Identifying occupancy patterns and profiles in higher education institution buildings with high occupancy density – a case study

Taylor & Francis
Taylor & Francis Group

Intelligent Buildings International

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tibi20

Identifying occupancy patterns and profiles in higher education institution buildings with high occupancy density – A case study

Bashar Alfalah, Mehdi Shahrestani & Li Shao

To cite this article: Bashar Alfalah, Mehdi Shahrestani & Li Shao (2022): Identifying occupancy patterns and profiles in higher education institution buildings with high occupancy density – A case study, Intelligent Buildings International, DOI: 10.1080/17508975.2022.2137451

To link to this article: https://doi.org/10.1080/17508975.2022.2137451

Published online: 06 Nov 2022.

INTELLIGENT BUILDINGS INTERNATIONAL
https://doi.org/10.1080/17508975.2022.2137451

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS

Check for updates

Identifying occupancy patterns and profiles in higher education institution buildings with high occupancy density – A case study

Bashar Alfalah[a,b], Mehdi Shahrestani [a] and Li Shao [a]

[a]School of the Built Environment, University of Reading, Reading, UK; [b]College of architectural and planning, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia

ABSTRACT

Building occupancy patterns are an important factor in considering the energy efficiency of buildings and a key input for building performance modelling. More specifically, the energy consumption associated with heating, cooling, lighting, and plug load usage depends on the number of occupants in a building. Identifying occupancy patterns and profiles in buildings is a key factor for the optimisation of building operating systems and can potentially reduce the performance gap between the planning stage and the actual energy usage. This study aims to identify the patterns and profiles of the occupants in a selected case study building in England. In this study, occupancy data were collected over 12 months at five minutes intervals. A sensor was used to obtain high accuracy occupancy data compared to previous studies that encountered uncertainties in data collection. A set of clustering analyses was carried out to identify occupancy patterns and profiles in the building. The results of this study identified three different occupancy patterns and profiles as well as four drivers that influenced the occupants in the case study building: the beginning of the academic term, the examination period, the weekday/weekends, and the vacation driver.

ARTICLE HISTORY
Received 18 February 2022
Accepted 11 October 2022

KEYWORDS
Occupancy; patterns and profiles; sensor; cluster analysis; educational building

1. Introduction

In 2020, the Department for Business, Energy and Industrial Strategy (2020) published a report on UK energy consumption, stating that the building accounts for approximately 40% of final energy consumption and around 35% of total energy-related carbon dioxide emissions. The understanding of occupancy patterns considerably helps to reduce the performance gap, as the use of occupancy data can be informative for buildings to operate more efficiently. However, occupants were not considered in the design phase of building energy systems. The elimination of an important factor, such as occupants, is one of the reasons for the high-performance gap between the predicted energy consumption of buildings at the design stage and actual energy consumption in operation. Previous studies on occupants and their influence on energy consumption have focused on residential (Ioannou, Itard, and Agarwal 2018; Jang and Kang 2016; Razavi et al. 2019; van den Brom, Meijer, and Visscher 2018) and commercial buildings (Ashouri et al. 2019; Azam et al. 2019; Hobson et al. 2020; Jia et al. 2018). However, there is a lack of occupancy data related to occupants' presence in educational buildings such as university buildings. Previous research shows that educational buildings not only consume a relatively large amount of energy but also tend to have high-performance gaps (Andersen et al. 2014; Bernardo and Oliveira 2018; Chatterton 2018). The lack of knowledge of the building occupants in such high-density buildings results in a high-performance gap, as reported in previous studies, to be up to 85% in electrical energy consumption (Amber 2017; Kim et al. 2021). Therefore, the knowledge and understanding of the occupancy profiles will contribute to the mitigation of the performance gap between

CONTACT Mehdi Shahrestani m.shahrestani @ reading.ac.uk

the estimated energy consumption by building simulation tools at the design stage and the actual energy consumption of buildings. Moreover, facility managers can operate building systems based on the density and presence of building occupants.

Several studies have aimed to identify occupancy patterns in different types of buildings. Studies such as Liang, Hong, and Shen (2016) aimed to identify occupancy patterns to predict future occupancy numbers in office buildings. The study collected occupancy data for one year using occupancy sensors. Unsupervised machine learning cluster analysis was utilised to group the data collected into a few cohesive clusters (Everitt et al. 2011), resulting in identifying four patterns. In another study, Mora et al. (2019) aimed to identify occupancy patterns in a university office building. The data was collected using occupancy sensors for 45 days. The application of hierarchical clustering techniques identified three occupancy patterns. In another study, Chang and Hong (2013) analysed occupancy status to find out its impact on energy consumption on three floors of an open-space office building. The presence and absence occupancy data were collected for six months using lighting switch sensors. The interpretation of the data collected identified five patterns. In Yuan et al. (2020), the aim was to assess the energy consumption of a single office in a university building. The occupancy data were collected for two months. A PIR was used to collect the occupants' status and the occupants were asked to record their arrival and departure times. The collected data was used to develop lighting system control. The results were predictive of one-day occupancy status. Another study carried out by Wang and Shao (2017) identified a set of occupancy patterns associated with one room of a library building. Wi-Fi technology was used to track the signal of the occupants' smartphones as they entered and exited a room for a month. The result of applying cluster analysis revealed four occupancy patterns that were used to identify energy waste. Similarly, Tekler et al. (2020) track the occupants using Bluetooth signals in office areas in an academic building. The data was collected for 25 days during office hours. The study identified the patterns and profiles of the occupants during workdays. Previous studies collected occupancy data for various periods. However, there was the exclusion of weekends and days with few or no occupants. The exclusion has limited the accuracy of identifying occupancy patterns and profiles. The major contributions of this study are listed as follows:

- Presenting the results of a comprehensive study on occupancy numbers in a higher education building, which has not been studied sufficiently in the previous research for an extended period of 12 months, covering all weekdays for more robust identification of occupancy patterns and profiles.
- Developing a robust data collection strategy using high-accuracy infrared video image sensors eliminates uncertainties associated with the collection of occupancy data and, consequently, the study results.
- Identifying the key occupancy patterns and profiles in a high-density building as well as the drivers that influence them. To the best of our knowledge, this study is the first to identify the occupancy patterns and profiles in such buildings for an extended period, which can contribute to reducing the uncertainties associated with the outcomes of building simulations as well as the safety and security of the building.
- Proposing a robust approach for data collection and analysis that can be adopted to develop a new benchmark for buildings of higher education institutions to be used in future studies.

2. Related work

Several techniques were used in previous studies to collect the number of occupants in different types of buildings. These techniques include observation, survey, secondary data, and technologies. In this study, the focus is on the advantages and features of technologies used to record the number of occupants, including video counters, WIFI trackers, infrared beams, thermal imaging systems, moving horizon estimation, and infrared video image analysis. Video counting features stereoscopic imaging using computer vision through an embedded device. The embedded system utilises an algorithm that provides accurate indoor or outdoor locations. Another technology is WIFI counting uses a WIFI which captures a unique signal emitted by smartphones within a certain range. This has a drawback due to the possibility that not everyone carries a smartphone (Jallow and Hong 2018; Liu, Makino, and Mase 2010). People counters such as infrared beam counters and thermal counters can be used in low light conditions and do not pry into people's privacy. Infrared beam counters are mounted across the entrance of a building and linked to a display unit. The number of people leaving and entering the building is recorded when broken the beam. Despite the relatively low accuracy of infrared bean counters, which is in the range of 60–80%, these counters are still in use for counting the occupancy in buildings due to their low maintenance and installation cost (Jallow and Hong 2018). Thermal counters use thermal imaging systems consisting of sensors that detect heat sources used for counting. This technology has an accuracy of 80% to 85% (Rashid, Chowdhury, and Nawal 2016). Moving horizon estimation uses the histories of CO_2 and airflow sensors to measure occupancy in a building. Although the technology is accurate, it can overestimate occupancy due to infinite-dimensional problems (Zavala 2014). Infrared video image analysis incorporates several features, resulting in high-quality images suitable for a wide range of applications requiring occupants counting or tracking with high accuracy of 98% (Zhang, Shen, and Zhang 2016). It has no privacy or security issues and features easy occupancy management through a live feed. Occupancy information such as occupancy numbers and patterns are essential in monitoring occupants in buildings to make the building's energy system consumption more efficient and optimise building spaces. Table 1 lists the commonly used technologies, the cost of installation, and the ability of the sensor to detect, estimate, or collect data, as well as the advantages and disadvantages of the technologies.

3. Materials and methods

This section presents the case study building and discusses how the occupancy data was collected and analysed. The methods adopted in this study aim to identify occupancy patterns and profiles along with established insights of the occupants in the Higher Education Institution (HEI) library building.

3.1 The case study building

The Urban and Regional Studies (URS) building was selected as a case study building in this study. The URS is a library building situated at the University of Reading Figure 1. The building consists of seven levels, the first three of which can be used by students, while the remaining four levels are for the use of library staff. The URS building remains open 24 h from Monday to Friday. On Saturday, the building closes at 9:00 PM and reopens on Sunday at 8:30 AM.

3.2 Adopted approach for data collection

This study collected occupancy data from the building using a reliable and high-accuracy technology sensor to avoid the uncertainties encountered by previous studies in data collection and to achieve more reliability in the results obtained. The Infrared video image sensor was selected for use in a high-density occupant case study. It consists of a thermal and video image which allows it to detect occupants with an accuracy of up to 98%. The ethical procedures were followed, and risk assessments were carried out for installing the sensors and data collection. The case study of this research, the URS library building, has

Table 1. Comparison of various occupancy tracking and detection techniques with regard to cost, purpose, advantages and disadvantages.

Technology	Cost	Purpose	Advantage	Disadvantage	Consists of
Ambient Sensors	Low for single sensor and high when it is combined	Detection, estimation, and collection	Easy to install, high accuracy for combined sensors, and operation for a long time	Delay problems, combined with other sensors for high accuracy, optimal location required, and high cost	Carbon dioxide, temperature, humidity, sound, and motion detectors
Camera	High cost in most of the types	Detection and collection	High accuracy in a short period	Expensive, privacy-invasive, and short-time operation	CCTV, Video camera, and thermal camera
Transmitted Signals	High cost in most of the types	Detection, estimation, and collection	High accuracy in a single space	Problem with privacy and short-time operation	Wi-Fi and Bluetooth
Passive Infrared	From low to medium based on the type	Detection and collection	Easy to install and low maintenance cost based on the type	Limited range, not suitable for counting or high-density area	PIR and beams counters

Figure 1. The external envelope of the URS building.

four entrances, but only the main entrance is operational following the refurbishment of the building. Therefore, the measurement unit was installed on the ceiling in front of the main entrance of the building to detect the number of occupants passing through the door in both directions of entering and leaving. Figure 2 shows the location of the unit and the dimensions of the space. The unit consists of infrared and video image sensors that work simultaneously for more accurate detection. The unit can be installed on a ceiling up to 11.7 m above the floor as the sensor lens has an angle of 90°. However, in this case study, the unit was installed at 3.4 m above the floor. The measurement unit has a built-in data logger for data acquisition to capture the occupancy data that can be retrieved from an online server and exported to an Excel spreadsheet. The sensor recorded the number of occupants for a period of 12 months (one academic year 2017/2018) at five minutes intervals. The data collection covered all weekdays, including holidays and vacations, to provide more robust results in identifying occupancy patterns and profiles. The infrared feature in the sensor with a video auditing system works simultaneously for more accurate data collection and to ensure occupant privacy. An experimental evaluation of the accuracy was done using head counts of the occupancy in the building for one month, covering several hours. The comparison shows the high accuracy of the sensor with an

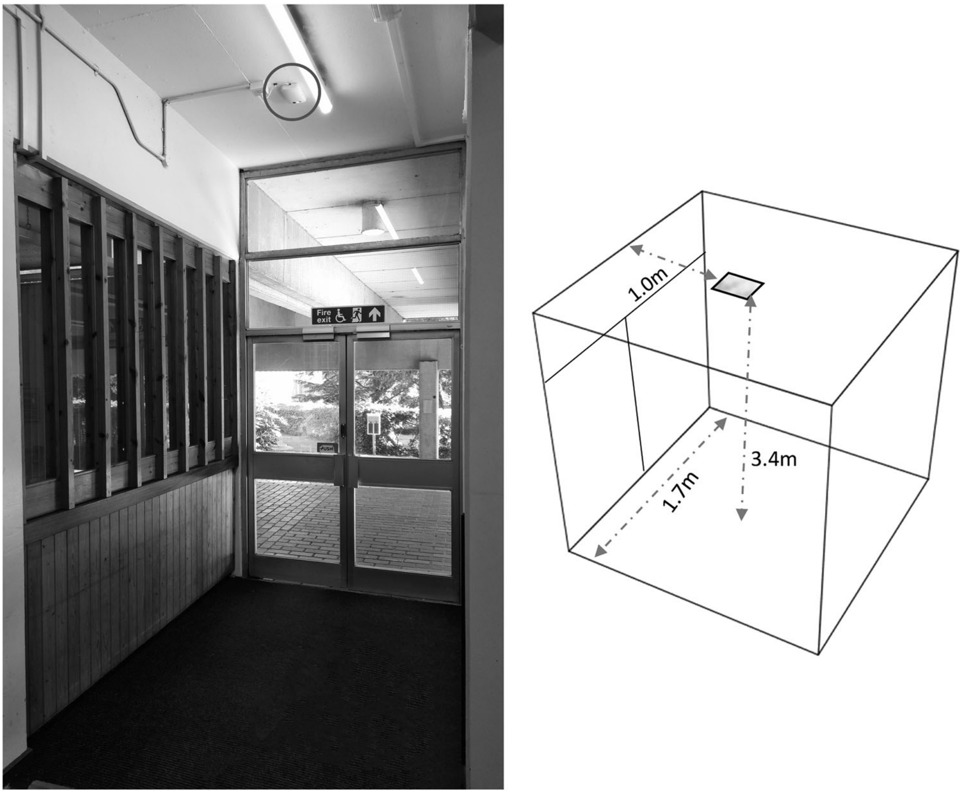

Figure 2. Sensor location (a) in front of the main entrance and (b) dimensions of the space.

error of 3.4%, and such a small error can be related to the ability of the sensor to handle a high volume of occupants.

3.3 Adopted approach for data analysis

In this study, several techniques were deployed to analyse the 12 months of occupancy data and identify its patterns and profiles in the case study building. Descriptive statistical analysis was used to interpret and gain insight into over 105,000 data points of occupancy data. This allows the specification of the occupancy number at different temporal resolutions (hour, day, week, month, and academic term). Secondly, heatmaps were used to visualise the influence of different drivers on occupants' presence. After that, cluster analysis of a K-means algorithm was deployed on the large set of occupancy data to be grouped into aggregated clusters sharing similar features (Hennig et al. 2015). The technique was applied to occupancy data of an office building and was able to effectively identify the occupancy patterns (D'Oca and Hong 2014). The number of clusters in this study was identified using the elbow method. The elbow method is useful for identifying the optimal number of clusters in a dataset, where the results are presented visually (Marutho, Handaka, and Wijaya 2018). The optimal number of clusters is determined when the distortion line graph starts to plateau (Syakur et al. 2018). A plateau occurs when the number of clusters increases, which leads to reduces the distortion value. Once the degree of distortion flattens, it indicates the optimal number of clusters. Moreover, clustering analysis is used to identify the drivers that influence the occupants in the building. Boxplot is a graphical tool that shows the distribution and spread of the data. Boxplot was used on the identified clusters to show how the days under each pattern are different from each other. Finally, Pearson Correlation Coefficient (PCC) analysis was conducted to find out the relationship between the weekdays in each pattern. The strength of the correlation between two variables is explained by the correlation coefficient (Julier and Uhlmann 2004). Figure 3 shows the framework of how the occupancy patterns and profiles were identified. The outputs of exploratory data analysis were used as an input for clustering analysis to identify the

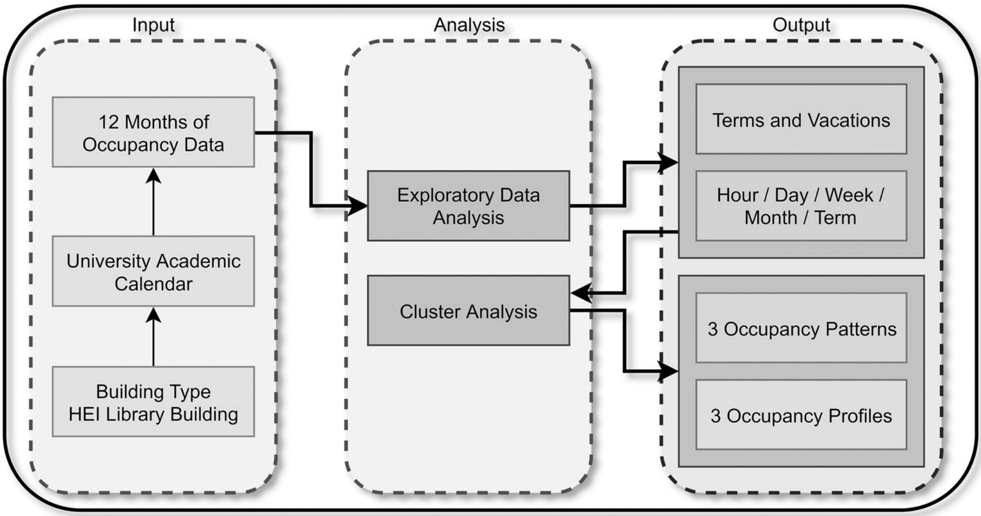

Figure 3. The framework of identifying occupancy patterns and profiles.

occupancy patterns and profiles. In this study, the MATLAB environment was used for performing data analysis and visualisation.

4. Results and discussions

The analysis of the 12 months of occupancy data identified occupancy patterns and profiles as well as discovered many drivers that influenced the occupants' presence in the building. The drivers include the beginning of the academic term, the examination period, the weekdays/ weekends, and the vacation driver. In this section, the analysis results are discussed and compared in detail using diverse techniques.

The 12 months of total occupancy numbers are plotted on a monthly scale in Figure 4. The plot shows an increasing trend in the number of occupants from January until April, with a slight decrease of 2% in May. After that, the trend drops significantly during summer vacation until September. It is noticed that the occupancy number increased by 303% in October, which can be associated with the start of the academic term after low occupancy numbers. On the other hand, the largest percentage change in the descending trend

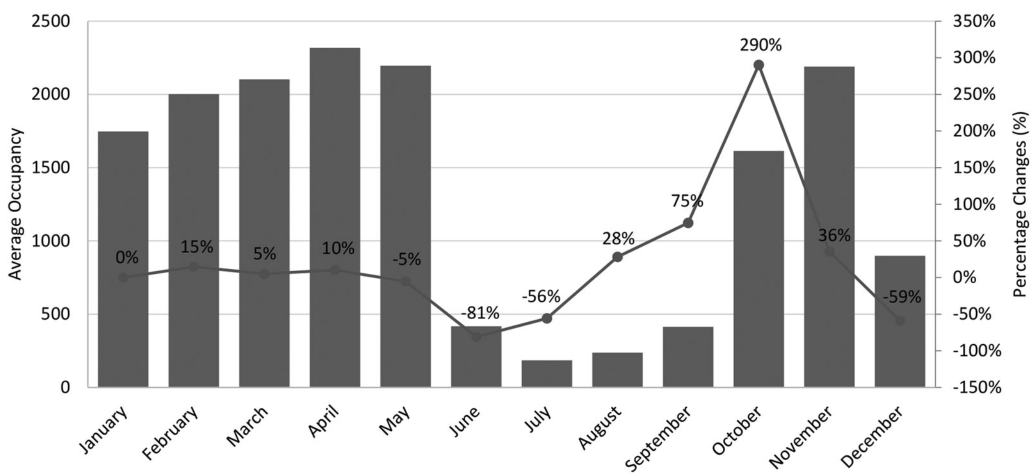

Figure 4. The average monthly occupancy number in the building includes the percentage of the changes from one month to another.

Table 2. Descriptive statistics of occupancy data covering 12 months.

	Measure of central tendency		Measure of variability			
Month	Mean	Median	SD	Min	Max	Sum
January	1,747	2,087	841	0	2,833	54,142
February	2,002	2,056	813	558	3,146	56,056
March	2,103	2,114	1,155	0	3,835	65,183
April	2,345	2,431	1,113	0	3,819	70,351
May	2,195	2,223	961	574	3,760	68,056
June	418	484	285	0	833	12,525
July	185	125	181	0	523	5,731
August	251	272	195	0	539	7,766
September	414	432	337	0	1,157	12,412
October	1,614	1,803	761	338	2,728	50,046
November	2,219	2,457	714	845	3,065	66,578
December	839	560	1,062	0	3,342	25,181

happened in June, when the summer vacation started. That shows the influence of the academic terms and vacations on the occupancy numbers. Table 2 shows the result of performing descriptive analysis, including the mean, median, standard deviation, minimum and maximum of the daily occupancy numbers for 12 months, and the total number of occupants in the building over 12 months. The highest monthly occupancy number in the building was 65,193 occupants during the examination month (April), while the lowest occupancy number was during the summer vacation in July with 5,731 occupants in the building.

Figure 5 shows the percentage of building occupied capacity on different days of the academic year. It shows academic terms and vacations, as well as the gradual changes regarding occupancy numbers across consecutive days. In the Spring and Autumn terms, the number of occupants increases with time due to increased student workloads that reach 80% of the building capacity. Unexpectedly, the highest capacity during the Spring term occurred on a Sunday, representing 79% of building capacity. On the other hand, the Summer Term was the highest period of building occupied, reaching 92% of the building capacity. The URS library building has a capacity of 830 seats, and it is worth noting that the building has never reached its 100% capacity. Even during the busiest examination period (Table 3). During the summer vacation, the number of occupants decreases sharply (the library is closed on weekends). After that, the number of occupants starts to increase again slightly before the end of the summer vacation as students are getting ready to be back at university. That proves that the examination driver has the most influence on the occupancy numbers, as the highest period occupied was during examination time.

The use of heatmaps was used to visualise how the drivers influence the occupants to present in the building. In order to illustrate the influence, four months with different occupancy densities were chosen, as shown in Figure 6. Figure 6a represents the number of occupants in January. The light-colour slots at the beginning of the month show a smaller number of occupants, which is because this period was towards the

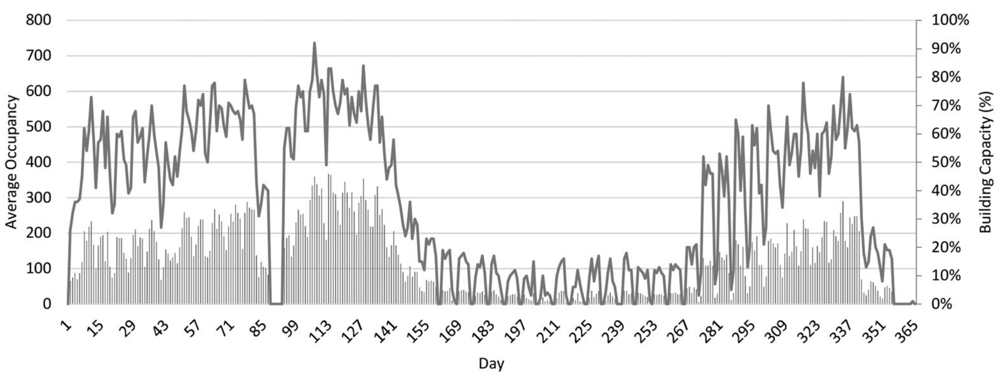

Figure 5. Average daily occupancy numbers and percentage of the building occupied capacity for one academic year on a daily basis.

Table 3. The highest capacity usage of different terms.

Term	Month	Day of the month	Day	Time	Occupancy number	Capacity usage
Spring Term	March	18	Sunday	3:30	653	79%
Summer Term	April	17	Tuesday	3:10	760	92%
Autumn Term	November	30	Thursday	3:15	664	80%

end of the Christmas vacation. The occupants were in lower numbers and with shorter stays (9:00 AM–6:00 PM). The interesting light-colour gradient in Figure 6b, which is in the middle of February, is related to the reading week where there are no classes timetabled. Figure 6 presents the occupancy heatmap in four months, January, February, April, and June. In April, Figure 6c, there are darker slots that cover a larger area of the heatmap. The occupants during this month stayed in the building for longer periods because of the examination period. Another observation in Figure 6d is that the colour of the heatmaps becomes sharply lighter in June. This coincides with the start of the summer vacation. During the summer vacation, the heatmap shows a lighter colour gradient, which is because the building closes during weekends and has limited access during workdays (5:00 PM). In summary, the identified drivers are associated with different months and academic terms.

4.1 Clustering analysis

A set of clustering analyses was carried out to identify occupancy patterns and profiles from the collected data. Figure 7 shows the average occupancy number in the building for 12 months in a one-day cycle. Several months presented show a similar pattern. The elbow method was deployed first to find the optimal number of clusters (Marutho, Handaka, and Wijaya 2018). The elbow method results are presented in Figure 8, showing that grouping the 12 months data into three clusters is the ideal grouping number. The total occupancy numbers for each month were clustered based on the number of occupants into three typical clusters graphically shown in Figure 9. Each cluster contains months that are similar in occupancy rate. Figure 10 presents the distribution of the average occupancy number of the patterns over a one-day cycle. There is a variation in the building capacity usage of 11% between Pattern 1 and Pattern 2, where the variation is more significant between Patten 2 and Pattern 3, up to 45%. The identified patterns based on the occupancy rates can be associated with different influence drivers. Pattern 1 has a medium occupancy rate associated with the beginning of the academic term driver, Pattern 2 has a high occupancy rate associated with the examination period driver, and Pattern 3 has a low occupancy rate associated with the vacation driver. The three daily patterns in more detail are as follows:

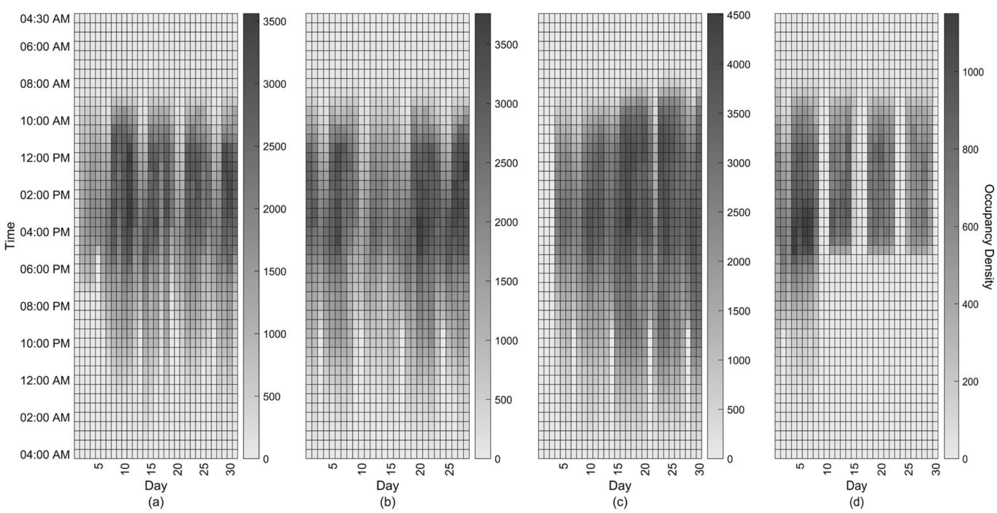

Figure 6. Heatmaps of the months; (a) January, (b) February, (c) April, and (d) June show the occupancy patterns across one day cycle.

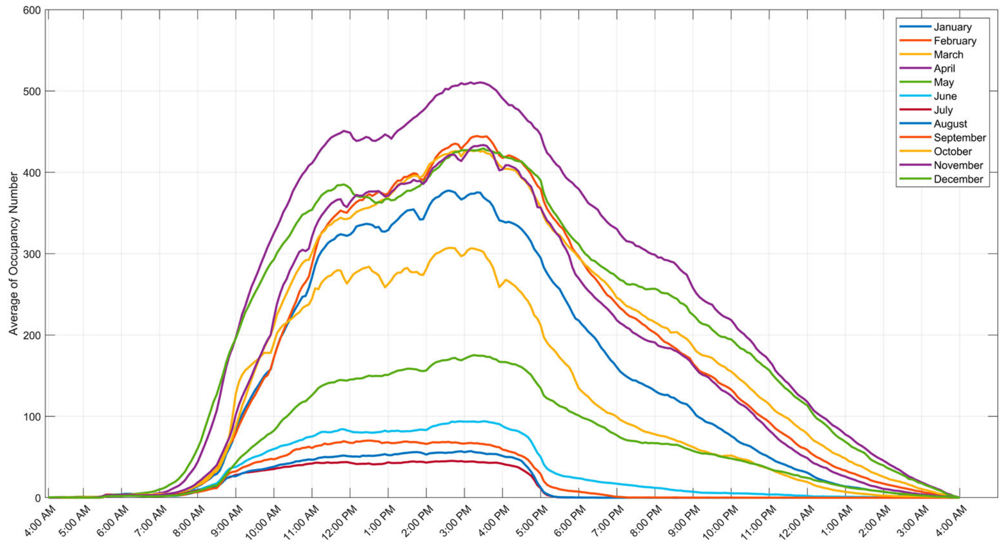

Figure 7. The average occupancy number in the HEI library building of 12 months distributed over one day cycle.

- Pattern 1 presents Monday to Sunday with a medium occupancy rate.
- Pattern 2 presents Monday to Sunday with the highest occupancy rate.
- Pattern 3 presents Monday to Sunday with the lowest occupancy rate.

As mentioned at the beginning, the body of the literature focuses on office buildings. The reason behind that is that office buildings have a benchmark for the occupancy diversity factor during weekdays that has been published in ASHRAE Standard 90.1 (2004). The availability of the benchmark gives the advantage of comparing and verifying the results of studies with the current standard. However, the case study building

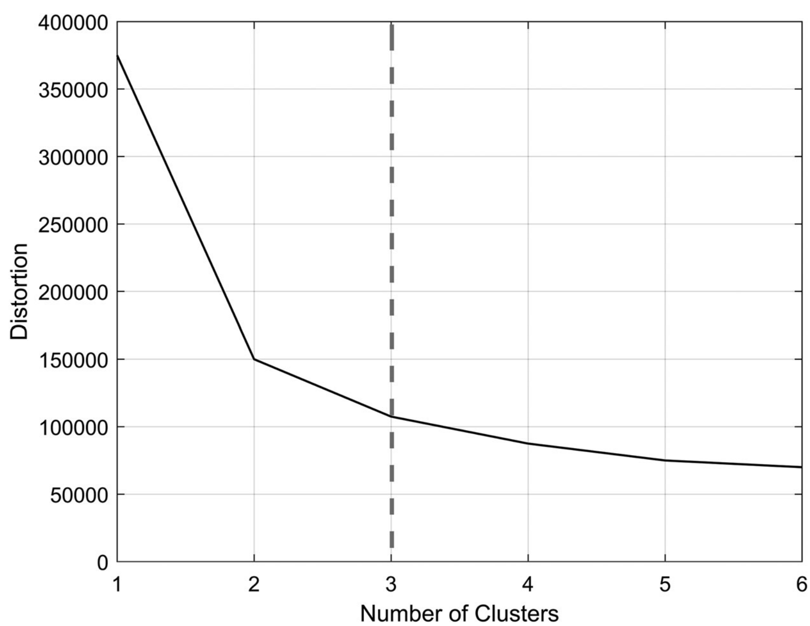

Figure 8. Elbow method to determine the optimal number of clusters.

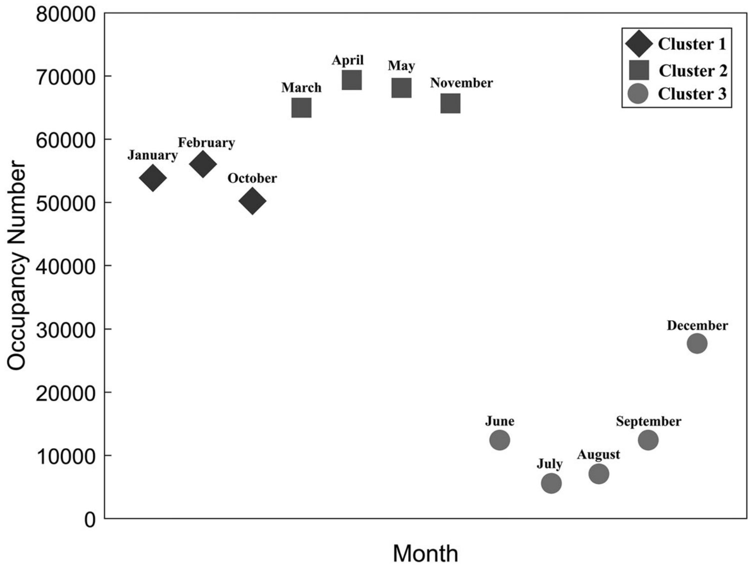

Figure 9. Graphical representation of data clusters and the months of each cluster.

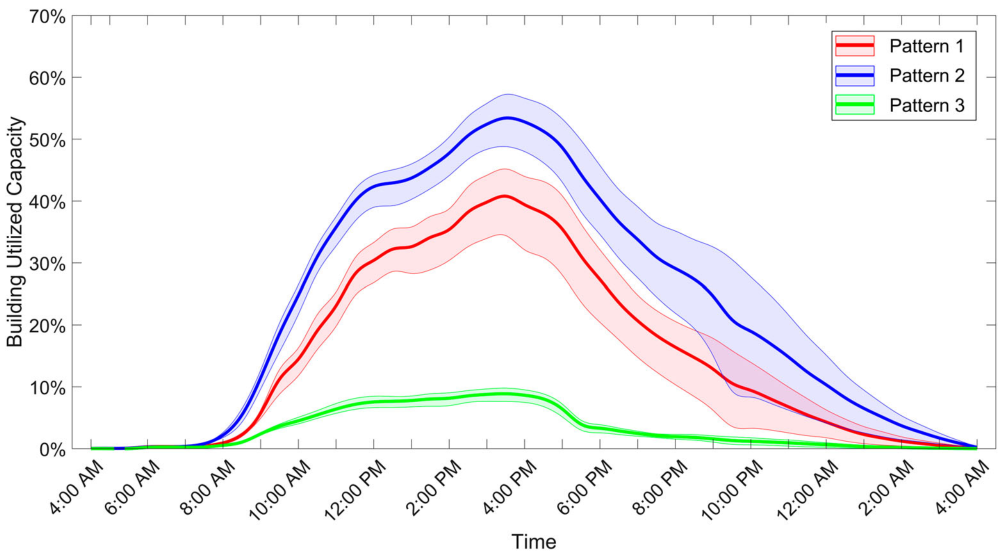

Figure 10. Presenting the minimum, average, and maximum building utilises capacity for the three patterns. The red colour is Pattern 1, the blue colour is Pattern 2, and the green colour is Pattern 3.

for higher education in this paper does not have an available benchmark for occupant density and distribution. This study compared the benchmark available for the office buildings with the patterns resulting from analysing 12 months of occupancy data in a higher education building to draw attention to different occupancy patterns associated with different types of buildings to avoid generalising these results. Figure 11 illustrates the occupancy percentage in buildings over one day cycle of weekdays. Figure 11 shows the ASHRAE benchmark during the workday has a symmetric pattern from the morning until the break hour at midnight and then to the end of the office hours. Comparing the ASHRAE trend with Pattern 2 to illustrate the difference in occupancy trend of two different building types. In Pattern 2, the occupancy numbers increased gradually, unlike in the ASHRAE trend, all the workers presented and left at specific hours. In contrast, the ASHRAE trend shows fewer occupants during the weekends, which is understandable because

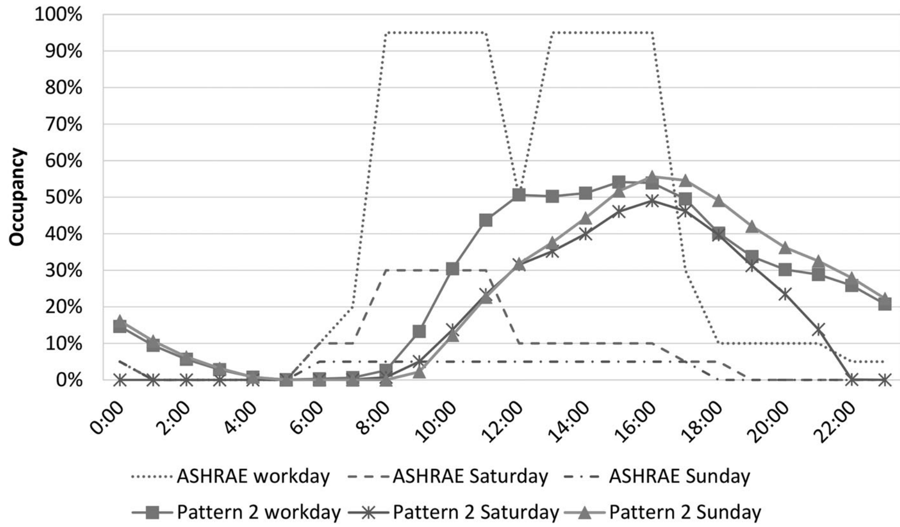

Figure 11. Comparing the ASHRAE benchmark with Pattern 2, which has the highest occupancy rate in the building.

of the building type and working hours. However, the occupancy trend for Pattern 2 during weekends was not significantly different from workdays with a large number of occupants in the building, which is related to the building type.

The comparisons illustrate the differences between the typical occupancy rate during weekdays of two different types of buildings. Office buildings have a stable rate of occupants during office hours and weekends for the majority of the year. Also, the number of employees in the building is known, as they have fixed working hours and a specified time to take a break. On the other hand, the number of occupants in higher education buildings is unknown. Many drivers influence the number of occupants during different periods of the year. Therefore, the availability of a benchmark for one type of building is not necessary that can be applied to other types of buildings, as explained and illustrated by the differences above. Thus, it demonstrates the importance of the patterns and profiles identified in this paper for a high-density higher education building, giving the possibility to inform the development of a benchmark for occupancy profiles for libraries in higher education buildings.

Figure 12 shows the average occupancy number of the three identified patterns covering workdays and weekends in the building distributed over a one-day cycle. Pattern 1 in Figure 12a and b show the occupancy disruption during workdays and weekends, respectively. Pattern 1 is associated with the beginning of the academic term driver. The occupant arrival time begins around 7:30 AM, and the number of occupants increases until the peak level. The highest peak level recorded during workdays was on Monday, with 54% of the building capacity, while it was 38% on weekends (Sunday). Friday had the lowest utilised capacity by 13% compared to Monday during the peak hour. Also, the starting slope rate after 9:00 AM, decreased by 26% on Friday compared to the other days. The high occupancy numbers of the staying duration on weekdays are between 11:00 AM and 4:00 PM, after which the occupants start to leave. Pattern 1 shows that some occupants stay in the building as late as 4:00 AM on workdays. Pattern 2 is presented in Figure 12c and d, showing an increase in the number of occupants associated with the examination period driver. Figure 12c shows Mondays recorded the highest number of occupants, which is 59% of building capacity, whereas Friday had the lowest workdays building capacity of 50%. The weekend in Figure 12d utilised 49% and 56% of the building capacity on Saturdays and Sundays, respectively. However, the library occupied these days is high, since students study for exams. The occupants begin to use the library as early as 6:30 AM on workdays and 7:30 AM on weekends, and the numbers of occupants continue to increase during the day to the peak hour. The early departure of the occupants on Saturdays had a similar reason as in Pattern 1, which was the library opening time. Patten 3 in Figure 12e and f show a significant fall in the occupancy number in the building. This is expected due to the end of exams and the start of the summer vacation, as this pattern is associated with the vacation driver. The number of occupants in the building during the weekdays shows that use of the library falls to zero at around 12:00 AM. The average peak level was 12% of building

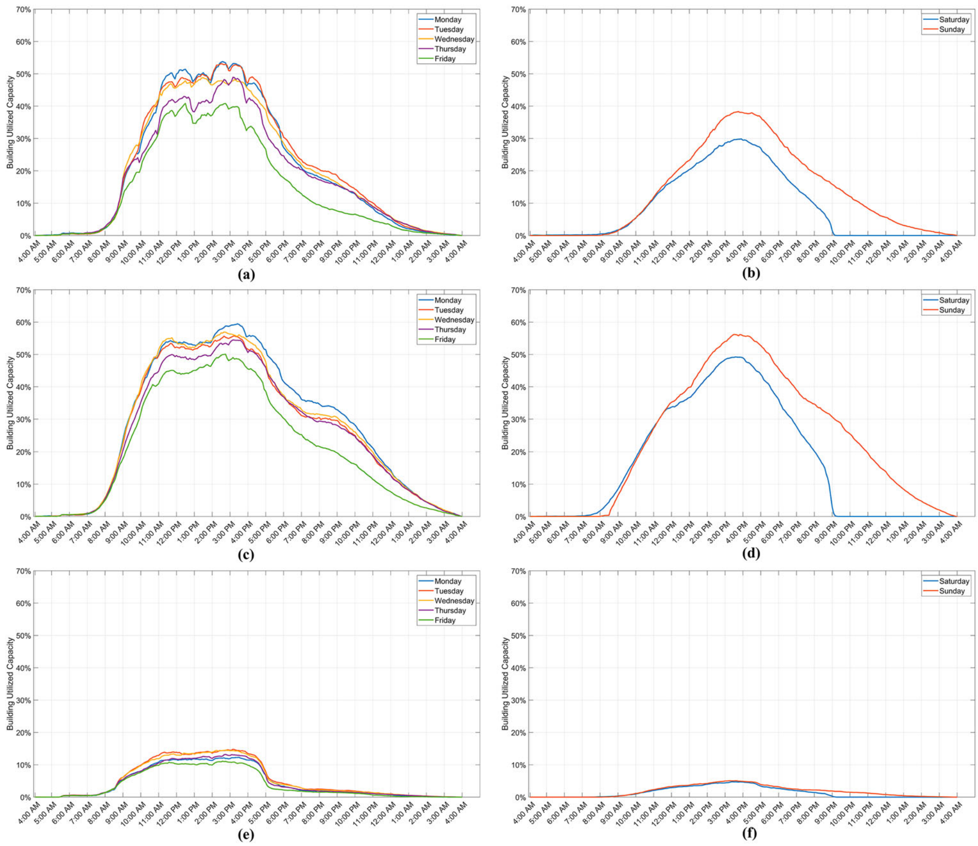

Figure 12. Distributing the average occupancy number over a one-day cycle Pattern 1 in plots (a & b), Pattern 2 in plots (c & d), and Pattern 3 in plots (e & f) for workdays and weekends.

capacity usage. During the weekends, as shown in Figure 12f, the occupancy number was much lower due to the building being closed on some days with limited building access time.

Boxplot was utilised to display the distribution of the data, the central value, and its variability. Boxplot used the three patterns identified to show how the maximum occupancy numbers during weekdays of each pattern. Pattern 1 in Figure 13a shows the highest mean of utilised building capacity started at the beginning of the week and then continued decreasing to the weekend, which can be associated with the beginning of the term driver. The busiest days were Monday and Tuesday, with a median utilised building capacity of around 60%. The median of utilised building capacity continued to fall, reaching around 47% on Friday. Tuesday has the highest variance, with maximum occupancy rates of 77.22%. The highest decrease can be observed on weekends when the median is approximately 35% on Saturday, and then on Sunday, the median increases slightly. Pattern 2 in Figure 13b is similar to Pattern 1, where the median starts high and decreases until the weekend, except on Tuesday, which has a high degree of variance of 91.56%. The outliers presented can be explained by the fact that the library building was closed during the Easter vacation. The highest median is registered on Monday, with 69.39% of the building utilised capacity. The rest of the days have an average median of 60%, except for Saturday, with the median being 52%. Pattern 3 in Figure 13c has a similar median utilised building capacity for workdays of an average of 13% and almost non-existent occupants on weekends. The low utilised building capacity is associated with the vacation driver. The outliers presented in a high range of utilised building capacity can be explained by the fact that there is a period in which the building is highly used (December).

After analysing the three patterns presented in Figures 12 and 13, taking into account several indicators such as the starting time, leaving time, fill-up slope, peak hour, and building utilisation capacity to facilitate further analyses. From that point, the occupancy profile identified and determined the typical occupancy schedule during the weekdays. The occupancy profiles in Figure 14a–c focus on the period of time when

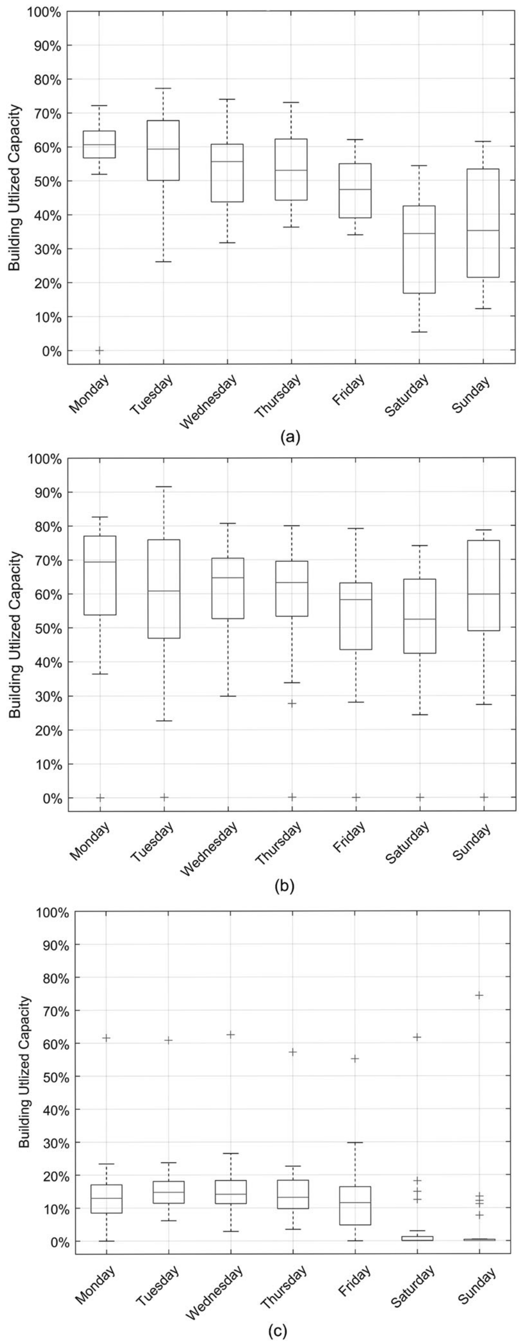

Figure 13. Boxplot of the three patterns showing the building sutilised capacity during weekdays, Pattern 1 (a), Pattern 2 (b) and Pattern 3 (c).

the occupants start to come to the building, the peak hour, the staying time, and leaving time, which is distributed the average occupancy number of each pattern over a one-day cycle. There was one hour difference between the occupancy profile of Pattern 1 and the other patterns with respect to the time that occupants began arriving. The filling-up rate can give an idea of how fast the building can fill up to the maximum, between 8:00 AM and 11:00 AM. Pattern 2 has the highest filling-up rate (Figure 14b), followed by Pattern 1 (Figure 14a) and Pattern 3 (Figure 14c). The filling rate in Pattern 1 is related to the beginning of the term

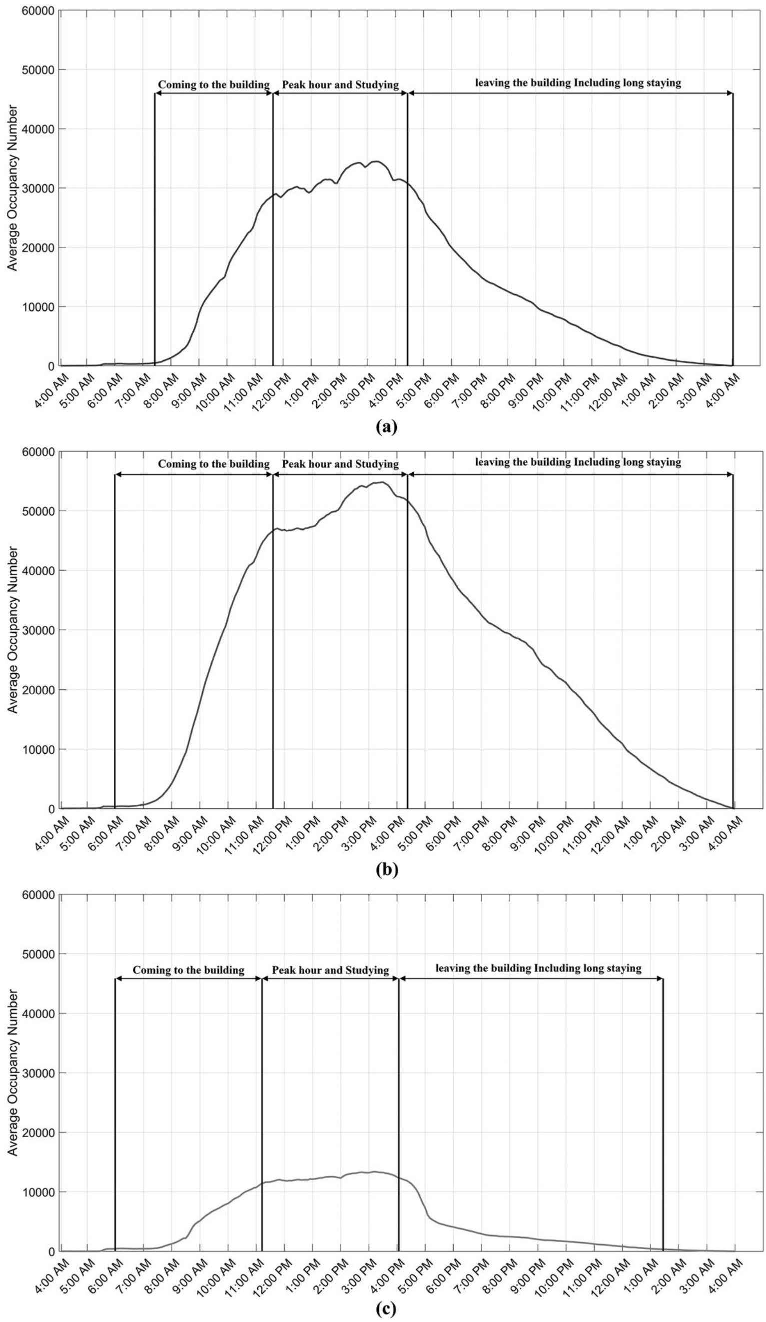

Figure 14. Occupancy profiles of the three patterns in the HEI library building during weekdays, plot (a) represents pattern 1, plot (b) represents pattern 2, and plot (c) represents pattern 3.

driver, and the explanation of the high filling rate in Pattern 2 is related to the examination period driver. On the other hand, the low rate of filling up in Pattern 3 is related to the vacation driver. The associated drivers also have an influence on the peak hour, on leaving the building, and on the number of occupants in the building.

PCC analysis was used to find out the relationship between the patterns by measuring the amount of linear association between them. Figure 15 shows the results of the PCC analysis to determine how strong the correlation is between the identified patterns. The analysis was performed on the hourly occupancy number for a

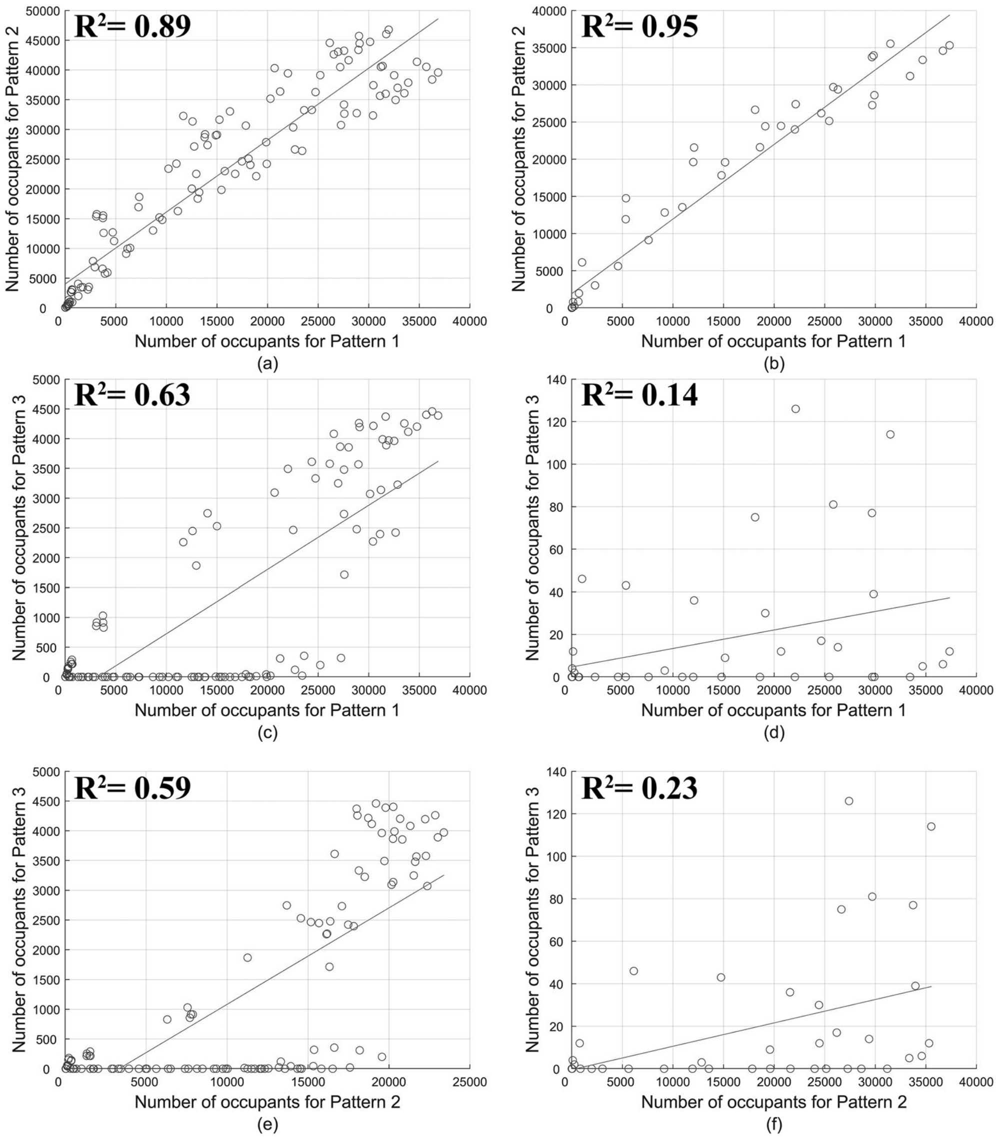

Figure 15. The result of performing the analysis of PCC between Pattern 1 and Pattern 2 during (a) workdays and (b) weekends, Pattern 1 and Pattern 3 during (c) workdays and (d) weekends, and Pattern 2 and Pattern 3 during (e) workdays and (f) weekends.

one-day cycle using one month from each of the three patterns during workdays and weekends. Figure 15a shows a strong correlation of 0.89 between the workdays of Pattern 1 and Pattern 2, with R^2 of 0.95 during weekends, as shown in Figure 15b. The second correlation presented in Figure 15c between Pattern 1 and Pattern 3 shows a moderate correlation of 0.63 during workdays and no correlation with the weekend Figure 15d. The low value can be explained by the fact that the building was open during the weekends for Pattern 1, unlike Pattern 3. The third correlation between Pattern 2 and Pattern 3 in Figure 15e shows a moderate of 0.59 during workdays and no correlation during weekends Figure 15f.

5. Conclusions

This study identified the occupancy patterns, profiles, and their associated main drivers in a high-density higher education building in England for an extended period of 12 months, covering all weekdays. The occupancy data were collected using an infrared video image analysis sensor to eliminate the uncertainties in data collection and, consequently, the results. A series of clustering analyses were carried out, identifying three patterns and profiles in the case study building. The key findings can be summarised as follows: Four drivers that significantly influence the identified occupancy patterns were identified. The first driver is the beginning of the academic term, which shows an influence on Pattern 1. The second driver is the examination period, which is the most influential factor in Pattern 2. The third driver is the weekend, which its effect never disappears, even during the examination period. The fourth driver is vacation, which has a meaningful influence on Pattern 3. There was a variation in the occupancy rates of almost 11% between Pattern 1 and Pattern 2. The variation increased to almost 45% between Pattern 2 and Pattern 3, as shown in Figure 10. Furthermore, Figure 14 shows the results of identified occupancy profiles in typical one-day cycle activities, including entering the building, peak hours, staying, and leaving the building.

The results of this study can contribute to knowledge by addressing the gaps found in the previous studies. First, the data collected in this research over 12 months addressed the existing shortage of long-term occupancy data related to higher education buildings, compared to 30 days of data collection in Wang and Shao (2017) in a selected room in an education building or 45 days in Mora et al. (2019) study. Second, using high-resolution infrared video image sensors eliminates the uncertainties associated with the robustness of collected occupancy data, especially for high-density buildings, compared to the issues faced by Chang and Hong (2013) using lighting switch sensors, Yuan et al. (2020), using PIR to collect occupancy status, or Tekler et al. (2020) using Bluetooth signals to track the occupants. Third, the approach taken in collecting and analysing real-world data can contribute to the development of a new benchmark for the HEI buildings to be used in future studies. As the comparison also illustrates the differences between the patterns identified in this study and ASHRAE Standard 90.1 (2004) during weekdays, where a unique standard is required for a higher education institution based on its occupancy rates and patterns. This imposes the importance of the results obtained in this study for higher education buildings.

Disclosure statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Notes on contributors

Bashar Alfalah is a Lecturer, in the College of Architectural and Planning, at Imam Abdulrahman Bin Faisal University, Saudi Arabia, He is also completing his doctoral thesis at the School of the Built Environment, University of Reading.

Mehdi Shahrestani is a Lecturer in sustainable technologies.

Li Shao is a professor in School of the Built Environment, at the University of Reading, Reading, UK.

ORCID

Mehdi Shahrestani http://orcid.org/0000-0002-8741-0912
Li Shao http://orcid.org/0000-0002-1544-7548

References

Amber, K. P., M. W. Aslam, A. Mahmood, A. Kousar, M. Y. Younis, B. Akbar, G. Q. Chaudhary, and S. K. Hussain. 2017. "Energy Consumption Forecasting for University Sector Buildings." *Energies* 10 (10): 1579.

Andersen, P. D., A. Iversen, H. Madsen, and C. Rode. 2014. "Dynamic Modeling of Presence of Occupants Using Inhomogeneous Markov Chains." *Energy and Buildings* 69: 213–223.

Ashouri, A., G. R. Newsham, Z. Shi, and H. B. Gunay. 2019. "Day-ahead Prediction of Building Occupancy Using WiFi Signals." In *2019 IEEE 15th International Conference on Automation Science and Engineering (CASE)*, 1237–1242. IEEE.

ASHRAE, ASHRAE Standard. 2004. *Standard 90.1-2004, Energy Standard for Buildings Except low Rise Residential Buildings*. American Society of Heating, Refrigerating and Air-Conditioning Engineers.

Azam, M., M. Blayo, J.-S. Venne, and M. Allegue-Martinez. 2019. "Occupancy Estimation Using Wifi Motion Detection via Supervised Machine Learning Algorithms." In *2019 ieee global conference on signal and information processing (GlobalSIP)*, 1–5. IEEE.

Bernardo, H., and F. Oliveira. 2018. "Estimation of Energy Savings Potential in Higher Education Buildings Supported by Energy Performance Benchmarking: A Case Study." *Environments* 5 (8): 85.

Chang, W., and T. Hong. 2013. "Statistical Analysis and Modeling of Occupancy Patterns in Open-Plan Offices Using Measured Lighting-Switch Data." In *Building Simulation*, 23–32. Springer.
Chatterton, T. 2018. Sustainable Developments? Assessing the Direct Energy (Building and Vehicles) Consumption of New Housing Developments in the United Kingdom Using Social and Structural Benchmarking.
Department for business energy and industrial strategy, Energy Consumption in the UK (ECUK) [Internet]. 2020. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/928350/2020_Energy_Consumption_in_the_UK__ECUK_.pdf.
D'Oca, S., and T. Hong. 2014. "A Data-Mining Approach to Discover Patterns of Window Opening and Closing Behavior in Offices." *Building and Environment* 82: 726–739.
Everitt, B. S., S. Landau, M. Leese, and D. Stahl. 2011. "An Introduction to Classification and Clustering." *Cluster Analysis* 5: 1–13.
Hennig, C., M. Meila, F. Murtagh, and R. Rocci. 2015. *Handbook of Cluster Analysis*. CRC Press.
Hobson, B. W., H. B. Gunay, A. Ashouri, and G. R. Newsham. 2020. "Clustering and Motif Identification for Occupancy-Centric Control of an air Handling Unit." *Energy and Buildings* 223: 110179.
Ioannou, A., L. Itard, and T. Agarwal. 2018. "In-situ Real Time Measurements of Thermal Comfort and Comparison with the Adaptive Comfort Theory in Dutch Residential Dwellings." *Energy and Buildings* 170: 229–241.
Jallow, B., and W.-T. Hong. 2018. "Evaluation of Indoor Positioning Based on IBeacon and Pi-Beacon." In 2018 International Conference on System Science and Engineering (ICSSE), 1–4. IEEE.
Jang, H., and J. Kang. 2016. "A Stochastic Model of Integrating Occupant Behaviour Into Energy Simulation with Respect to Actual Energy Consumption in High-Rise Apartment Buildings." *Energy and Buildings* 121: 205–216.
Jia, M., R. S. Srinivasan, R. Ries, and G. Bharathy. 2018. "Exploring the Validity of Occupant Behavior Model for Improving Office Building Energy Simulation." In *2018 Winter Simulation Conference (WSC)*, 3953–3964. IEEE.
Jin, Y., D. Yan, X. Zhang, J. An, and M. Han. 2020. "A Data-Driven Model Predictive Control for Lighting System Based on Historical Occupancy in an Office Building: Methodology Development." Building simulation.
Julier, S. J., and J. K. Uhlmann. 2004. "Unscented Filtering and Nonlinear Estimation." *Proceedings of the IEEE* 92 (3): 401–422.
Kim, Y. K., L. Bande, K. A. Tabet Aoul, and H. Altan. 2021. "Dynamic Energy Performance gap Analysis of a University Building: Case Studies at UAE University Campus, UAE." *Sustainability* 13 (1): 120.
Liang, X., T. Hong, and G. Q. Shen. 2016. "Occupancy Data Analytics and Prediction: A Case Study." *Building and Environment* 102: 179–192.
Liu, X., H. Makino, and K. Mase. 2010. "Improved Indoor Location Estimation Using Fluorescent Light Communication System with a Nine-Channel Receiver." *IEICE Transactions on Communications* E93-B (11): 2936–2944.
Marutho, D., S. H. Handaka, and E. Wijaya. 2018. "The Determination of Cluster Number at k-Mean Using Elbow Method and Purity Evaluation on Headline News." In *2018 international seminar on application for technology of information and communication*, 533–538. IEEE.
Mora, D., G. Fajilla, M. C. Austin, and M. De Simone. 2019. "Occupancy Patterns Obtained by Heuristic Approaches: Cluster Analysis and Logical Flowcharts. A Case Study in a University Office." *Energy and Buildings* 186: 147–168.
Rashid, M. T., M. S. S. Chowdhury, and M. F. Nawal. 2016. "Composite Indoor Localization and Positioning System Based on Wi-Fi Repeater Technology." In *2016 IEEE Region 10 Conference (TENCON)*, 2130–2133. IEEE.
Razavi, R., A. Gharipour, M. Fleury, and I. J. Akpan. 2019. "Occupancy Detection of Residential Buildings Using Smart Meter Data: A Large-Scale Study." *Energy and Buildings* 183: 195–208.
Syakur, M. A., B. K. Khotimah, E. M. S. Rochman, and B. D. Satoto. 2018. "Integration k-Means Clustering Method and Elbow Method for Identification of the Best Customer Profile Cluster." In *Iop conference series: materials science and engineering*, 12017. IOP Publishing.
Tekler, Z. D., R. Low, B. Gunay, R. K. Andersen, and L. Blessing. 2020. "A Scalable Bluetooth Low Energy Approach to Identify Occupancy Patterns and Profiles in Office Spaces." *Building and Environment* 171: 106681.
van den Brom, P., A. Meijer, and H. Visscher. 2018. "Performance Gaps in Energy Consumption: Household Groups and Building Characteristics." *Building Research & Information* 46 (1): 54–70.
Wang, Y., and L. Shao. 2017. "Understanding Occupancy Pattern and Improving Building Energy Efficiency Through Wi-Fi Based Indoor Positioning." *Building and Environment* 114: 106–117.
Yuan, et al . 2020. "A data-driven model predictive control for lighting system based on historical occupancy in an office building: Methodology development." *Building Simulation* 14 (1): 219–235.
Zavala, V. M. 2014. "Inference of Building Occupancy Signals Using Moving Horizon Estimation and Fourier Regularization." *Journal of Process Control* 24 (6): 714–722.
Zhang, B., Y. Shen, and L. Zhang. 2016. "Inferring Building Occupancy Based on Statistical Modeling of Multi-Sensor Data." In *2016 IEEE international conference on automation science and engineering (CASE)*, 809–814. IEEE.