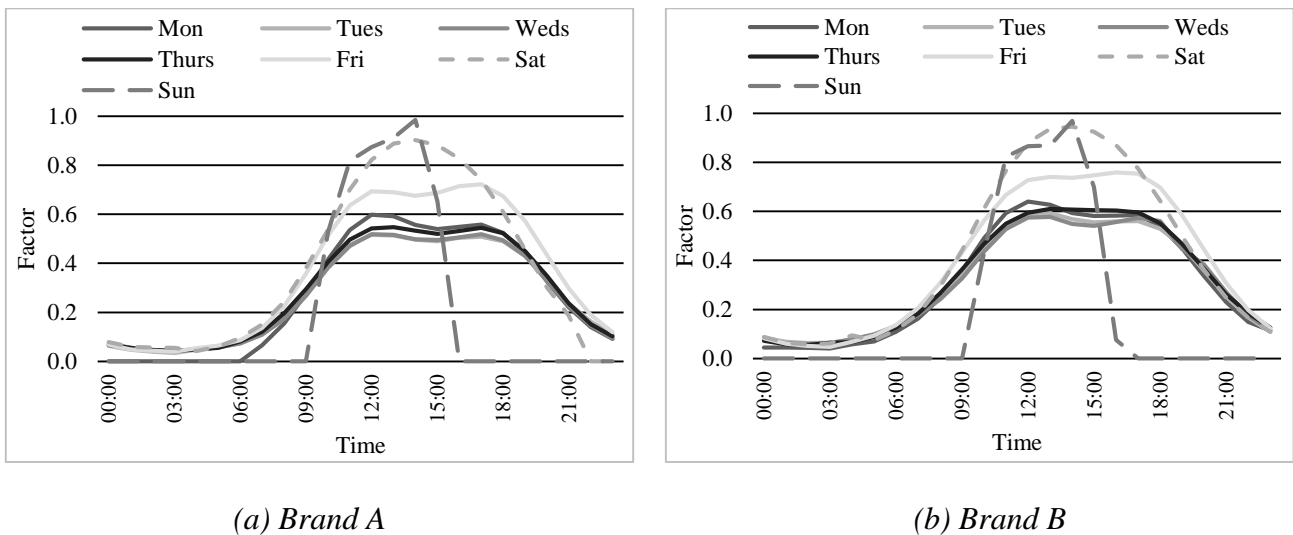
Fig. 2. Average daily metadata occupancy schedules for stores with 24 hour opening.

Existing schedules included within the NCM specify discrete profiles for weekdays, Saturdays and Sundays (although these are identical as illustrated in Fig. 4). The metadata collected for the stores with 24 hour opening indicates that it is reasonable to use a single weekday schedule as weekdays follow a similar pattern for both brands. This pattern differs significantly from the NCM profile, which includes a drop in occupancy levels around midday. The metadata suggest that occupancy actually

peaks during these periods. To confirm the distinct differences in occupancy profiles between days, the Euclidean distance between the daily average profiles was calculated and hierarchical clustering was carried out to identify any grouping. The dendogram of this data is displayed in Fig. 3, where the y axis denotes the Euclidean distance between the groups which are connected by a horizontal line. It is apparent that Monday to Thursday have a small difference in distance suggesting they are very similar, with the Friday profile lying at a slightly greater height suggesting some similarity, albeit with a weaker relationship. The Saturday and Sunday profiles are noticeably different.

Fig. 3. Dendogram illustrating relationship between metadata occupancy schedules for stores with 24 hour opening.

The metadata also confirm that the separation of schedules for Saturday and Sunday is realistic as they differ by a sufficient degree from the weekday profiles, and from each other, for the creation of discrete schedules to be justified. Occupancy on Saturdays follows a similar pattern to the weekdays but has higher peaks during the middle of the day. The different occupancy pattern for Sundays is dictated by legislative opening hours but peak occupancy in all but one of the sample stores occurred

on a Sunday. The average metadata schedules for 24 hour stores are compared with the corresponding NCM schedules in Fig. 4. As with the weekday case, the weekend metadata schedules peak around midday rather than fall as in the NCM estimates.
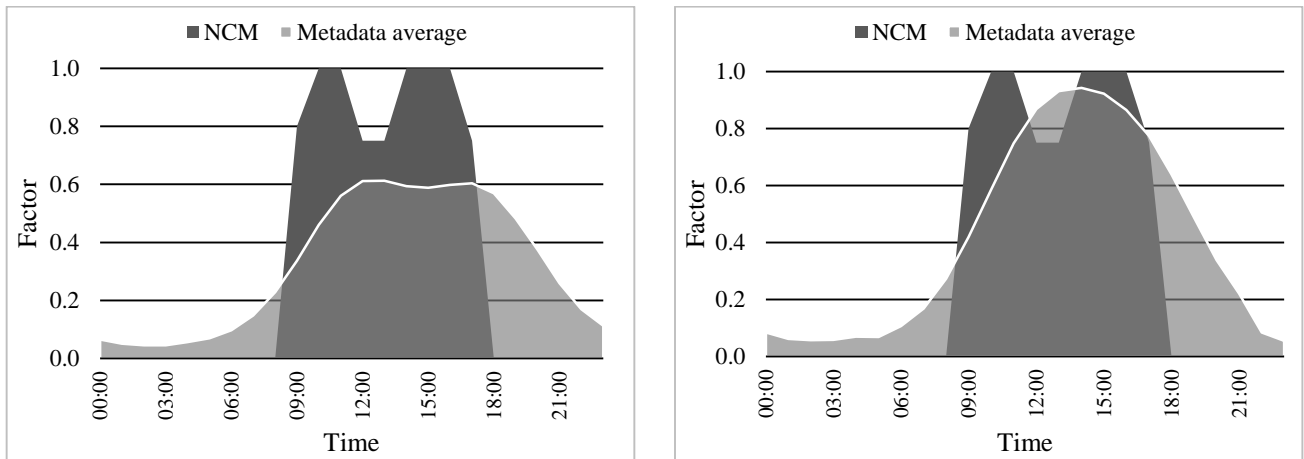

(a) Weekdays

(b) Saturdays
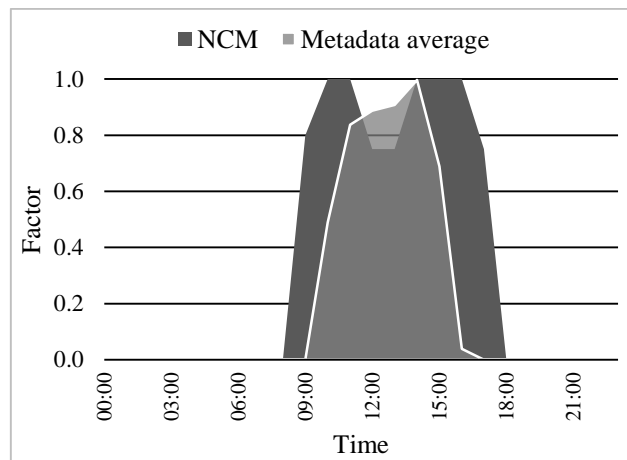
(c) Sundays

Fig. 4. Average metadata occupancy schedules for stores with 24 hour opening.

Although the profiles of the NCM and the metadata schedules are visibly different, the difference was quantified statistically. To achieve this, particular groups were selected (Weekday, Saturday or

Sunday) and a Wilcoxon signed rank test was performed on each time bin of the data. This test is used to determine if the NCM profile occupancy values lie outside the region predicted by the new profile data. A p-value of less than 0.05 was taken as being indicative of a statistically significant difference, and times at which both models predict an occupancy of 0 were ignored. The percentage of the time bins for which the p-value was less than 0.05 was then calculated for each profile. A percentage of 100 would therefore suggest all of the data points in the new profile are statistically different from the NCM values. The percentage test results were 100% for Monday-Friday, 96% for Saturday and 89% for Sunday, demonstrating a large statistical difference between the NCM profiles and all of the metadata schedules.

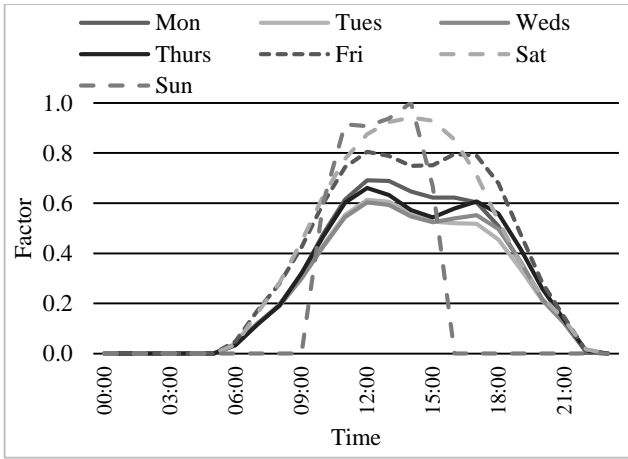## 4.2. Stores with daily opening patterns

A wider sample of data was used for stores with the more commonly used daily opening patterns. Metadata was collected for 10 stores from each of the 6 leading supermarket brands (anonymised here as brands A-F). The average daily schedules for all brands are shown in Fig. 5.
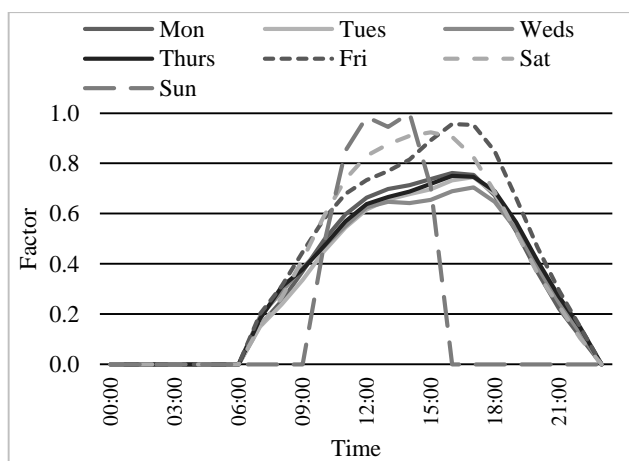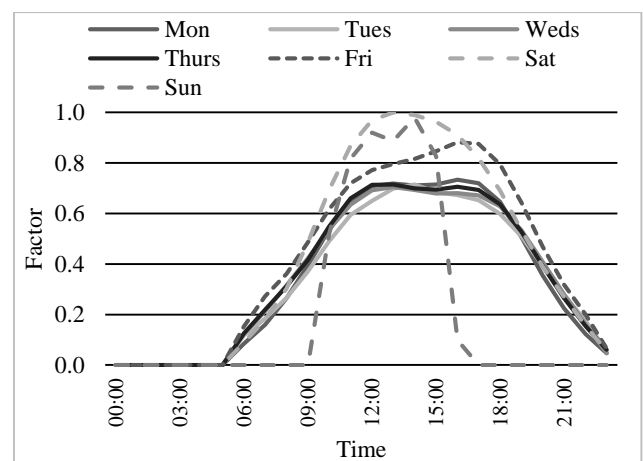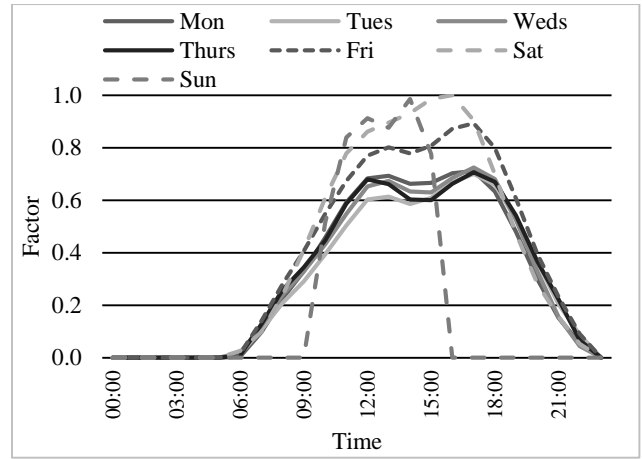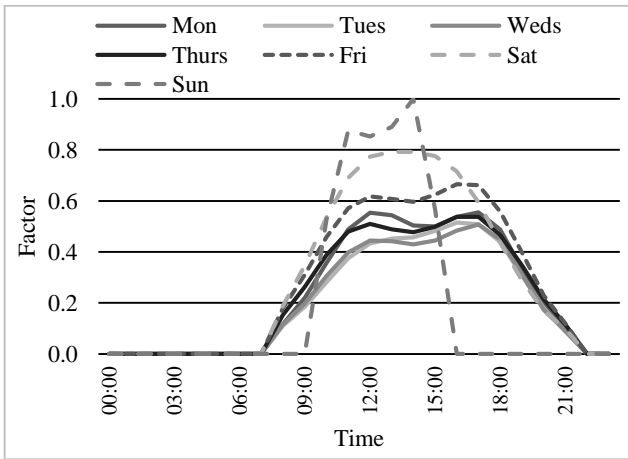

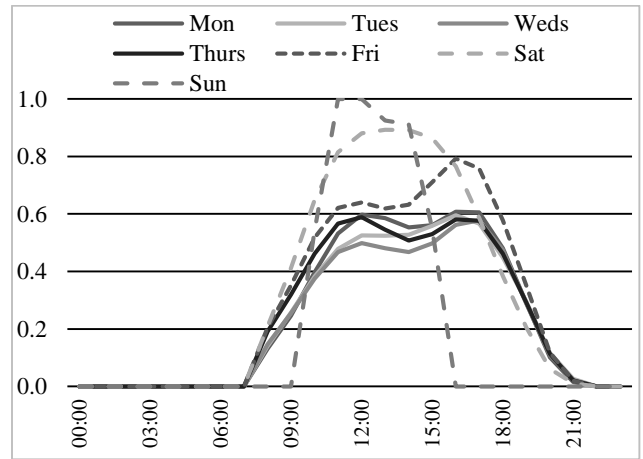*(a) Brand A*   *(b) Brand B*

(c) Brand C

(d) Brand D

(e) Brand E

(f) Brand F

Fig. 5. Average daily metadata occupancy schedules for stores with daily opening.

There is more variance found in the opening times of specific stores that use daily patterns. This can differ between brands but also between stores within the same brand group. However, a similar pattern of occupancy is evident in the metadata schedules across the group of stores with daily opening patterns. Following the same procedure described in section 4.1, a dendogram was plotted to illustrate the relationships between daily occupancy profiles. The Monday to Thursday values again lie very close to each other, with results indicating that bespoke profiles are required for occupancy for Fridays, Saturdays and Sundays. The results of this exercise are illustrated in Fig. 6.
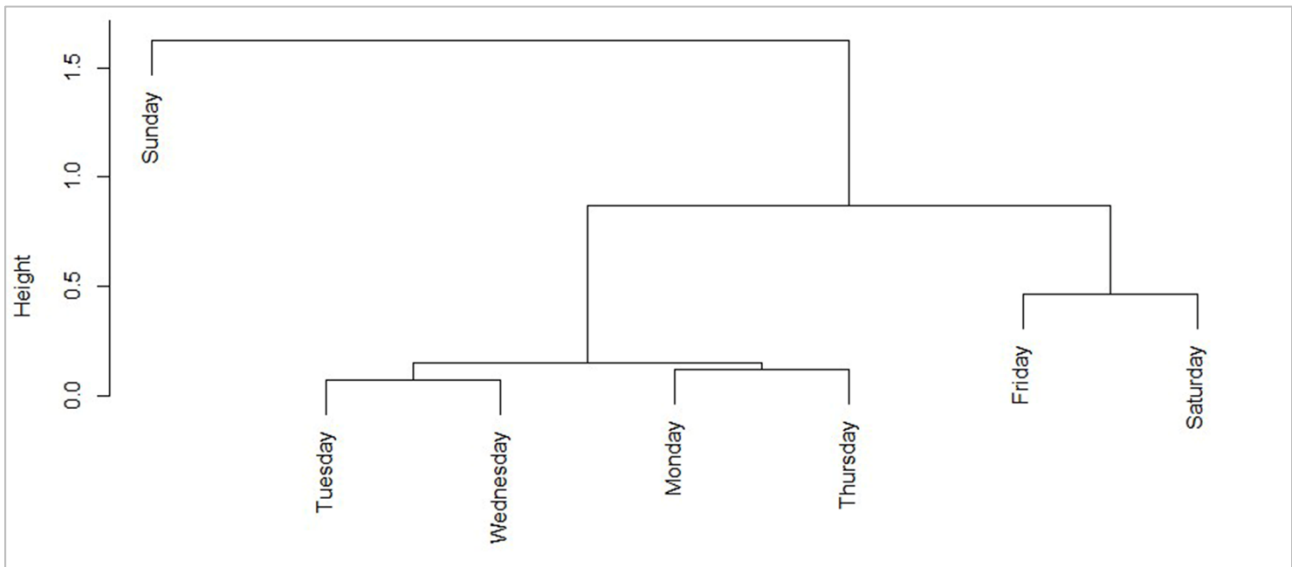
Fig. 6. Dendogram illustrating relationship between metadata occupancy schedules for stores with daily opening.

In keeping with the stores using 24 hour opening patterns, the peak occupancy levels occur at the weekend for stores with a daily opening pattern. Although there is some consistency between most weekdays, both the visual pattern and statistical analysis confirm that Fridays for all the brands, include significantly higher peaks than the other weekdays. Therefore, a separate average schedule was created for Fridays for stores using daily opening patterns. The average occupancy schedules for stores with daily opening are shown in Fig. 7.
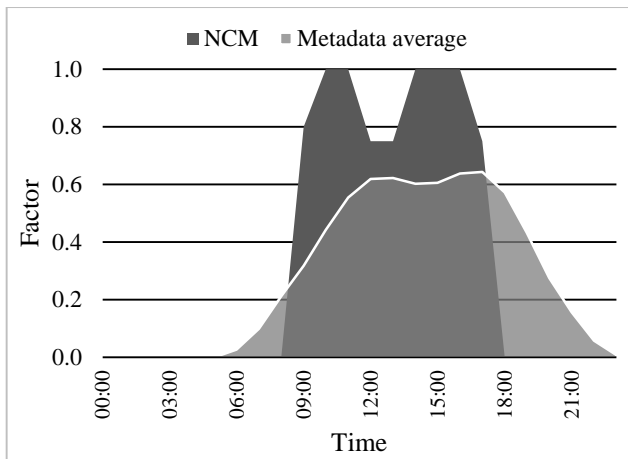
(a) Monday – Thursday
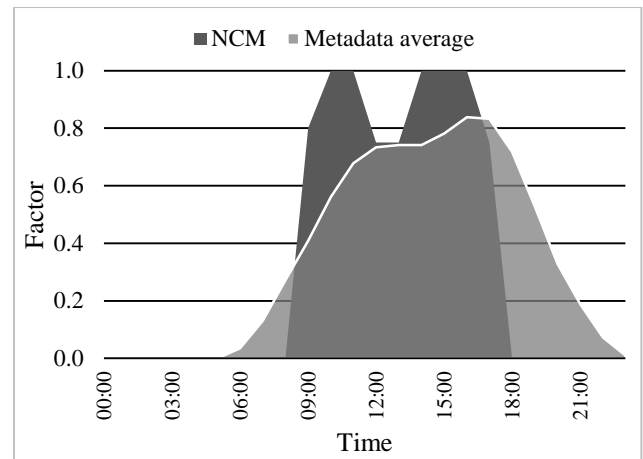
(b) Fridays
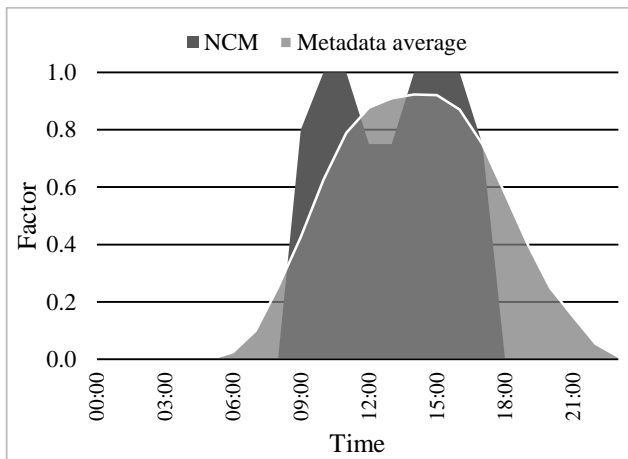
(c) Saturdays
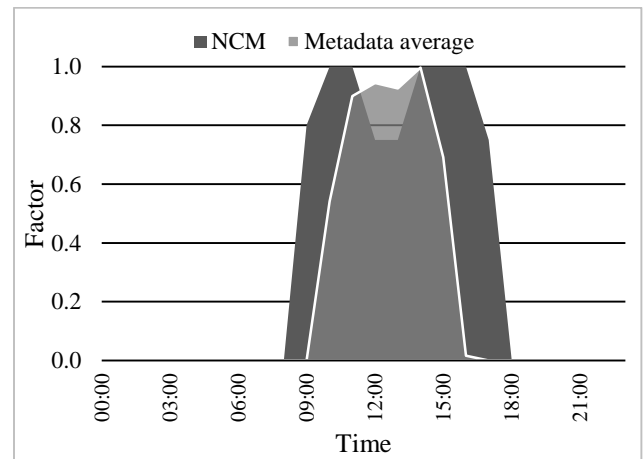
(d) Sundays

Fig. 7. Average brand specific and aggregated metadata occupancy schedules for stores with daily opening.

As with the 24 hour opening profiles, there is again no fall in occupancy around midday and stores are open for longer periods than the default NCM estimates (apart from Sundays). Occupant density from Monday to Thursday is also lower than the NCM estimates. Peak occupancy for all of the discrete schedules actually occurs when the NCM profiles estimate occupant density to begin falling. These differences have an impact on the energy consumption and peak demands predicted by building energy models as demonstrated in the section 4.3. The percentage of data points which are statistically

different to the NCM values were again calculated following the same procedure as described in section 4.1. The percentage values were 100% for Monday-Thursday, 96% for Saturday and 89% for Sunday. Although the similarity between the metadata profiles and the NCM profile is very low in all cases, it can be seen that this is particularly the case for weekdays for both 24 hour opening and daily opening stores.

*4.3. Impact of metadata profiles on building energy model outputs*

To quantify the impact of using the metadata schedules in DSM calculations, they were used as inputs in 20 different supermarket building energy models. Outputs were then compared with those from baseline models using the NCM occupancy schedules. Outputs were compared for total annual boiler heating and chiller cooling energy consumption and peak heating and cooling power demand. Outputs for the peak percentage of people dissatisfied (PPD) and internal CO_2 concentration (in parts per million - ppm) for the shop floor areas (accounting for the majority of total floor area in all cases) were also compared to illustrate the impact on internal thermal performance and air quality. Model numbers 1-10 are supermarkets using 24 hour opening patterns, model numbers 11-20 are supermarkets using daily opening patterns. To allow for results to be more easily visualised, the y-axis in Fig. 8. has been truncated at -75% as some percentage differences in chiller energy consumption were very large. This is due to the very low chiller energy demand in the baseline models of smaller stores; although absolute changes in terms of energy consumed are relatively small, the proportionate change is large. Specifically, store 16 had a decrease of -142.86%, store 19 of -180.00% and store 20 of -88.46%.
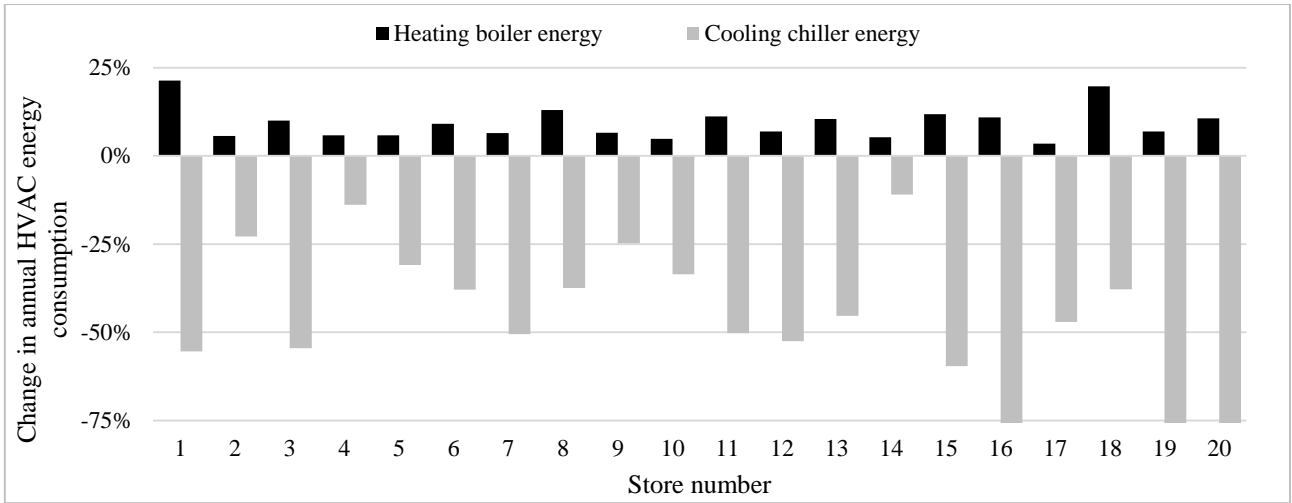
Fig. 8. Percentage difference between baseline model and metadata model annual space conditioning energy consumption.
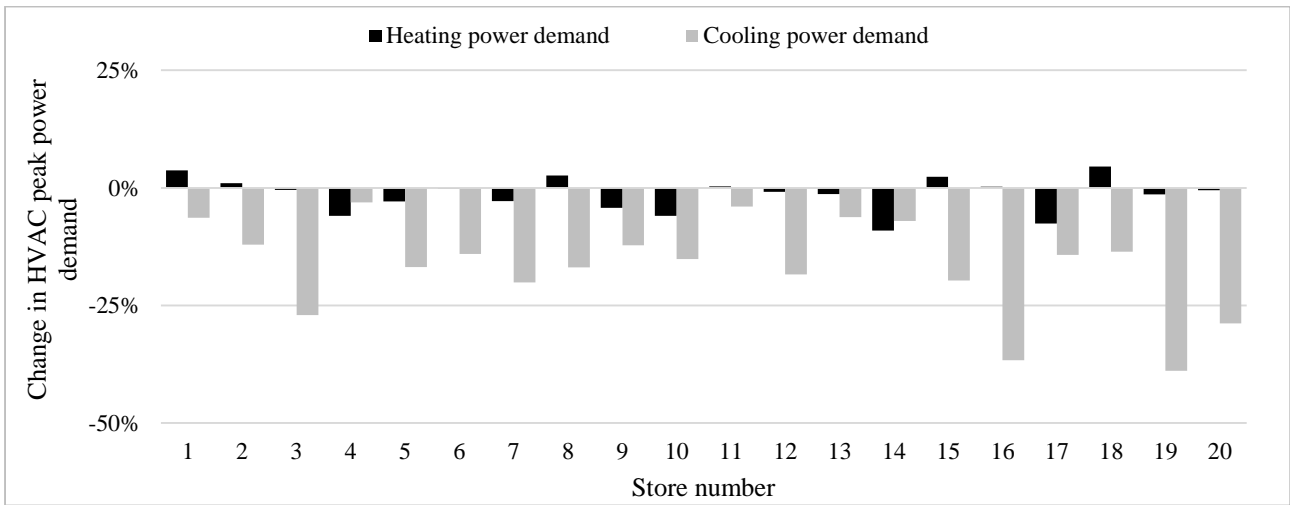
Fig. 9. Percentage difference between baseline model and metadata model peak space conditioning power demand.
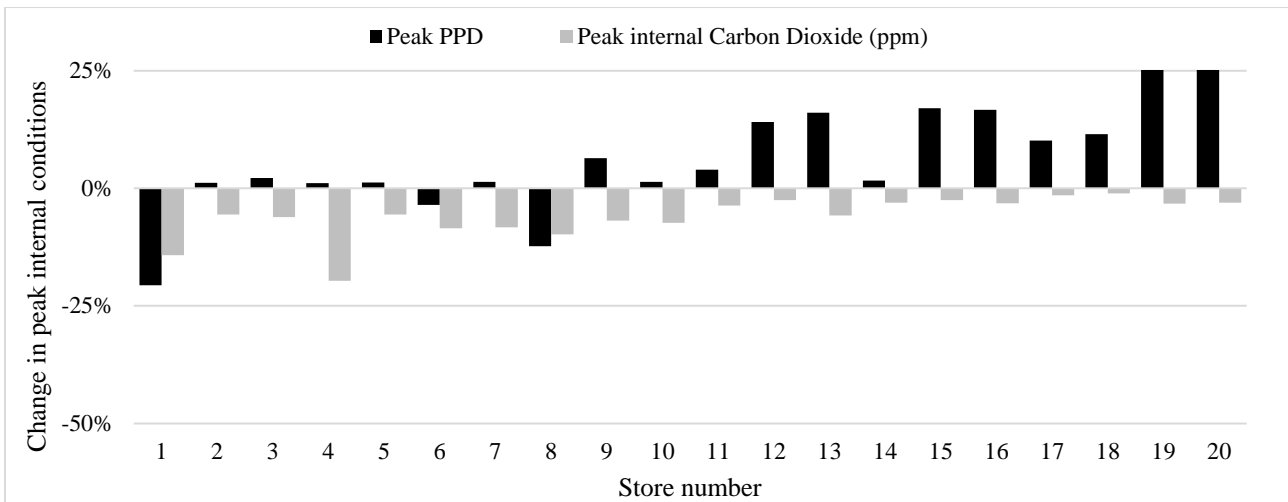
Fig. 10. Percentage difference between baseline model and metadata model peak internal conditions.

In all cases, the estimated total annual boiler energy consumption for space heating increases, as can be seen in Fig. 8. The average increase in boiler heating energy was 9.32%, stores with 24 hour opening had a slightly lower average increase of 8.89% and stores with daily opening increase by an average of 9.76%. These results are due to the decrease in cumulative internal heat gains associated with the building occupants and also the increased heat loss driven by natural ventilation that is in turn due to extended periods of occupancy. As stated in section 3.3, the natural ventilation air change rate through customer doorways is controlled using the occupancy schedule. The mechanical ventilation rate is also linked to the number of occupants. Although there will be less heat loss associated with natural ventilation compared to the times that the NCM occupancy profiles are active, using the metadata schedules results in additional heat loss outside of these times, when external temperatures are often lower due to diurnal cycles. In contrast to this, annual cooling energy consumption decreases in all cases due to the lower density of occupants during the weekdays in particular. The average for all models decreased by 53.82%, with 24 hour stores decreasing by an average of 36.17% and daily opening stores by an average of 71.47%. As many of the daily opening stores were smaller in size and had relatively low cooling consumption in the baseline models, the average value for these stores is

exaggerated. These differences would contribute significantly to gaps in performance between the estimated and actual energy consumption of each store.

Peak cooling power demand reduces in all cases for the same reasons that influence the fall in overall consumption as can be seen in Fig. 9. There is a certain degree of coincidence associated with this, as, although the selected models are in different geographical locations, the peak external temperatures for all cases happen to fall on weekdays when occupant density is lower and internal heat gains are reduced. It is possible that these could increase if these peak temperatures occurred around midday over a weekend period. The average change in peak cooling demand was -16.56% for all stores, with the 24 hour stores having an average value of -14.37% and the daily opening stores an average of -18.75%. Changes in peak heat demand are inconsistent between stores. This is again due to when peak demand occurs. All of the stores with an increase in peak heating demand are in stores where the peak occurs when the NCM occupancy schedules are not active. The stores with a fall in peak heating demand experience their peak during times when the metadata schedules dictate a higher occupancy rate (between 11:00 and 13:00 hours). Differences in estimated peak power demand could have implications for HVAC plant sizing. They could also have an impact on forecast demands for local power grids.

The peak percentage of people dissatisfied (PPD) increases most consistently in the stores with daily opening patterns. As mentioned previously, many of these are smaller stores and the increase in PPD relates to a period of higher occupancy in warmer conditions. The average change in all stores was 6.46% although 24 hours stores actually saw an average decrease of -2.16%, with the daily opening stores increasing by 15.08%. Peak internal CO_2 levels reduce in all cases, by an average value of -6.07% (-9.18% for 24 hour stores and -2.96% for daily opening stores). In total, internal CO_2 levels would be expected to fall as occupancy peaks are not as prolonged as those found in the NCM. There is a fixed mechanical ventilation air volume per person and natural ventilation rates are linked to the occupancy profile, meaning that the metadata models with less dense occupancy experience a

reduction in internal CO_2 levels. This change would affect the analysis of retrofit CO_2 demand control ventilation strategies. The increased mechanical ventilation under the NCM profiles could lead to higher estimated heat loads and in turn mean that this type of retrofit could appear to be financially unviable.

5.   Discussion

Holistically, the location metadata used in this work has a very broad range of potential applications that fall outside the scope of this work. The source of the primary data comes from analytic services delivered by Google that provide businesses with valuable commercial data, useful for the management of their assets and their commercial development. It also has the scope for wider public applications, such as to aid the management of urban infrastructure systems in future smart cities. Subjectively however, it is useful for the type of modelling work described here, particularly when no empirical occupancy data is available.

There are two ways in which the metadata occupancy schedules described in this paper can be applied to building energy models: creating bespoke occupancy profiles for specific buildings; and the creation of occupancy profiles for building archetypes. In the primary data, there are differences between the daily profiles created for individual stores and other stores in the brand group. Similarly, there were differences between the average daily profiles for each brand which is evident when comparing the charts shown in Fig. 5. The results illustrated in section 4.3 show that the difference in profiles has an impact on modelled energy performance. It could therefore also be useful to create building-specific schedules and average schedules for a particular brand's own portfolio of stores. It is worth noting that the DSM models used to quantify the impact of the metadata schedules on energy performance are relatively simple in terms of geometry, the number of zones and construction types. Further work is required to quantify the impact these schedules would have in more complex building

energy models. For example, there were no changes made to profiles that could respond to occupancy schedules, such as those controlling lighting and equipment.

The second application of the occupancy data is the creation of average occupancy profiles for building types that could be used by energy modellers and ultimately developed for use in the NCM in the UK and other similar regulatory compliance calculations in the rest of the world. This paper demonstrates that this methodology can work in theory and further work is required for this method to be refined, though it does appear that the existing NCM profiles need updating if models are intended to provide a relatively realistic prediction of performance. The first area of further work will focus on a more efficient method of harvesting the metadata at source. The method described in this paper is time consuming and a more automated means of collecting it would make it more practicable.

Further work should also aim to check the validity of the personal location metadata itself. At present, the data displayed by Google represent an average over several preceding weeks. Before wider application, it will be necessary to regularly check the metadata to identify any temporal differences in schedules. The data collected in this work covers a limited geographical range, so similar checks would need to be made to compare any geographic differences in the data for specific building types. It will also be necessary to check the validity of the metadata schedules against detailed occupancy data collected within specific buildings. Despite there being some evidence of a wide spread of demographic groups using smart phones [13], the potential bias caused by metadata only representing users with smartphones and location data turned on should also be investigated further.

There is also potential bias relating to the location positioning technology. Currently, Google use a combination of three data sources to locate devices: GPS: satellite location data, usually accurate to within a few metres; Wi-Fi: local Wi-Fi connections help position users; Cell Tower: connection to a cellular network, only accurate to a few thousand metres [51]. In many instances, the GPS location data will suffice to differentiate between adjacent stores, especially larger buildings that are detached from other structures. This does however represent another limitation as it only helps to resolve the

problem on a horizontal plane. For example, there are facilities in major cities located on different storeys of the same building. It may be that there is more of a reliance upon tagged Wi-Fi connections in these circumstances. Google does produce follow up data requests, such as requests to confirm a user was in a recent location which could also help improve location accuracy over time. Further research aimed at validating the metadata against *in situ* monitored occupancy data can also be designed to test the effectiveness of the positioning technologies for specific facilities.

It may be possible that the metadata schedules can be used in conjunction with other occupancy monitoring or modelling approaches. A simplistic example of this would be using a known maximum absolute occupancy value for the peak hour of the profile and calculating all other hourly values using the metadata schedules. A more sophisticated example would be to use the metadata schedules as a baseline to develop stochastic agent-based models to capture the dynamic nature of real occupancy patterns. It may also be the case that the metadata schedules help to capture some of the dynamic changes in building occupancy due to them being iteratively updated from the latest location data.

Despite there being further work required before this methodology can be used at scale, the scope for its application is potentially very broad. Returned results from search engines confirm that this type of data is available for many commercial buildings and some public facilities, including restaurants, bars, shops, cinemas, leisure centres, university buildings, museums and airports for example. This type of data is not limited to the UK and search results for similar facilities in other countries have the same results, meaning that this method has the scope for global application.

6.   Conclusion

The work undertaken here has demonstrated that publically available personal location metadata can be collated to create occupancy schedules for use in building energy models. These schedules can be created for specific buildings, groups of buildings within a brand portfolio or for different building

types. For the supermarket building-type used here, there is suitable consistency between the different store and brand profiles for average building-type level schedules to be created.

The metadata schedules defined in this work differ considerably from those currently used for regulatory compliance calculations in the UK and have a significant impact on predicted energy performance. They have the potential to offer a more realistic estimate of occupancy profiles that can be used in future updates of the NCM guidance, not only for this building-type, but for a wide range of commercial and non-domestic building-types on an international scale. In the small sample presented here, use of the estimated NCM profiles resulted in heating energy consumption being underestimated by an average of 9.32% and cooling loads being underestimated by an average of 53.82%. The accuracy of building modelling can be improved by removing elements of uncertainty. In relation to model inputs, it is always preferable for them to be based upon the most reliable source of data. The method proposed in this paper offers a means of helping to remove uncertainty and therefore iteratively improve the accuracy of model predictions.

This work provides a proof of concept and demonstrates its application, further work is required to rationalise the data collection methodology. In co-operation with Google it may be possible to use recorded long-term historical data to create bespoke annual schedules that could be used in detailed model calibration for specific stores. In particular, this would improve the accuracy of any DSM retrofit scenario analysis and evaluation. The use of long-term historical data could also be used to create annual occupancy profiles that capture seasonal fluctuations for groups of stores and building types.

References

[1] T.F. Stocker, D. Qin, G.K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, P.M. Midgley, IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change., in, Cambridge, 2013, pp. 1535.

[2] The Carbon Trust, Building the Future Today: Transforming the Economic and Carbon Performance of the Buildings We Work in, in, London, 2009.

[3] X. Feng, D. Yan, T. Hong, Simulation of occupancy in buildings, Energy and Buildings, 87 (2015) 348-359.

[4] D. Yan, W. O'Brien, T. Hong, X. Feng, H. Burak Gunay, F. Tahmasebi, A. Mahdavi, Occupant behavior modeling for building performance simulation: Current state and future challenges, Energy and Buildings, 107 (2015) 264-278.

[5] D. Coakley, P. Raftery, M. Keane, A review of methods to match building energy simulation models to measured data, Renewable and Sustainable Energy Reviews, 37 (2014) 123-141.

[6] E. Azar, C.C. Menassa, A comprehensive analysis of the impact of occupancy parameters in energy simulation of office buildings, Energy and Buildings, 55 (2012) 841-853.

[7] D. Yan, T. Hong, Definition and Simulation of Occupant Behaviour in Buildings, in, International Energy Agency, Paris, 2014.

[8] C. Wang, D. Yan, Y. Jiang, A novel approach for building occupancy simulation, Building Simulation, 4 (2) (2011) 149-167.

[9] HM Government, National Calculation Methodology (NCM) modelling guide (for building other than dwellings in England and Wales), in, BRE Ltd, London, 2013.

[10] HM Government, Developing DECC's Evidnece Base, in: D.o.E.a.C. Change (Ed.), Crown Copyright, London, 2014.

[11] P.M. Nadkarni, What Is Metadata?, in: Metadata-driven Software Systems in Biomedicine: Designing Systems that can adapt to Changing Knowledge, Springer London, London, 2011, pp. 1-16.

[12] D. Macro, Building and Managing the Meta Data Repository: A Full Lifecycle Guide, John Wiley and Sons, New Jersey, 2000.

[13] E. Marsden, P. Lee, There's no place like phone: Consumer usage patterns in the era of peak smartphone, in, Deloitte LLP, London, 2016.

[14] J. Palmer, N. Terry, P. Armitage, Building Performance Evaluation Programme: Findings from non-domestic projects, in, Innovate UK, London, 2016.

[15] A.C. Menezes, A. Cripps, D. Bouchlaghem, R. Buswell, Predicted vs. actual energy performance of non-domestic buildings: Using post-occupancy evaluation data to reduce the performance gap, Applied Energy, 97 (2012) 355-364.

[16] P. Raftery, M. Keane, J. O'Donnell, Calibrating whole building energy models: An evidence-based methodology, Energy and Buildings, 43 (9) (2011) 2356-2364.

[17] P. Raftery, M. Keane, A. Costa, Calibrating whole building energy models: Detailed case study using hourly measured data, Energy and Buildings, 43 (12) (2011) 3666-3679.

[18] J.M. Parker, D. Farmer, M. Fletcher, Calibrating whole house thermal models against a coheating test, in: System Simulation in Buildings 2014 Proceedings of the Ninth International Conference, December 10-12, 2014., Atelier des Presses, Liege, 2015, pp. 211-219.

[19] Y. Heo, R. Choudhary, G.A. Augenbroe, Calibration of building energy models for retrofit analysis under uncertainty, Energy and Buildings, 47 (2012) 550-560.

[20] M. Manfren, N. Aste, R. Moshksar, Calibration and uncertainty analysis for computer models - A meta-model based approach for integrated building energy simulation, Applied Energy, 103 (2013) 627-641.

[21] N. Aste, F. Leonforte, M. Manfren, M. Mazzon, Thermal inertia and energy efficiency - Parametric simulation assessment on a calibrated case study, Applied Energy, 145 (2015) 111-123.

[22] K. Sun, T. Hong, A simulation approach to estimate energy savings potential of occupant behavior measures, Energy and Buildings, 136 (2017) 43-62.

[23] Q. Wang, G. Augenbroe, J.H. Kim, A framework for meta- Analysis of the role of occupancy variables in the energy use of commercial buildings, in: 14th International Conference of IBPSA - Building Simulation 2015, BS 2015, Conference Proceedings, 2015, pp. 2278-2285.

[24] C. Liao, Y. Lin, P. Barooah, Agent-based and graphical modelling of building occupancy, Journal of Building Performance Simulation, 5 (1) (2012) 5-25.

[25] R. Melfi, B. Rosenblum, B. Nordman, K. Christensen, Measuring building occupancy using existing network infrastructure, in: 2011 International Green Computing Conference and Workshops, IGCC 2011, 2011.

[26] J. Yang, M. Santamouris, S.E. Lee, Review of occupancy sensing systems and occupancy modeling methodologies for the application in institutional buildings, Energy and Buildings, 121 (2016) 344-349.

[27] T. Hong, S.C. Taylor-Lange, S. D'Oca, D. Yan, S.P. Corgnati, Advances in research and applications of energy-related occupant behavior in buildings, Energy and Buildings, 116 (2016) 694-702.

[28] R. Ward, R. Choudhary, Y. Heo, A. Rysanek, Exploring the impact of different parameterisations of occupant-related internal loads in building energy simulation, Energy and Buildings, 123 (2016) 92-105.

[29] S. D'Oca, T. Hong, A data-mining approach to discover patterns of window opening and closing behavior in offices, Building and Environment, 82 (2014) 726-739.

[30] S. D'Oca, T. Hong, Occupancy schedules learning process through a data mining framework, Energy and Buildings, 88 (2015) 395-408.

[31] K.U. Ahn, C.S. Park, Correlation between occupants and energy consumption, Energy and Buildings, 116 (2016) 420-433.

[32] Y.S. Kim, M. Heidarinejad, M. Dahlhausen, J. Srebric, Building energy model calibration with schedules derived from electricity use data, Applied Energy, 190 (2017) 997-1007.

[33] Y.S. Kim, J. Srebric, Improvement of building energy simulation accuracy with occupancy schedules derived from hourly building electricity consumption, in: ASHRAE Transactions, 2015, pp. 353-360.

[34] M. Wang, X. Wang, G. Zhang, C. Li, Occupancy detection based on Spiking Neural Networks for green building automation systems, in: Proceedings of the World Congress on Intelligent Control and Automation (WCICA), 2015, pp. 2681-2686.

[35] T. Ekwevugbe, N. Brown, V. Pakka, D. Fan, Real-time building occupancy sensing using neural-network based sensor network, in: IEEE International Conference on Digital Ecosystems and Technologies, 2013, pp. 114-119.

[36] Q. Hua, H.B. Chen, Y.Y. Ye, S.X.D. Tan, Occupancy detection in smart buildings using support vector regression method, in: Proceedings - 2016 8th International Conference on Intelligent Human-Machine Systems and Cybernetics, IHMSC 2016, 2016, pp. 77-80.

[37] H. Zhao, Z. Qi, S. Wang, K. Vafai, H. Wang, H. Chen, S.X.D. Tan, Learning-based occupancy behavior detection for smart buildings, in: Proceedings - IEEE International Symposium on Circuits and Systems, 2016, pp. 954-957.

[38] Z. Chen, Y.C. Soh, Comparing occupancy models and data mining approaches for regular occupancy prediction in commercial buildings, Journal of Building Performance Simulation, (2016) 1-9.

[39] M.S. Gul, S. Patidar, Understanding the energy consumption and occupancy of a multi-purpose academic building, Energy and Buildings, 87 (2014) 155-165.

[40] A. Mahdavi, F. Tahmasebi, M. Kayalar, Prediction of plug loads in office buildings: Simplified and probabilistic methods, Energy and Buildings, 129 (2016) 322-329.

[41] A. Mahdavi, F. Tahmasebi, Predicting people's presence in buildings: An empirically based model performance analysis, Energy and Buildings, 86 (2015) 349-355.

[42] D. Wang, C.C. Federspiel, F. Rubinstein, Modeling occupancy in single person offices, Energy and Buildings, 37 (2) (2005) 121-126.

[43] J. Hutchins, A. Ihler, P. Smyth, Modeling count data from multiple sensors: A building occupancy model, in: 2007 2nd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, CAMPSAP, 2007, pp. 241-244.

[44] S. Meyn, A. Surana, Y. Lin, S.M. Oggianu, S. Narayanan, T.A. Frewen, A sensor-utility-network method for estimation of occupancy in buildings, in: Proceedings of the IEEE Conference on Decision and Control, 2009, pp. 1494-1500.

[45] C. Martani, D. Lee, P. Robinson, R. Britter, C. Ratti, ENERNET: Studying the dynamic relationship between building occupancy and energy consumption, Energy and Buildings, 47 (2012) 584-591.

[46] L. Pitt, P.R. Green, B. Lennox, A sensor network for predicting and maintaining occupant comfort, in: 2013 IEEE Workshop on Environmental, Energy and Structural Monitoring Systems, EESMS 2013 - Proceedings, 2013.

[47] T. Hong, H. Sun, Y. Chen, S.C. Taylor-Lange, D. Yan, An occupant behavior modeling tool for co-simulation, Energy and Buildings, 117 (2016) 272-281.

[48] X. Luo, K.P. Lam, Y. Chen, T. Hong, Performance evaluation of an agent-based occupancy simulation model, Building and Environment, 115 (2017) 42-53.

[49] C. Duarte, K. Van Den Wymelenberg, C. Rieger, Revealing occupancy patterns in an office building through the use of occupancy sensor data, Energy and Buildings, 67 (2013) 587-595.

[50] X. Liang, T. Hong, G.Q. Shen, Improving the accuracy of energy baseline models for commercial buildings with occupancy data, Applied Energy, 179 (2016) 247-260.

[51] Google, Google My Business Help: Popular times and visit duration, in, Google Ltd, 2016.

[52] J.M. Parker, F. Thomas, M. Fletcher, D. Glew, C. Gorse, Accounting for refrigeration heat exchange in energy performance simulations of large food retail buildings, Building Services Engineering Research and Technology, 30 (10) (2016) 355-365.

[53] ASHRAE, ASHRAE Guideline 14-2002, Measurement of Energy and Demand Savings, in, Atlanta, 2002.