Considering Culture-Specific Occupant Behaviour in Energy Performance Evaluation: A Case Study of Residential Houses in Riyadh, Saudi Arabia.

Mohammed Aljammaz

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield

04/11/2021

Abstract

Energy consumption prediction is a critical aspect of home design that can help homeowners manage energy use. At present, the performance gap between energy consumption prediction and actual energy use could reach two to five times higher than current predicted performance rates (Zou, Wagle & Alam 2019). Lack of information about the specific factors that control energy use is a main contributor to energy performance gaps (Yoshino, Hong & Nord 2017). However, research has shown that human behaviour significantly impacts energy performance gaps, even more than building design.

Recent technological advances have led to the development of energy simulation software that helps predict energy use. The software includes various inputs to support energy consumption predictions. This study aims to address the lack of information related to Saudi cultural occupant behaviour in an effort to enhance load schedule inputs into simulation software to improve energy use predictions. This research focuses on single-family houses in Riyadh, Saudi Arabia.

For this study, a mixed-methods approach was used. A combination of both quantitative and qualitative methods allows for an in-depth exploration and understanding of cultural behaviour. In total, ten case studies were used and 1600 survey responses were collected from participants using face-to-face surveys and online forms. The researcher found that privacy is one of the main factors that shape Saudi culture. This is reflected in the behaviour of occupants, house design, and circulation. To prevent circulation overlap between guests and family members, houses are divided into three main zones: a male guest zone, a female guest zone, and a family zone. This research addresses the gap in the literature regarding unique behaviour for targeted zones. The experimental study showed that guest zones covered 30% to 40% of the total home area. Each of these zones had unique time-use behaviour. Results of data analysis found that the average time-

use of the guest zones was 36 days per year. The time-use data (TUD) survey results were used to develop an occupancy prediction tool using Random Forest. The tool predicted the time-use for the male guest zone, the female guest zone, and the living halls across five case studies. The five case studies were used to evaluate the occupancy prediction tool and test its applications in energy simulation. Without considering the cultural behaviour, experiment results show the gap between energy prediction output and real energy use could reach as high as 52%. After applying the prediction tool, the gap was reduced to a difference of 4%.

This study focuses on the single-family who lives in a detached house in the capital of Saudi Arabia, Riyadh. It considers the cultural behaviour related to occupants' use of guest zones in their homes. The research did not consider other behaviour that may be driven by Saudi culture, such as the size or the time-use of windows, which may also impact energy consumption. There is potential for future researchers to investigate the impact of culture on other features in the house, to collect more data considering the other building and family types in Riyadh or other cities in the country, and to validate the model and the impact of the cultural behaviour on the energy consumption using more case studies.

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1. Predicting energy consumption in buildings

Five main inputs are required to predict energy consumption for any building. The inputs are climate, geometry, material, load schedules, and systems. The accuracy of predicted energy consumption rates depends on whether these five inputs are representative of actual energy use. At the inception of energy prediction methods in the 1960s, the gap was estimated to be 2.5 times the real energy use (de Wilde 2014) and considered an acceptable magnitude of difference (Clarke 2001). However, by the mid-1990s, this gap became unacceptable due to concerns about increasing energy prices. As a result, researchers, scientists, and consumers began to focus on reducing energy use.

In response to these concerns, governments around the world set energy codes for commercial and residential buildings. In addition, they began educating people about the importance of reducing energy use to thwart the negative impact of global warming. Subsequently, homebuilders and homeowners demanded more accurate energy use predictions to ensure buildings met code requirements at the design stage of building prior to construction. As part of this demand, software developers began designing sophisticated simulation methods to reduce energy performance gaps (Sunikka-Blank & Galvin 2012). Simulation software was developed and helped reduce energy performance gaps to approximately 30%. Through these developments, software outcome gaps were identified in other indices, such as thermal comfort, indoor air quality, acoustic performance, and daylight levels. The improved accuracy of data inputs in the software yielded more accurate energy prediction outcomes (Oberkampf & Roy 2010).

Occupant behaviour load schedules are one of the main factors that need to be considered to close the performance gap (Martinaitis et al. 2015; Sunikka-Blank & Galvin 2012). Load schedules are the parameters representing the end-use energy activities of occupants and pose a significant challenge for modelling due to the unpredictable and dynamic nature of human behaviour (Paatero & Lund 2006). Inaccurate assessment and a lack of sufficient knowledge of people's behaviour and activities are critical factors that have led to previous inaccuracies in simulated estimations (Yoshino, Hong & Nord 2017).

Considering people’s behaviour can provide a better understanding of building energy use and help define the necessary calculation inputs for energy simulations. Several initiatives have been undertaken to understand behaviour pertaining to energy usage. Currently, technology does not afford researchers exact energy use predictions in the real world. At present, standardised assessments assume a very high energy usage that does not reflect realistic assessments of occupant behaviour (Yu et al. 2011). However, researchers are working on reducing the gap to a more acceptable level. Higher accuracy is achievable when exact input data are available, especially information on building occupancy patterns and activities (Korjenic & Bednar 2012).

Several researchers have discovered that default load schedules do not represent real occupant behaviour in their countries because they do not appropriately account for cultural differences in behaviour (Ali et al. 2020; Cruz 2019; van Dronkelaar et al. 2016). In response, they have conducted time-use data (TUD) surveys to obtain more accurate behaviour data to generate new occupancy schedules (Capasso, Lamedica & Prudenzi 1994). Widén et al. (2009a) developed a computational framework to generate a comprehensive thermal model considering realistic energy behavioural patterns. The aim was to prove that the TUD had a high potential for load modelling, which could introduce realistic behavioural patterns into various simulations. In addition, Statistics Sweden (SCB) conducted a TUD survey as part of a pilot study in 1996 to assess load schedules for occupant behaviour in Sweden. Further, Wilk (2012) presented a bottom-up stochastic method to simulate domestic occupancy behaviour in France. Further, he generated occupancy-driven load schedules from French time-use survey data during the period 1998–1999. Since then, TUD surveys have been adopted widely by researchers from the United Kingdom, France and Sweden to incorporate occupants' behaviour as inputs into energy performance simulations to improve energy consumption predictions. Researchers have also used new technology, such as machine learning (ML) algorithms to train and validate the data and predict more accurate occupant behaviour input to improve accuracy of energy predictions. They have determined that ML is an efficient mechanism to predict occupant behaviour input for energy simulations.

1.2. Background of the research

Prior to 2016, Saudi Arabians were not concerned about energy consumption because the government subsidised electricity use at an affordable fixed price (0.1 Saudi Riyal for 1 KWH). However, on 25 April 2016, Crown Prince Mohammed ibn Salman Al Saud and the Saudi Arabian government announced that, by 2030, the country's income would no longer be based on oil revenues. This change was important to save the Saudi economy and prevent it from future crises. However, the shift away from oil revenues subsequently affected the cost of electricity.

Compared to energy prices, energy use in recent years demonstrates the impact of the Saudi government intervention. Buildings in Saudi Arabia consume approximately 80% of the total electricity generated in the country, with approximately 50% of usage attributed to residential buildings (Ahmed Felimban et al. 2019). Given this reality, the government announced its intention to discontinue support for electricity use. Beginning in 2017, the government started gradually removing subsidies with the aim of completely removing support by 2020. After 2020, electricity bills would be based on fuel export prices. In response, Saudi people have begun to consider energy use and reducing their overall energy bills as an important element when designing their homes. Energy simulations to predict and understand energy consumption is a key factor in building design.

Occupant behaviour is one of the main factors that impact energy prediction outcomes. Occupant behaviour includes dynamics of how individuals operate and use building space. Saudi Arabian culture has a direct influence on Saudi lifestyle and behaviour. This is reflected in building preferences and building use. Several studies have examined characteristics of Saudi culture, such as how city and building designs reflect the Islamic religion and ideas of privacy. However, research is limited in fully understanding the impact of cultural behaviour on energy use predictions. Further, research has not examined the impact of cultural and behavioural considerations as energy simulation inputs to improve the accuracy of predictions. This study investigates occupant choices and usage behaviour to address this gap in the literature. Findings from this study can improve the accuracy of energy simulation inputs that will help people have more accurate energy predictions for their homes.

Load schedules are one of the five main inputs for energy modelling and are driven by occupancy behaviour. Current load schedules used in simulations do not account for culture-driven occupancy behaviour and far exceed actual estimations. Case studies have shown an expected 40% excess of predicted energy use over actual use in residential buildings in Riyadh, Saudi Arabia (Aljammaz 2016).

As noted earlier, occupant behaviour influences energy consumption, which is a critical factor in the gap between predicted and actual energy use (Menezes et al. 2012). One of the unique patterns of behaviour influenced by the culture in Saudi Arabia is that every villa must have two separated gender-specific guest zones (a male guest zone and a female guest zone). These guest zones have an entrance, reception room, dining room, and bathroom. Smaller houses have just one reception used for male and female visitors; however, the space is used for each gender at different times. In these houses, the reception room is usually only used for guests a few days a year.

1.2.1. The growth of Riyadh

This research focuses on single-family homes in Riyadh, the capital of Saudi Arabia (located at a Latitude of 24.000 North and a Longitude of 45.000 East). Riyadh has a desert climate. The weather is typically hot and dry in summer, with average temperatures ranging from 81 – 109 °F (27.2 – 42.7 °C) (Fig. 1.1). Due to these dry conditions, there are frequent dust storms. The dust is often so thick that the visibility falls below 10 m (33 ft). The winter climate in Riyadh is cool and dry, with average temperatures ranging from 47–68 °F (8.3 – 20 °C). Rain is rare in Riyadh; the average annual rainfall in the city is 4 inches (100 mm).

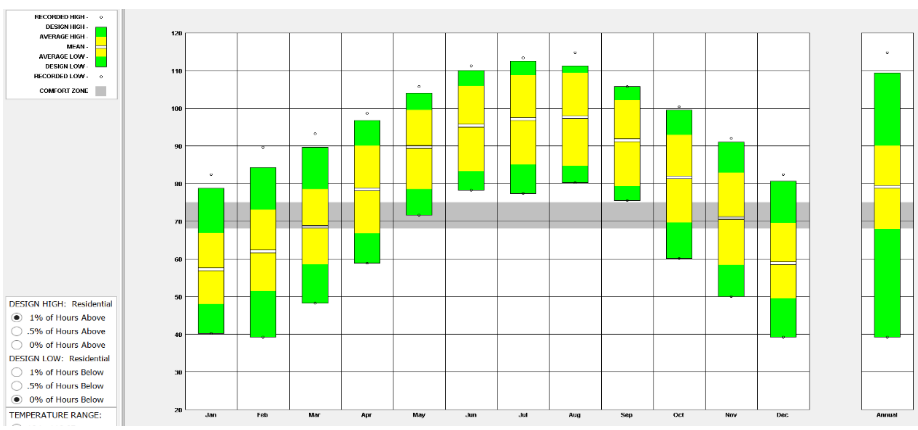


Fig. 1.1. Riyadh annual temperature range.

Riyadh is the largest city in Saudi Arabia by physical area, covering 1,798 km (Alahmadi, Atkinson & Martin 2013). Riyadh is also the most populous region, comprising 24% of the total population of Saudi Arabia. Saudis constitute 61% of the total population in the city; immigrants and expatriate workers make up the remaining 39%. Riyadh is also one of the fastest-growing cities in the Middle East. Table 1.1 shows population growth from 1950 to 2020, with an average population growth of 9.6%.

Table 1.1
Riyadh population growth.

Year	Population	Growth %
1950	111 000	15.6
1960	156 000	4.05
1970	408 000	16.1
1980	1 055 000	15.8
1990	2 325 000	12.03
2000	3 567 000	5.34
2010	5 220 000	4.63
2020	7 231 000	3.85

Source: United Nations World Population Prospects

Part of the Saudi Vision 2030 is to make Riyadh one of the largest ten city economies in the world. While also aiming to reduce fuel use for electricity and replace it with alternative renewable sources. Currently, Riyadh is the 40th largest city economy in the world. To reach the Saudi Arabia vision, the government aims to increase the current number of people living in the city from 7.5 million to between 15–20 million by 2030 (Amlôt 2021).

The current study is informed by the Saudi Arabian government's concurrent goals of reducing electricity usage and dramatically increasing the population of Riyadh. The government vision 2030 to reduce electricity usage (and subsequent increase in electricity prices) has generated a great deal of concern about energy consumption for Saudis. especially as new building structures are planned and designed. In addition, plans to increase the population in the capital city of Saudi Arabia has made the current study even more essential. This study addresses the influence of culture on occupant behaviour and its impact on energy use. This will improve the accuracy of the energy prediction and help to better estimate the energy use during the design stage.

1.3. Significance of the study

Closing the gap between energy prediction and real energy use of buildings will increase energy performance. As discussed earlier, occupant behaviour is one of the main reasons causing the gap in energy use predictions (Menezes et al. 2012). In Saudi Arabia, only a few studies have investigated the impact of Saudi culture on occupant behaviour. However, there are no studies that have investigated the impact of Saudi cultural behaviour on energy use. This study will investigate the influence of occupant behaviour concerning the time-use of rooms, which is unique to Saudi homes. This study will address the gap in the literature related to understanding the impact of Saudi-specific occupant behaviour on energy use. Furthermore, findings from this study will contribute to improving the accuracy of load schedules, with more accurate time-use data for guest rooms that have unique characteristics, strongly influenced by Saudi culture.

1.4. Primary research questions

- Will cultural occupant behaviour play a significant role in energy use predictions in Saudi houses?

1.4.1. Sub-questions

- How do cultural aspects of behaviour affect the energy consumption of houses in Riyadh?
- Without considering cultural aspects, what is the gap between energy simulation predictions and real energy use for houses in Riyadh, Saudi Arabia?
- To what extent does considering people's behaviour in energy simulations close the gap between the predicted energy use and real energy use for houses in Riyadh, Saudi Arabia?

1.5. Research design

This study focuses on Saudi Arabian cultural influences on occupant energy use behaviour in residential buildings in Riyadh. In particular, this research aims to close the gap between energy predictions and real energy use by considering cultural behaviour that affects energy consumption. This research focuses on three main areas: 1) energy performance, 2) occupant behaviour, and 3) culture. The researcher identified these three areas as critical factors in addressing research questions. The literature review (chapter 2) discusses previous studies for all three areas. Previous studies identified culture as having a big impact on house design and people's behaviour. However, the researcher found little information regarding the impact of culture on occupants' behaviour in Saudi Arabian homes.

Occupant behaviour, which is of course influenced by culture, is not very well understood. To identify exact behaviour influenced by culture, the researcher began by identifying and analysing five case studies. The purpose of the case study analysis was to compare energy predictions with and without considering cultural behaviour and comparing them with real energy use. The review included factors related to building design, circulation, privacy, and energy use to identify occupant behaviour that significantly impacts energy consumption in residential houses. The aim of this experiment was to identify the energy performance gap with and without considering this behaviour (chapter 4).

The next stage of the study included the collection of time-use data (TUD) using surveys. The researcher conducted two surveys. He started using a 21-question closed-ended quantitative multiple-choice survey that was organised into four sections. The aim of the initial survey was to understand which rooms had unique time-use patterns and the variables that influenced those patterns. The researcher prepared a second survey to gather additional data, using an open-answer design. The survey included 11 questions focused on cultural behaviour (chapter 5).

After gathering the survey data, the researcher used the data to improve the accuracy of the time-use for the surveyed rooms. This was achieved by creating an occupancy prediction tool and generating new time-use schedules for the targeted zones using a machine learning algorithm. The model was created for use with single-family homes in Riyadh, Saudi Arabia. The researcher's final step was to validate the tool and define how much it could close the gap between energy predictions and real energy use. This was achieved by running an energy simulation experiment for five representative case studies, using the tool to improve the time-use input, and then comparing the outcome with the real electricity bill.

1.6. Definition of terms

This section includes common terms and their respective definitions used throughout this study.

Building Energy Simulation or Modelling: Building Energy Simulation or Modelling is a software system that can analyse different types of building input data to predict or estimate the energy that is being used in buildings at various time frequencies (i.e., at distinct points in time, hourly or inclusive of the entire year of operation). Energy programs typically use open platforms and different data resources such as climate data, common envelope materials and insulation, equipment, internal gain from lighting, occupants' schedule, cooling and heating system, and ventilation. The outputs of energy use predictions are usually categorized in different schedules: heating and cooling, artificial light, electricity, gas, fan, utility rate input, and energy cost (Rosenbaum 2003).

Building performance: Building performance is the quality of the integration of architecture, occupant behaviour, engineering sciences, and facilities management (Amaratunga & Baldry 2001).

Circulation: Circulation refers to people's movement in the house from one location to another.

Courtyard: A courtyard is an open, unroofed area located in the middle of a building and surrounded by the building on all sides. The access to the courtyard area is through the building.

Culture: The way of life for a particular group of people at a particular time, especially in their general customs and beliefs.

Modern house: A modern house was a new building style that became popular in the 1950s. Modernism was a global style that rejected all traditional house styles and reflected differences in people's behaviour and culture. Modern houses used new construction technologies such as glass, steel, and concrete. The layout of modern houses included a setback layout instead of a courtyard.

Overlooking: Overlooking is the visual observation of the yard or rooms by neighbours or strangers from outside the property without the landlord's permission.

People's behaviour: The behaviour of a person is defined as the ability to experience physical, mental, emotional, and social activities during one's life. The behaviour of a person is mostly dictated by ethics, society, culture, morals, and values.

Privacy: There are different definitions for privacy. The best definition for the current research is the protection of family members from visual observation by guests who are not of the same gender.

Setback: Setback is a layout style for buildings. It is a design that places houses in the middle of the land and leaves space between the building and the property boundary. The setback layout could be a choice made by the owner or rules governed by the council of a city. In Riyadh city, the rules indicate the owner must leave at least a two-meter space from their neighbours on the side of the house. For the front of the house, the space must be at least 1/5 the width of the street. For example, if the width of the street is 30m, then the space must be 6m.

Traditional house: A traditional house is typically defined as a single-family (usually 2-story) home. Traditional houses are affected by societal culture and human behaviour that has remained unchanged in a group of people for a long time. Before the discovery of oil, traditional houses were the only house design in Riyadh. They were made of mud and included a courtyard (Fig. 1.2). After oil was discovered in 1938, government buildings and some homes were built with concrete. Despite the increased use of concrete, traditional houses were the most common type of homes in Riyadh until the 1960s.

Villa: A villa is a freestanding house, set back in boundary from all sides of the lot. Villas usually have two floors and stairs to reach the roof (Fig. 1.3).

Fig. 1.2. Example of the Najdi traditional house in Riyadh, Saudi Arabia. The example shows no windows facing outside. All the rooms in the traditional houses face the courtyard, except the male guest zone.

Fig. 1.3. A standard villa in Riyadh; the rooms in the villa are facing outside.

1.7. Summary

This research aims to close the gap in predicting energy consumption for single-family residential houses in Riyadh, Saudi Arabia. In particular, the study focuses on the time-use of guest room spaces that are significantly affected by unique occupancy and occupant behaviour related to Saudi culture. The research methodology includes the collection of time-use survey data and developing a machine learning (ML) model to predict time-use patterns for rooms in single-family detached homes. Findings from survey data and the generated ML model were evaluated with case studies using actual metered data and predicted energy consumption via simulation. This methodology addressed how culture-influenced occupancy patterns enable accurate energy consumption predictions for residential buildings in Riyadh, Saudi Arabia. The introductory chapter covered the research background, the problem statement with research questions, and the purpose of the study, its significance and limitations. The remaining structure of the dissertation is as follows:

- The second chapter includes the literature review, covering the research's focus areas: energy performance, occupant behaviour, and culture.
- The third chapter reviews the methodology of the dissertation, which includes data collection and analysis methods, as well as the method used to develop and evaluate the occupant prediction model.
- The fourth chapter includes information from the two pilot studies, conducted to investigate more deeply occupant behaviour influenced by Saudi culture in selected homes.
- The fifth chapter includes an overview of the data and analysis from the two surveys (primary survey and main survey).

- The sixth chapter describes the time-use prediction model and includes two main sections: 1) the model development and 2) the evaluation of the model.
- The last chapter is the discussion of the dissertation, which includes research findings, study limitations, and future work.

2. Literature review

The goal of this research is to improve the accuracy of residential building energy predictions in Saudi Arabia. To achieve this goal, the researcher examined the influence of Saudi culture on occupant behaviour for individuals living in single-family homes in Riyadh, Saudi Arabia. This research focused on privacy, which is a significant part of Saudi culture (Edwards et al. 2004).

Chapter two reviews cultural behaviour, which impacts the type of homes and daily lifestyles of individuals living in Saudi Arabia. The chapter also includes a review of the most recent techniques, methods and tools used to examine the impact of culturally influenced occupant behaviour on energy prediction. A top-down approach is used to review and discuss the most important resources related to this research (Denney & Tewksbury 2013). The top-down method begins with a review of the impact of Saudi culture, specifically the circulation and privacy of spaces, on house types and configurations in the city of Riyadh. The chapter then follows with a top-down critical assessment of three areas: 1) the main factors that contribute to energy performance gaps, 2) the techniques used to resolve performance gaps, including machine learning (ML), and 3) an examination of approaches used for ML data collection, specifically cultural data.

2.1. Energy demand

Energy use in buildings has a significant impact on global warming (Santamouris 2014). In Europe, approximately 40% of total energy consumption occurs in buildings (27% in residential buildings and 13% in commercial buildings), and 36% of energy consumption is released into the environment in the form of CO_2 emissions (European Commission 2016). In the United States, many studies have shown that approximately 22% of total energy consumption occurs in residential buildings and about 19% in commercial buildings (Mendes et al. 2014). Building energy efficiency could have significant economic, social, and environmental benefits.

In 2014 in Saudi Arabia, 40% of CO2 emissions were caused by electricity consumption (Wogan, Carey and Cooke 2019). Between 2002 and 2018, the average energy consumption increased by 7% (193,472,186 MWh). Residential buildings are the highest demand sector, making up 44% of energy consumption (Fig. 2.1) (Amran et al. 2020 ; Demirbas, Hashem & Bakhsh 2017).

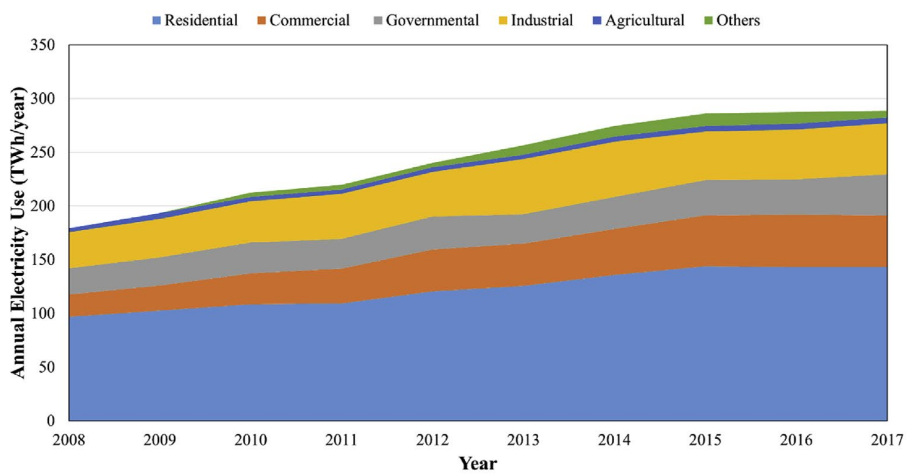


Fig. 2.1. Annual electricity by sector, Saudi Arabia (Krarti et al., 2020)

This problem is not limited to Saudi Arabia but is a worldwide issue. Researchers have forecasted that by 2050, buildings will account for more than one-third of total energy consumption and will be responsible for a third of global CO2 emissions (International Energy Agency 2014).

The literature cites many reasons for the rise in energy consumption, including increases in electrical processes, the ever-growing population, higher standards of living leading to increased consumerism, and urbanisation. People now spend more time working, studying, and living in buildings (Jing et al. 2017). With increased energy consumption demand in buildings resulting in increased CO2 emissions, these concerning facts demonstrate the need to manage energy consumption within buildings efficiently.

In Saudi Arabia most of the energy consumption comes from electricity. Compare the energy use intensity (EUI) in Saudi Arabia with its neighbouring countries (Fig. 2.2), which have similar weather/climate conditions. Saudi Arabia has the highest energy consumption for residential buildings compared to the other countries, followed by Egypt. Located in the middle of the country, the capital Riyadh has the highest energy consumption compared to other regions. The energy consumption in the west, where holy Makkah is located, also has very high energy consumption (Fig. 2.3).

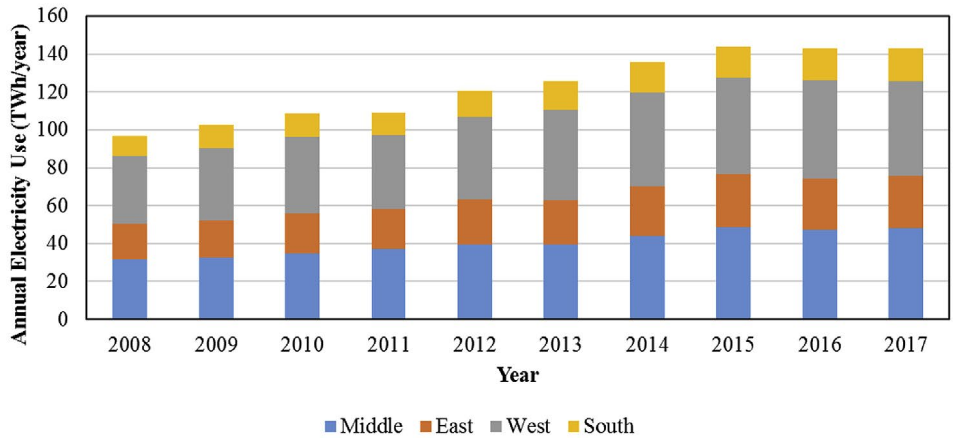


Fig. 2.2. Annual electricity consumption for residential buildings in Saudi Arabia (Krarti et al., 2020)

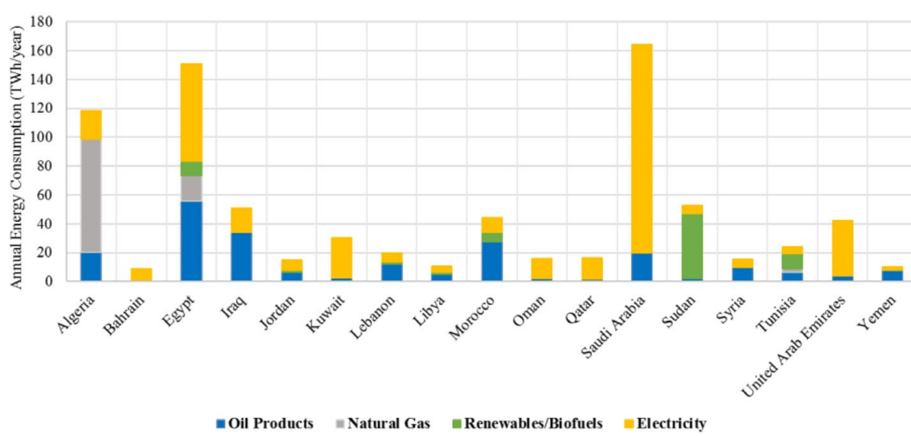


Fig. 2.3. Comparison of annual energy consumption in Saudi Arabia and neighbouring countries (Krarti, 2019)

Governments around the world are taking various measures to reduce energy consumption and mitigate its environmental impacts. In Europe, objectives for 2020 included increasing the amount of renewable energy to 20% of the total energy production mix and reducing gas emission levels by 20% (Sandels, Widén & Nordström 2014). China achieved a 2020 goal of 680 GW installed renewable power capacity and established a goal of a 20% share of non-fossil fuel total consumption by 2030 (Renewable energy market analysis: The GCC Region 2016). According to the Dubai Clean Energy Strategy 2050 report, Dubai aims to generate 7% of its total energy consumption from clean energy sources by 2020, 25% by 2030, and 57% by 2050 (UAE Government News 2016). Many other countries around the world have also set plans to generate renewable energy within the next ten years.

World governments have also encouraged the use of renewable energy and sustainable techniques by creating new policies on building electricity use. The United States of America (Energy Independence and Security Act of 2007) and the United Kingdom have released new building regulations to be implemented by 2030 (Briller 2013; Menezes et al. 2012). In Saudi Arabia, two of the main 2030 goals are to reduce energy consumption and increase the use of renewable energy.

In 2017, the Electricity and Cogeneration Regulatory Authority (ECRA) in Saudi Arabia announced that there were 7 million residential consumers. Most of the power generated by residential consumers was because of HVAC system usage due to the hot climate in Saudi Arabia (AlGhamdi 2019). The ECRA increased the electricity tariff threefold to encourage people to use sustainable techniques (Vision Progress | Saudi Vision 2030 2018). By the beginning of 2018, the government had increased the residential sector electricity tariff policy to 250% of previous tariff rates (Harbi & Csala 2019). This strategy was intended to reduce energy consumption and advance the use of more renewable energy resources during a decade-long transformation period (2020 to 2030) from an oil-reliant economy to a knowledge-based economy.

To move towards the use of sustainable and renewable energy in buildings, energy simulation software must be considered during the building design stage. The energy industry is still considered to be relatively new. Beginning in the early 1990s, the design of passive houses started to develop from conceptual ideas to actual buildings (Ionescu et al. 2015). One of the biggest ongoing problems the energy industry faces is the energy performance gap—the gap between energy consumption predicted in the design stage of a building and real energy use after the building was built and occupied. Evidence suggests energy use in buildings could reach 2.5 times the predicted energy consumption rate (de Wilde 2014). This huge gap could have a massive impact on the management of energy supply and demand (de Wilde 2014).

To reduce energy performance gaps, it is important to study the main inputs for building energy consumption. Simply using energy simulation software is not sufficient to achieve a sustainable solution that can bridge the energy performance gap. Software users need to be aware of the factors that influence a building's design, including building location, occupant choices, social behaviour, economy, and culture. These factors must be considered to identify ways to conserve energy and also to calculate energy consumption accurately. The next section reviews how culture, specifically privacy, impacts house designs in Riyadh, Saudi Arabia.

2.2. Cultural influence on occupant behaviour in Riyadh, Saudi Arabia

Culture plays an important role in shaping the lifestyle of Saudi Arabians. Saudi people are conservative and family-oriented, reflected in their behaviour and home design, operation, and facilities. As such, Saudi people prefer a design plan for their homes that facilitates a high level of privacy (Saleh 2001). The following section delves deeper into Saudi culture from the perspective of privacy and examines how it influences building design, occupant behaviour, and energy use in Riyadh. It should be noted that other regions of Saudi Arabia share similar cultural considerations; however, climate variations impact home designs differently.

2.2.1. Privacy culture in Saudi Arabia

The Islamic religion is the primary influence on many cultural components in Saudi society. As noted, citizens of Saudi Arabia are conservative and family-oriented (Al Surf, Susilawati and Trigunarsyah 2012). These cultural qualities have a direct impact on residential building energy consumption. For example, in Saudi Arabia, it is common for multiple generations to live in the same house at the same time, which leads to higher levels of energy consumption (Grube & Michell 1978).

Residential building design and occupant behaviour in terms of privacy differ in multiple families versus single-family homes (Alshahrani & Boait 2019). Saudi society has undergone many changes due to increases in development and urbanisation. Many customs and traditions have been subjected to these changes, leading to the gradual reduction of multiple family homes. In both multi- and single-family homes, many cultural values and principles are still reflected in how homes are designed to preserve privacy (Babangida & Katsina 2018). Despite this, there are still houses where multiple families live together. This, however, is out of the scope of this study.

The next sections review home designs in terms of privacy. This is divided into two main periods: 1) the traditional house period from 1824 to 1960 after Riyadh became the capital of Saudi Arabia and 2) the modern house period after 1960 when the oil industry started to mature, leading to increased economic growth and greater urban development and buildings.

2.2.2. The houses in Riyadh (1824 to 1960)

Between 1824 and 1960, Islamic beliefs and culture heavily influenced the urban design of cities and the lifestyle of Saudi Arabians. There was a strong relationship between culture and house design (Babangida & Katsina 2018). Urban designs enhanced participation in economic and religious life within the community. The city was divided into public, semi-private, and private areas. Fig. 2.4 shows how the public area was defined by the mainline in the middle of the city and included streets that housed public facilities, such as the main mosque, shops, and cafes. Private narrow streets were designated for residential buildings, as depicted in Fig. 2.5.

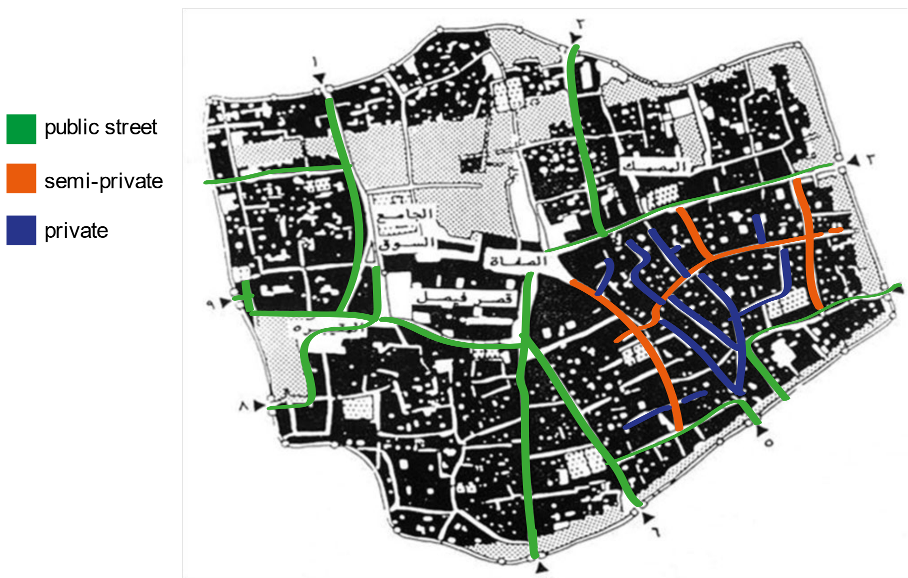


Fig. 2.4. Riyadh, 1863 (Centre for the Study of the Built Environment).

Fig. 2.5. The traditional city showing some of the sustainable techniques, natural materials, narrow streets, overhang, and self-shading.

There were no building codes in Saudi Arabia between 1824-1960. A courtyard technique was commonly used for house designs and most houses were built using mud (Al-Hathloul 2017). Courtyards helped promote privacy. Windows and openings typically faced the courtyard and also allowed for greater ventilation to address the heat associated with the desert climate (Alwetaishi 2018). Most residential buildings were left incomplete to allow for home additions to accommodate family growth. Fig. 2.6 depicts home expansions and how homes were linked to each other. As shown in the image, the householder's son would build his house next to his family, creating a family-house complex (multi-family house). Until the 1960s, family complexes were allowed as long as the privacy of extended families was not violated (Grube & Michell 1978).

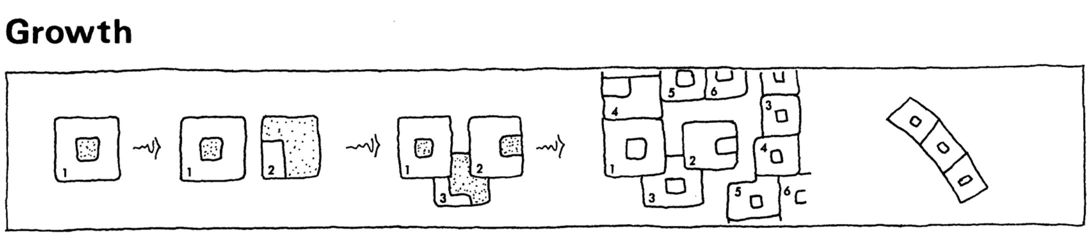


Fig. 2.6. The urban growth in Riyadh (Grube and Michell, 1978).

Akbar (1980) explained how homes were designed with privacy in mind. Fig. 2.7 illustrates a plan of a traditional house in Riyadh during this period. The reception (also known as the male guest area) was located next to the entrance, which allowed for easy access and a high degree of visual privacy. Every house had at least one reception space for visitors. All rooms had access to the courtyard, except the guest room, which had less privacy. All windows and openings faced the courtyard except the guest room window, which faced the street.

Before the 1960s, many buildings in Riyadh were designed based on the Najd style. A triangular shape and courtyard were common features of the Najd style. In warmer months, Saudi people used to sleep on the roof terrace at night. The roof was also used for family celebrations, including weddings and gatherings of female guests. Traditional houses had a family gathering room that was used daily, and families used the courtyard for occasional parties and larger gatherings.

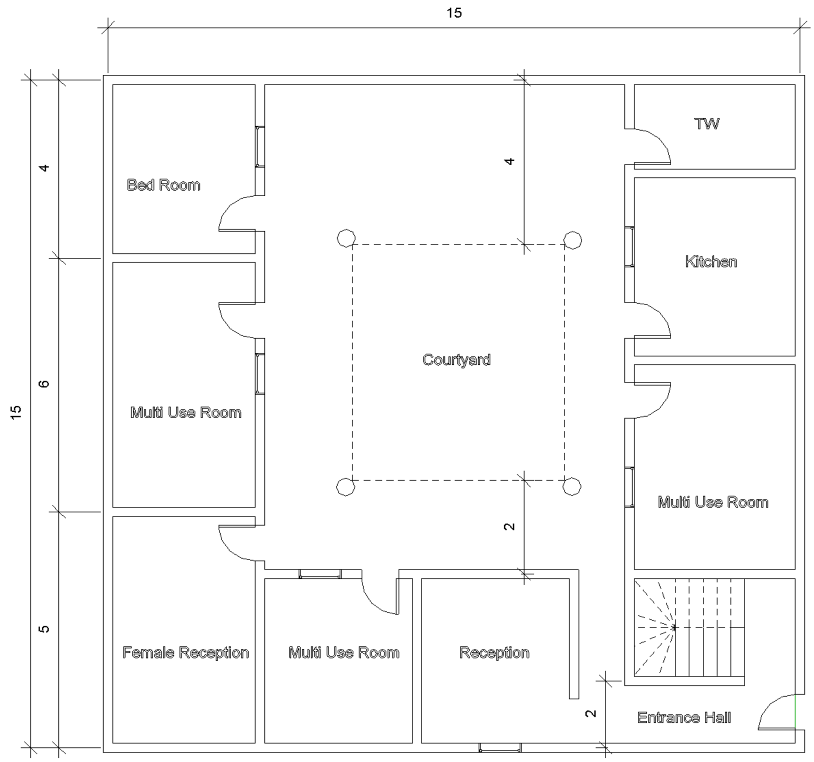


Fig. 2.7. Houses 1824 to 1960 (Akbar 1980).

2.2.3. Houses in Riyadh after 1960

Until the 1950s, comprehensive town and city planning was unknown to the Saudi Arabian government. In the 1950s and 1960s, the growth of the oil industry caused Saudi cities to develop and change (Nurunnabi 2017). However, there was not enough time and capacity to adapt to the country's growth while working with traditional principles. Urban planning for a new Riyadh was needed. Due to a lack of local expertise, foreign consultants assisted with initial planning efforts; however, many of these experts failed to understand local environmental issues and cultural preferences regarding buildings. As a result, Doxiadis Associates were hired to design the master plan for the development of Riyadh (Alkhedheiri 1998). Fig. 2.8 shows the design plan of Riyadh using a grid-planning concept created by Doxiadis Associates.

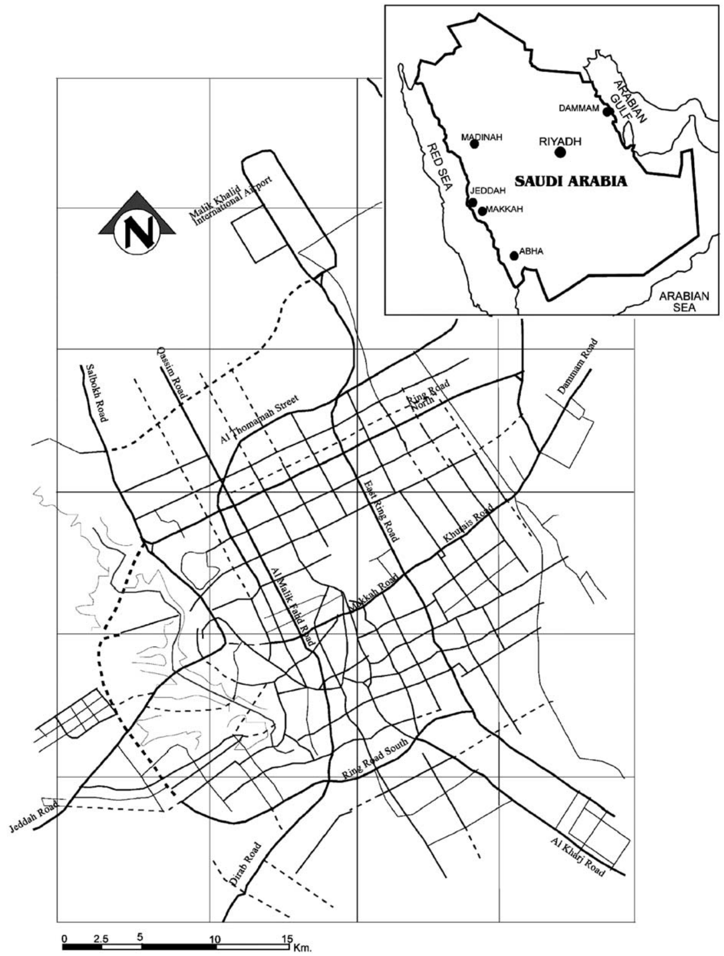


Fig. 2.8. The plan of Riyadh after the traditional period, design by Doxiadis Associates using a grid-planning concept (Garba 2004).

In 1953, government agencies were relocated from Makkah to Riyadh. An area called Al-Malaz was chosen as the location to build the first modern complex project in the city. Fig. 2.9 shows the plan of the Al-Malaz district in Riyadh. Foreign architects who planned Al-Malaz used a gridiron pattern with wide streets. Al-Malaz, located 4.5km north-east of Riyadh, had 754 villas and 180 apartments. The area included public buildings (such as a library), gardens, a zoo, and other facilities, which were new to Riyadh. It also included the first university in Saudi Arabia. Al-Malaz was approximately 500 hectares, which was significantly larger than the old city. With the new layout, the population density was reduced to one-fifth of the traditional city. However, the street area was three times larger (Al-Said 2003).

In Al-Malaz (Fig. 2.9), the concept of concrete villas was introduced. The 25m x 25m two-story villa was built as a setback design. Windows faced the front and back yard (Mubarak 2004). A solid fenced wall surrounded the parameters of home property areas to maintain high levels of privacy. These new buildings and urban designs were adopted in other cities of Saudi Arabia.

Fig. 2.9. Al-Malaz district in Saudi Arabia Riyadh, new street pattern, grid, and house form (Centre for the Study of the Built Environment).

The grid planning concept and the setback technique for homes resulted in limited social and cultural activities. Compared to previous years, children could no longer safely play in the street (Almahmood et al. 2017). This new setback technique reduced building performance and led to social problems (Eben Saleh 1997). The setback technique changed the orientation of windows and openings from facing the courtyard to facing outwards, so people in new villas could see their neighbours' front yards (Fig. 2.11). This change violated the privacy of neighbours (Fig. 2.10). In addition, the setback technique increased the time that buildings were exposed to the sun, thus increasing the level of heat inside villas (Fardous 2019; Giddings, Almehrej & Cresciani 2020).

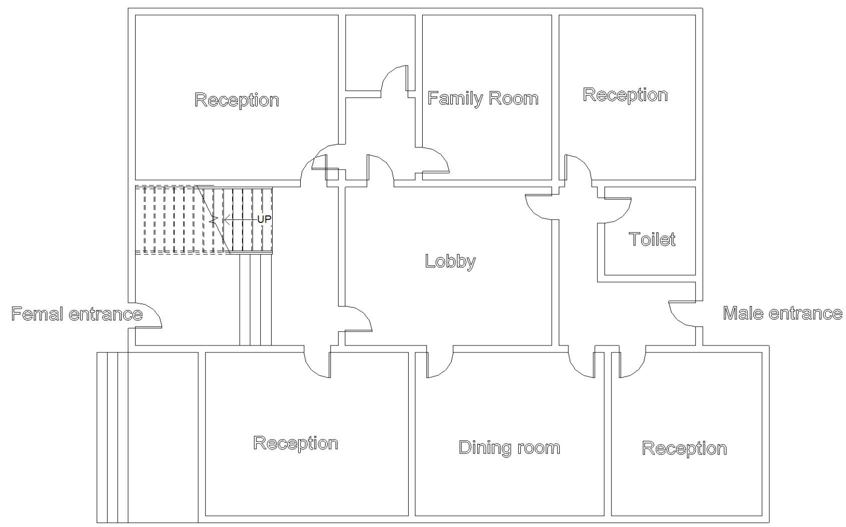


Fig. 2.10. Example of the new type of house built in 1970.

Fig. 2.11 illustrates residential building regulations after 1960. Householders were required to keep two meters between their residence and their neighbours' land. From the street side, the landlord was required to keep an open space that was 20% of the street width. For example, if a street were 20 meters in width, the open-space requirement would have been 4 meters (i.e., 20% of 20 meters).

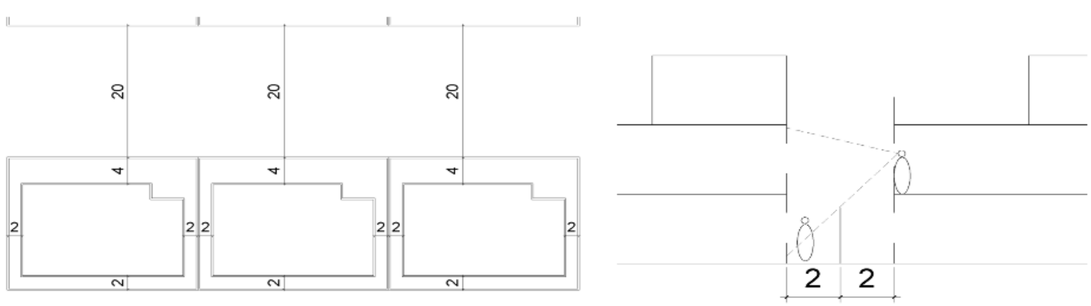


Fig. 2.11. The rule for building homes in Riyadh is to keep 2 meters of space to the neighbouring land at all sides and 20% of the street width for the front side.

Because of this, people started to use their front and back yards less because of the loss of privacy; neighbours could see their family members. To address issues of privacy, Saudi people added steel structures on top of the wall between homes. Fig. 2.12 shows an image of residential building steel walls. As depicted in Fig. 2.13, these structures could have been as tall as two stories to ensure their neighbours' view was blocked.

Fig. 2.12. Residential buildings' steel walls.

Fig. 2.13. The new pattern style.

Al-Hemaidi (1996) investigated whether the new building style offended occupational privacy. His survey covered 61 villas in one of the modern districts of Riyadh. He discovered that only six families said that they used the front yard for family activities. Interestingly, to protect their family from being observed by their neighbours, these six families surrounded their homes with a steel structure on the top of the wall. In addition, less than a third of residents indicated that they opened their living room windows. Al-Hemaidi stated that the main reason for this was because their neighbours could see them through the windows.

As part of his master thesis, Al-jammaz (2016) investigated the privacy practices of 80 families who lived in modern-family houses in Saudi Arabia. His findings reported that 36% of the sample indicated that they did not use their windows during the day. 44% reported that they spent less than an hour in front of their windows. In addition, 40% of families said they closed their curtains all day, 27% said the curtains remained closed for more than 12 hours a day, and 72% of the sample indicated they did not open curtains during the night. Aljammaz (2016) also examined occupant preference of using the front or backyard for family gatherings. In his sample of participants, the front yard had less privacy (as it was closer to the male guest zone). The backyard had greater privacy (which was closer to the family zone). He found that 87% of the sample preferred using the backyard. Further, he noted that 73% of families indicated they did not like their family members to be seen by neighbours while playing in the front yard.

2.2.4. Conclusion

In Saudi culture, family ties and relationships with neighbours were major themes that impacted the evolution of house designs in Riyadh (Fadan 1983). Culturally, Saudi people used to invite visitors and neighbours to their houses a great deal of the time. This strong cultural element still affects how houses and other building types are designed today.

Saudi Arabian citizens want to have visitors while also not compromising their privacy. Home designs that divided homes into two or three zones offered families the opportunity to maintain their privacy, while also affording their social and cultural practices of gathering. The first zone was located near the entrance for visitors and had less privacy. The second level was for female guests. The third level was located further away from the entrance and was intended for family members. Family zones had a family room, living hall, and kitchen. The second and third levels were not physically or visibly accessible to male visitors. Guest zones were required in every house, even if they were only to be used a few days a year. These home designs that promoted privacy took different shapes in different Muslim societies; however, the consensus to prioritize privacy was very similar.

In 1981, Jamel Akbar created a diagram to define home privacy levels from the public perspective, ranging from number one (little privacy) to number five (a very high level of privacy). In this experiment, house spaces were defined according to privacy level, determined by the category of the individual. Categories include male visitors, relatives, and inhabitants and female visitors, relatives, and inhabitants.

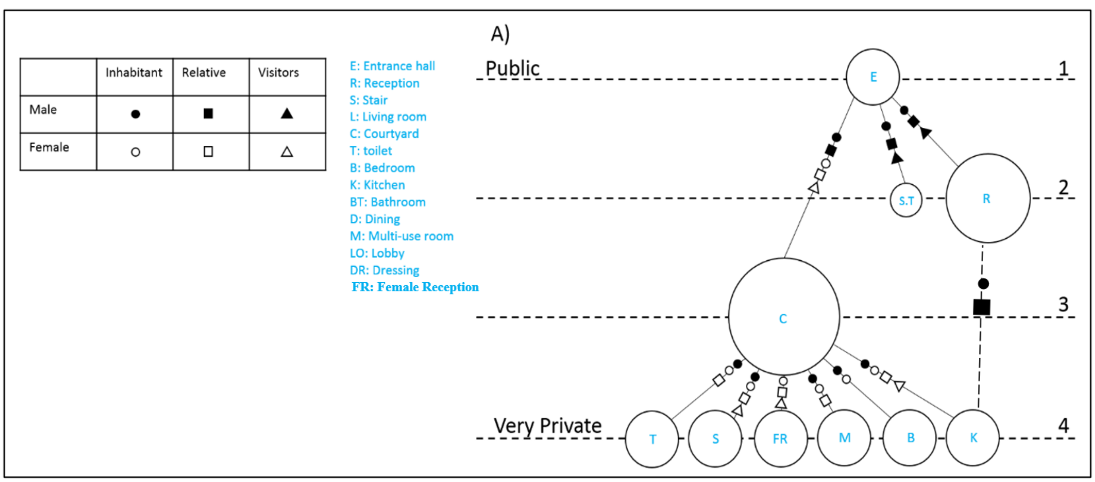


Fig. 2.14. Traditional house privacy level diagram.

Fig. 2.14 represents a traditional house privacy level diagram during the period 1824 to 1960. The diagram shows that in traditional houses, male guests could only use the entrance and the reception. The family could go from the entrance to the courtyard without passing the reception. They could also go to any room that had doors and windows facing the courtyard.

Fig. 2.15 illustrates modern house privacy level diagrams (Diagram B and C) for homes built after the 1960s. In modern homes with the setback technique design, the courtyard was altered into a living room and dining room comprised of two doors for the male guest zone and family zone. Diagram C shows a modern house, but with a separate female guest zone. The entry is separated by gender; there is an entrance, reception, and dining room for males and a separate entrance, reception, and dining room for females. The diagram also shows that visitors do not pass the third line from the public entrance to the private level; however, an exception is made for relatives who may go farther.

The diagram is used later in this document (in Chapter 4) to highlight changes in privacy levels based on occupant behaviour over the past 50 years in Riyadh, Saudi Arabia. This research includes a pilot study that examined impact of culture on occupant behaviour, in terms of design, privacy, circulation and energy use.

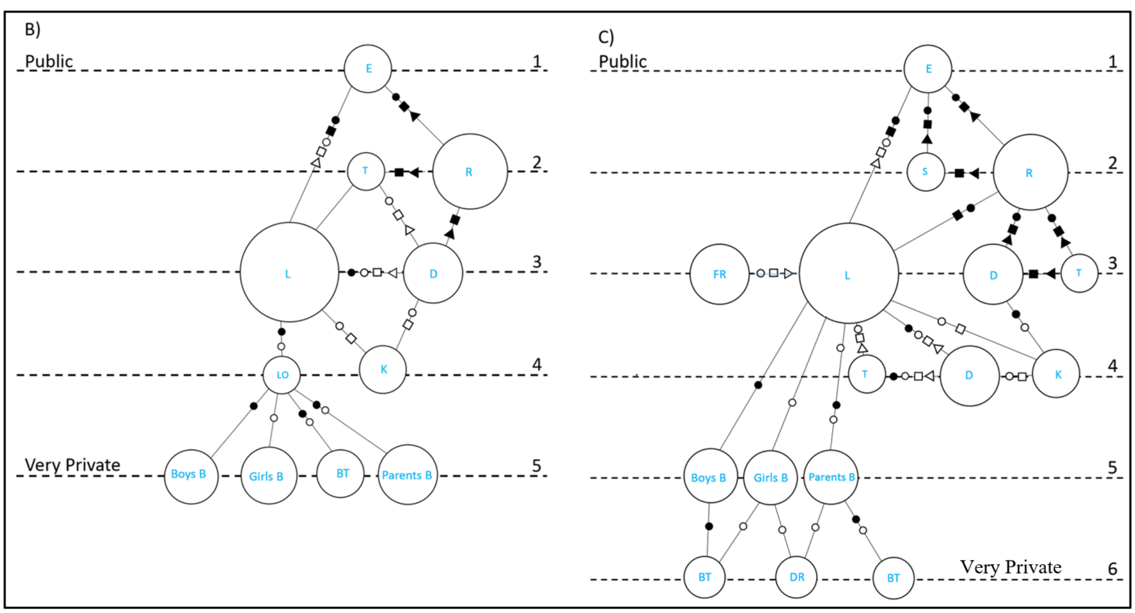


Fig. 2.15. Figure B modern house privacy level diagram; C modern house with separated female zone.

2.3. Energy performance gap

De Wilde (2014) referred to the energy performance gap as the gap between the energy prediction of a building during the design stage and the real energy use during the operation of the building. Three factors can contribute to this gap, including 1) the building design, 2) the construction process, and 3) the operation or mode of building use.

Buildings should be functionally designed to suit the users' requirements. During the design stage, miscommunication or misunderstandings between and among the design team and/or the client could lead to a design error in some spaces, subsequently creating a gap in the total energy performance. The complex process of transforming the building's technical attributes (due to uncertainties at various design, development, and construction phases) into a computer model might also contribute to this gap (Yoshino, Hong & Nord 2017). Furthermore, the incorrect use of simulation software tools could cause a mismatch in energy predictions and actual energy use. Given these considerations, software users must have an appropriate level of knowledge (Niu, Pan & Zhao, 2016).

Reports from the construction industry have shown that construction quality does not always match design specifications. According to Almeida et al. (2010), there is often a lack of attention given to insulation materials. Similarly, inadequate contractor knowledge and/or ability to interpret simple or complex design details and specifications (such as quality and dimensions) also cause problems. Some contractors use substandard materials to spend less on projects and extort construction fees. These problems are difficult to solve after construction and ultimately cause a gap in the building's energy performance (Turner & Frankel 2008).

The energy performance gap can also be attributed to the user operation of buildings (Menezes et al. 2012). For example, the amount and duration of lighting, ventilation, equipment, and HVAC system usage for cooling and heating impact energy consumption. Because this behaviour changes over time and is hard to accurately foresee, it is nearly impossible to predict the operations and occupant behaviour of buildings. However, researchers are now utilising behaviour data to more precisely predict occupant behaviour to close energy performance gaps.

Many researchers also focus on bridging the performance gap between building design and operations (Niu, Pan & Zhao 2016). To fill this gap, designers make assumptions and approximations to predict energy consumption using guideline designs and standards. For example, the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) uses climate data to predict what systems are suitable for a specific climate zone and the amount of energy needed to cool or heat a specific space to reach thermal comfort for building users. These standards use the climate of a location as a parameter for predictions and predict fixed loads and schedules for all buildings in the same climate (Ding et al. 2015).

With the current rising demand in the energy industry, researchers are working to make continuous improvements in energy predictions (Dwyer 2013). Evidence suggests that the use of Artificial Intelligence (AI) can improve building performance and efficiency. However, at present, there is a dearth of information to improve data inputs. Further research efforts can leverage the use of more data and AI (Wang & Srinivasan 2017).

The next section focuses on previously conducted studies that aimed to examine occupant behaviour. A review of the literature includes key considerations and data collection methods related to occupant behaviour. The section also highlights suitable approaches used to achieve this goal.

2.4. *Occupant behaviour*

2.4.1. *Data collection and analysis*

This section focuses on closing the energy performance gap by improving occupant behaviour inputs (Hu et al. 2020). Researchers have grouped this data collection method into two main categories: 1) post-occupant data collection and 2) pre-occupant data collection (Breadsell, Byrne & Morrison 2019). Occupant data collection involves collecting detailed data about occupant behaviour for a specific group of people or buildings. It is difficult for participants and researchers to garner all the details by interview alone, so sensors, cameras, and other devices in the building are used to collect data in detail.

The post-occupant method is used to collect data for an existing building. This method is also used to evaluate the building and the experience of its occupants and find and solve operational problems. For example, if the goal of an experiment were to find the relationship between the number of employees entering the office at a certain time and the amount of lighting they need, then the post-occupant method would be the most suitable approach. The advantage of this method is the accuracy of the devices to collect detailed data (Li, Ng & Skitmore 2018). For example, placing sensors in the office to collect the number of employees and the amount of lighting used for some time (e.g., one year) will give a clear understanding of their behaviour. The disadvantage of this method is that the outcome cannot be applied in other buildings because every building and occupant is different (Jia, Srinivasan and Raheem 2017).

The pre-occupant method is considered a new approach in the field of building energy performance gaps. The pre-occupant method involves collecting occupant behaviour data during design stages to predict future behaviour (Hajj-Hassan et al. 2020). This can be achieved by taking advantage of new energy simulation software (Yaman et al. 2021). The advantage of this method is that it gives insights into occupant behaviour in real-life settings during the design stage. These data enable the designer and the owner to evaluate the design and energy use before a building is built (Uddin et al. 2021). Changing the building design during the design stage to reduce future real-life energy consumption is easier and cheaper than reducing energy consumption after the building is built. The pre-occupant data method, also known as virtual building simulation input, is the method that would be used to test such designs. Pre-occupant data is gathered from a database of similar cases. The disadvantage of this method is that the accuracy of predictions is not as good as post-occupancy findings (Panchabikesan Haghighat & El Mankibi 2021).

Both pre-occupant and post-occupant data collection methods can be analysed using statistical analyses or data mining (Dong et al. 2018). Statistical analysis is the most common method researchers use to understand occupant behavioural patterns and factors influencing the behaviour that affects energy consumption (Muroni et al. 2019). Doing so allows researchers to better predict energy usage and to help close the energy consumption gap. Using machine learning to analyse occupant behaviour is more effective than statistical analyses, especially for complex data. Data mining is an efficient approach to predict occupant behaviour and can include examining data related to an occupant's time in a building or a specific space (Jia et al. 2021; Laaroussi et al. 2020).

2.4.2. Calculating energy consumption models

Researchers have discovered that energy simulation requires a human-in-the-loop (HITL) model to improve accuracy. The HITL model is an interactive model that needs human operators to perform simulations (D'Oca, Hong & Langevin 2018). Energy simulations are categorised into three main models: the white box, grey box, and black box.

The white box approach, or engineering method, is a method that provides the energy model with physical knowledge (Li, Han & Xu 2014). The input describes building details and physical behaviour. The white box works as an energy simulation for different applications and scales. For example, a simple method takes the average monthly temperature and uses it to predict the annual energy use of buildings (Westphal & Lamberts 2004; White & Reichmuth 1996). The white box approach requires more data and complicated calculations to render outputs. Many software programs have been invented for this purpose. They provide a mechanical system for physical data, calibration, and calculation of outcomes in an energy tool (Crawley et al. 2008). Examples of white box application tools include as EnergyPlus, DOE-2, and ESP-r. ("EnergyPlus | EnergyPlus" 2021; "DOE2.Com Home Page"2021, 2; "Welcome to ESP-r" 2021).

The black box approach involves the use of a database for prediction through artificial intelligence. The model can predict occupant behaviour, including the pattern of space-use in a specific building or predict the operating system, such as cooling or heating, of a specific room (Hu & Xiao 2020). Unlike the white-box approach, the black box method does not require physical inputs. Instead, it uses fitting techniques (means of sample data linked to each other and the description of behaviour) to predict energy consumption (Ciulla & D'Amico 2019). It also requires pre-selected statistical models and training data to improve the model's accuracy (Li, Han and Xu 2014). With the improvement of technology, more researchers are using this relatively new approach to improve the accuracy of occupant behaviour inputs (Carlucci et al. 2020). Different methods used for this approach include multiple linear regression, statistical regression, neural network, support vector machine, and Random Forest (Li & Yao 2020). There are many different machine learning algorithms available to improve occupant behaviour predictions. Researchers choose the algorithm based on the data type, number of parameters, number of samples, and the goal of the model (Ahmad, Mourshed & Rezgui 2017).

The grey box approach combines physical knowledge data and fitting techniques (Zou et al. 2018). The concept of combining the two approaches is to take advantage of both methods and generate a more accurate energy model. For example, researchers can integrate the use of EnergyPlus for physical knowledge (such as design, structure, operation, and climate) with occupant behaviour (Jia & Srinivasan 2020). The reason for this integration is that the black box prediction for occupant behaviour is more accurate than the white box approach (Laaroussi et al. 2020). Many researchers use this method in different ways based on their project type and goals (Jia & Srinivasan 2020). This research uses the grey-box method (using the black-box approach to predict occupant behaviour and integrating it with an energy model) using EnergyPlus to examine the data for the cultural behaviour impact on the accuracy of the energy prediction.

2.4.3. Occupant behaviour, modelling, and simulation

Using occupant behaviour input increases the accuracy of energy simulation outcomes (Clevenger & Haymaker 2006). Predicting random human behaviour is difficult. However, predictions can be improved by collecting data for specific behaviour and identifying links between behaviour and other relevant variables (Langevin, Wen & Gurian 2016). For example, in residential buildings, the social, economic, and climate of the building location influences occupant behaviour (Zou et al. 2018). Behaviour, such as occupant presence in a particular zone, opening/closing of windows, and other activities (Diao et al. 2017), impact energy use in buildings. Using artificial intelligence and the black or grey box approach to train the data about such behaviour will improve the accuracy of predictions. Due to the erratic and unpredictable nature of human behaviour and the limited technology we have available now, predictions will not be exactly accurate. However, they will improve the accuracy of inputs to inform predictions (Ding et al. 2020).

Another important aspect of occupant energy models is the data collection approach. Fig. 2.16 illustrates the structural framework diagram of behaviour modelling and simulation. There are two main ways to collect occupant behaviour data: 1) a technology-based approach, including the use of sensors, or 2) a survey system approach, such as Time-Use Survey Data (TUD) (Du, Pan & Yu 2020). Collecting behaviour data and variables related to behaviour should be linked with a specific behavioural model approach. The existing modelling approach for occupant behaviour inside buildings can be categorised into four main quantitative modelling approaches: 1) stochastic modelling, 2) statistical modelling, 3) data mining, and 4) agent-based modelling (Uddin et al. 2021). Stochastic modelling is used to estimate the probability or patterns of behaviour driven by statistical information. Statistical modelling is used to predict occupant behaviour in indoor and outdoor conditions. The data mining technique is used to study, investigate, and find trends of occupant behaviour patterns. Data mining requires an enormous database to predict an accurate outcome. Agent-based modelling is a prediction model approach that focuses on a personal level or a specific group of people and how the group interacts with one another within indoor building systems (Jia, Srinivasan & Raheem, 2017). The following section will cover stochastic modelling studies. Other modelling approaches are out of the scope of this research.

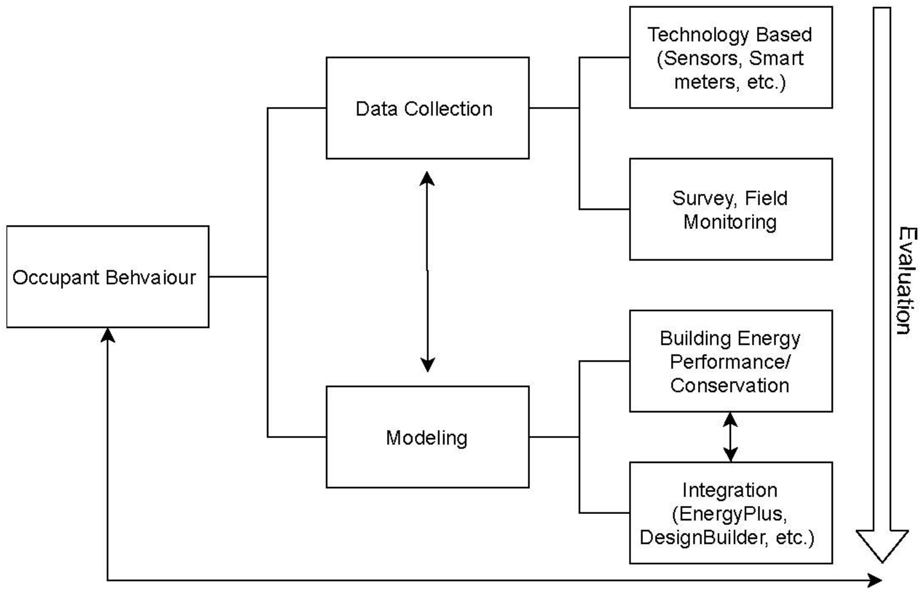


Fig. 2.16. Structural framework diagram of behaviour modelling and simulation (Uddin et al. 2021).

2.4.4. *Stochastic modelling*

Stochastic modelling, otherwise known as the probabilistic model, is used to understand occupant movements and actions inside a building. This approach has four different models: 1) Bernoulli's process, 2) Markov chain, 3) survival analysis, and 4) statistical analyses. Bernoulli's process is a sequence of Bernoulli random variables. It can be used, for example, to find the probability of using spaces in a building (Mun, Kwak & Huh 2021; Sparks 2021). The idea of the Markov model is to find the probability of being in a particular place at a specific time. The Markov chain order is based on probabilities of previous time steps. These transition probabilities, from one state to another, are held in "transition probability matrices", which are obtained by observing occupant states (Sun et al. 2020). A more detailed model called Hidden Markov uses smart metering, observations, and characterisation of occupant presence and behaviour in residential buildings (Liisberg et al. 2016). Survival analysis is a time-to-event outcome variable used to evaluate the period of a status or event (Jia, Srinivasan & Raheem 2017). It is also used to evaluate how long a building could stay unaffected by occupants or users (Barthelmes et al. 2018). Survival analysis uses a regression model to find the relationship between occupant behaviour and electricity use in indoor environments or through a time series model (Panchabikesan, Haghighat & El Mankibi 2021). The probability of this behaviour could be predicted by relevant parameter inputs. This model trains and validates the data and finds a relatively steady relationship between occupants' status. Researchers mostly use this model for indoor environmental data and the presence or absence of information (Laaroussi et al. 2020). This research uses statistical analyses to answer the research questions.

2.5. Limitations of occupant behaviour models

This section reviews the limitations of occupant behaviour models in line with the scope of this research. The six driven factors based on The International Energy Agency Energy in Buildings and Communities Programme (IEA EBC) (Annex 53) are: 1) climate, 2) building envelope, 3) building energy and services systems, 4) Indoor Environmental Quality (IEQ), 5) building operation and maintenance, and 6) occupant behaviour. The last three factors are related to occupant behaviour. One of the findings of Annex 53 is that these three factors could influence as much as the first three or more in terms of building energy use (Yoshino, Hong & Nord 2017, p. 53). Researchers need to understand architecture, building science, computer modelling, and simulation to consider these factors for improved energy predictions (de Wilde 2014). Furthermore, researchers need this knowledge to understand total building energy use to address the energy performance gap.

2.5.1. Social factors

The first three factors in the IEA EBC Annex 53 report are related to the building envelope, structure, and systems. These fixed physical elements, however, are influenced by building operations. According to researchers, physical elements are the base of building energy simulations and they are known as human-influenced factors (Du, Yu & Pan 2020). Merging these factors gives the real energy use of buildings. In the case of new building developments, factors driven by occupant behaviour are found in standards like ASHRAE or owner requirements. However, researchers have compared outcomes from software energy predictions during the design stage with actual building usage and found that the calculated energy performance in software programs did not reflect actual energy use (Yu, Du & Pan, 2019).

Social factors are unique to human behaviour, usually for people living in the same country or region. Social factors influence energy use and have a large number of parameters, such as energy prices, political conditions, local habits, and occupant attitudes. The influence of social factors on energy use varies between countries (Yoshino, Hong & Nord 2017). For example, local habits driven by strong cultural behaviour in one country would have a different (and greater) impact on energy use predictions compared to other countries with lesser cultural considerations (de Wilde 2014; Tam, Almeida & Le 2018).

2.5.2. Level of complexity

The level of model complexity or resolution targeted to improve energy performance must be defined. The level of complexity reflects the details of the data type (Gaetani, Hoes & Hensen 2016). For social statistical analysis models, the number of parameters after analysis is less when the data targets a large number of buildings (Diao et al. 2017). The details and complexity of the data increase when the prediction is only for a few case studies. The level of complexity for social input data (such as family information, related energy use, or local behaviour) is categorised into three levels. They are:

1) Simple: for statistical analyses of a large number of buildings and energy-use data (monthly or annually),

2) Middle: for a few case studies, energy-use data (monthly or daily), and

3) Complex: for detailed diagnostics, energy-use data (daily or hourly).

2.5.3. Probability

One of the problems researchers face is the weak connection between the technology-social interaction and the human-building interaction, both of which affect building energy consumption. Reviews from the IEA EBC Annex 66 indicate that simulations of occupant behaviour in buildings using the probability method can improve energy predictions. In these studies, researchers used the probabilistic models of behaviour for a large-scale survey. The goal was to develop a stochastic occupant-presence model. Lawrence Berkeley National Lab (LBNL) in Berkeley, California and Yan (2017) conducted state of the art energy studies and implemented machine learning algorithms (ML). Algorithms were used to train and validate the data collected for occupant behaviour to enhance building performance (Hong et al. 2020).

Machine learning algorithms build a model based on sample data. It is a smart tool that can learn from training data and find relationships between variables to make a prediction. The algorithm should be chosen based on model goals and the data collection approach to ensure an accurate ML prediction model. Fabi, Andersen, and Corgnati (2015) used regression models to predict occupant model presence behaviour in spaces of buildings. The results of the model are validated using either boot-strap validation, cross-validation, or random-sample validation methods. The most popular regression algorithms used to solve this problem are Support Vector Regression (SVR) and Random Forest (RF). For example, in South Korea, geographical, cultural, and social factors influencing the time-use of HVAC systems were considered in order to predict the ON/OFF function for HVAC use in residential buildings. Data collected was trained and then compared using the two algorithms SVR and RF. Findings show that the RF predicted a more accurate outcome.

The Random Forest method defines the prediction outcome by generating several predictions. This ranges from training data to choosing random samples, then letting individual prediction models vote (Mun, Kwak & Huh, 2021). One study predicted hourly time-use in two educational buildings in Northern Central Florida by comparing data using three ML models: 1) Regression Trees (RT), 2) Support Vector Regression (SVR), and 3) Random Forest (RF). The outcome shows that the RF model was better in the first building by 14.5% RT and SVR. The second building had 25-5.5% RT and SVR (Wang et al. 2018). The comparison between regression algorithms is the most common approach used to choose the most suitable ML for particular datasets and project goals. Some researchers compare a greater number of models by testing all ten regression models in the default Scikit-learn library (Huchuk, Sanner & O'Brien 2019).

2.6. Summary

Saudi Arabia is a country with a strong culture. Culture, as mentioned previously, is a social factor that influences occupant behaviour. This has an impact on lifestyles, building, and, ultimately, energy use. Despite this, there are as yet no occupant behaviour models to predict occupant behaviour from this cultural perspective. This study aims to collect data to fill the lack of information regarding how cultural behaviour impacts energy use. The next section reviews previous studies on Saudi houses and social factors that influence design and occupant choices. The research covers old traditional mud houses to new model houses. It also covers social and behavioural changes throughout Saudi history. This will give a deeper insight into the performance gap problem and offer suggestions to address these gaps through this research.

This research aims to improve energy performance gaps by filling the lack of information in the literature related to culturally influenced occupant behaviour. This study addresses this gap by considering the impact of culture on occupant behaviour as a key factor in energy simulation inputs. The study focuses on single-family houses in Riyadh, Saudi Arabia, and investigates three areas: 1) energy performance evaluation, 2) occupant behaviour, and 3) culture. Fig. 2.14 shows a diagram that includes an explanation of these three areas and how they overlap. The zones highlighted in four different colours (green, blue, yellow, and red) in the diagram are covered in this research. The red colour in the middle of the diagram (where all three areas intersect) highlights the problem this research attempts to solve. To understand and solve the lack of information in this area, the research covers:

- (A: blue) Occupant behaviour / Energy performance.
- (B: green) Culture / Occupant behaviour.
- (C: yellow) Culture / Energy performance.

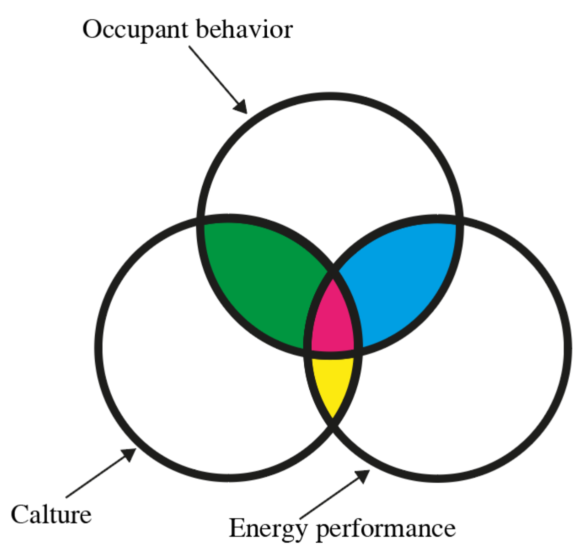


Fig. 2.17. Diagram of the areas this research covers.

A) Occupant behaviour / Energy performance is one of the five main inputs for building energy performance. Chapter 2 covered the previous data collection methods to understand occupant behaviour patterns and analysis approaches used for different occupant behaviour data types. The last section reviewed the most effective methodologies and technologies used to improve occupant behaviour inputs. Results show that using a survey data methodology is suitable to understand patterns of cultural behaviour. The ML model has been tested to train and validate data to predict behaviour and proved that it provides highly accurate occupant behaviour predictions for energy simulations. Researchers found using Regression, SVR, and Random Forest gave highly accurate probability predictions. Furthermore, reviews suggested that researchers test all these algorithms using a comparison approach, explore advantages and disadvantages, and then determine the best algorithm to address the model gap. Findings from this investigation will improve occupant behaviour input from the perspective of culture.

B) Culture / Occupant behaviour. This part covers the impact of culture on occupant behaviour in terms of space usage. This helps tackle behaviour related to culture that is not considered in occupant behaviour inputs in energy simulations for houses in Saudi Arabia. Studies throughout Saudi history show that occupants prefer high levels of privacy, which is driven by their culture. To understand the impact of privacy and how it is reflected in home design and circulation, the author uses a privacy diagram method created by Jamel Akbar. The Akbar diagram ranks spaces into five privacy levels, from public to very private. He also categorised occupants into six categories, from guests to family members. Lastly, Akbar defined privacy levels based on occupants who were allowed to enter different rooms. This study shows that Saudi houses are divided into three zones (male guest zone, female guest zone, and family zone). The separation of zones is determined by the privacy levels occupants prefer in their homes. This area has not been well covered in previous studies. Therefore, the pilot study chapter covers further investigation and analyses using case study experiments to understand how culture impacts occupant behaviour and energy use of single-family houses in Riyadh, Saudi Arabia.

C) Culture / Energy performance. This section aims to understand how culture impacts energy performance. This is achieved through a historical review of the structural evolution of homes in Saudi Arabia. This area covers the impact of culture on the layout of cities and house designs. Privacy is the main element shaped by Saudi culture. The investigation shows that this cultural behaviour is reflected in their houses.

Saudi people are proud of their social norms, such as inviting guests, helping travellers, and showing generosity. In the country's history, when there were no hotels or restaurants, people would invite travellers into their homes to stay and have food. This hospitality is reflected in house layouts. Even if they are rarely used, guest zones are a major space in Saudi houses. However, guest zones and the areas previously discussed have not yet been investigated in depth. Therefore, this research answers the main question: Does the influence of Islamic culture play a significant role in the accuracy of simulations related to energy performance gaps?

3. Methodology

This study aims to understand Riyadh citizens' energy use and culturally influenced occupant behaviour and use these data to improve energy performance predictability. The research is intended to make a valuable contribution to the field of energy conservation, which is a growing global concern. Energy conservation concerns are especially significant in Saudi Arabia, where plans for stricter energy use regulations are underway.

Chapter 2 reviewed the literature and discussed the critical importance of privacy in Saudi Islamic culture and its impact on house design. Saudi people divide their homes into separate zones: a male guest zone, a female guest zone, and a family zone dedicated to home occupants. Saudi occupant behaviour is yet to be researched in terms of energy performance. Occupant behaviour affects home energy use and is a significant contributor to energy performance gaps. This chapter is divided into four main sections:

- The first section explains the methodological approach and underlying processes used to define and address the research problem.
- The second section describes data collection methods and tools used to develop a time-use prediction model.
- The third section discusses data analysis approaches.
- The fourth section reviews the methodology used to evaluate the impact of culturally influenced behaviour in energy simulation outcomes.

Low-accuracy modelling of occupant behaviour inputs can lead to a large gap in energy performance predictions. To improve predictions, researchers must consider and understand the interplay of culture and occupant behaviour. This study explores whether Saudi Islamic culture plays a significant role in the accuracy of energy performance predictions. To answer this main question, three sub-questions were devised. The first sub-question is “How do cultural aspects affect the energy consumption for the houses in Riyadh?” The second sub-question is “Without considering cultural aspects, what is the gap between the energy simulation prediction and the real energy use for houses in Riyadh, Saudi Arabia?” The third sub-question is “To what extent does people’s behaviour in the energy simulation close the gap between predicted energy use and real energy use for houses in Riyadh, Saudi Arabia?” Comparing the answers to the second and third sub-questions with the real energy use will illuminate the benefits of integrating cultural considerations into energy simulation predictions.

For this study, a mixed-methods approach was used. A combination of both quantitative and qualitative methods allows for an in-depth exploration and understanding of cultural behaviour. The data were collected using ten case studies, interviews, and surveys (ethics approval number 018753; approval letter included in the appendix). Fig. 3.1 illustrates the structure of the methodology.

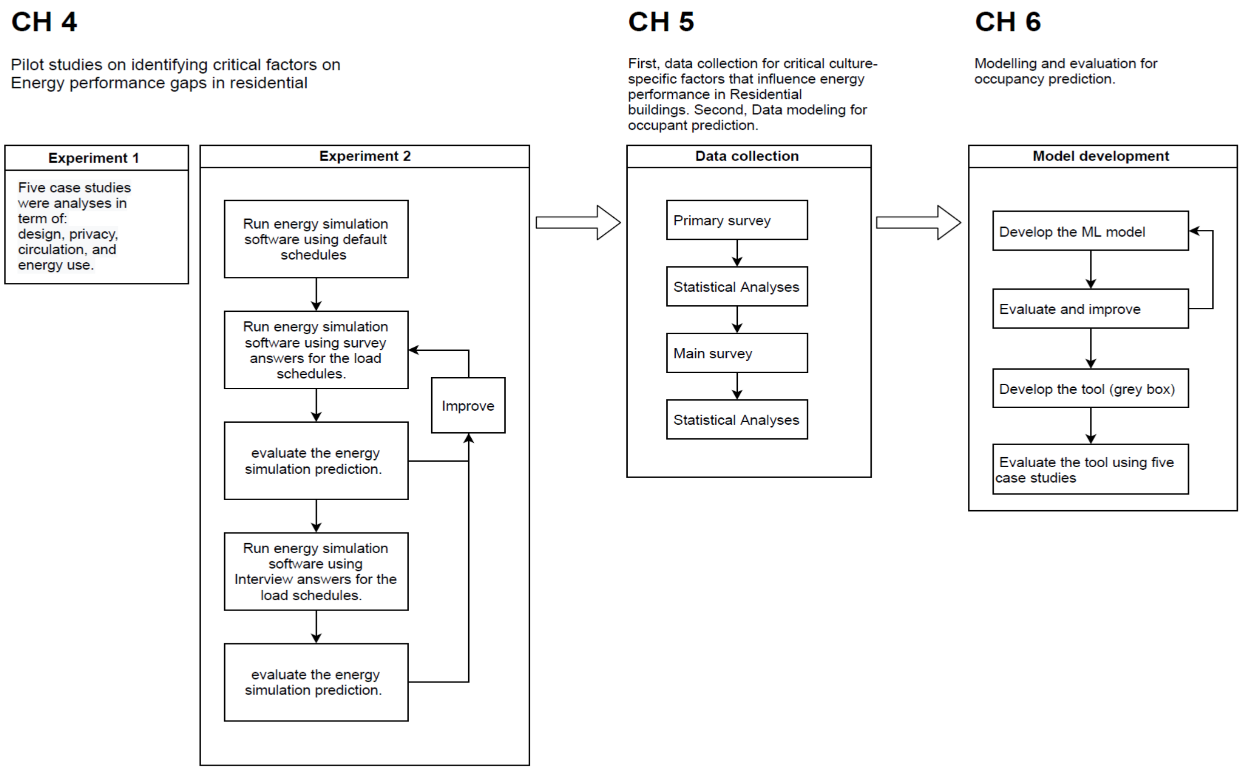


Fig. 3.1. The framework for the methodology.

As shown in Fig. 3.1, Chapter 4 discusses two experiments that served as pilot studies. As shown in Chapter 2, the literature provides no clear understanding of the impact of the Saudi culture on occupant behaviour or the impact of this behaviour on Saudi house design and energy performance. Therefore, the two experiments designed for this study investigate this area further. The first experiment includes an analysis of five case studies. Each case represents a different decade of home design. The analyses examine house layouts, privacy diagram analyses, circulation, and energy simulation predictions. The second experiment uses a case study to perform energy simulations. The experiment's goal is to test the impact of culture on occupant behaviour to close the gap between energy simulation predictions and real energy use. In addition, the case study was used to improve survey questions before publishing.

Chapter 5 discusses the implementation, data collection, and analysis of two surveys. The primary survey intends to understand and define critical culture-specific factors that influence energy performances in residential buildings. The second survey gathered data used to train and validate the occupancy prediction model.

Chapter 6 reports on the occupancy prediction tool development. The researcher discusses the framework used to develop the time-use prediction tool for targeted zones and the tool integration with the energy simulation process. The framework used survey data to generate a statistical analysis model using the machine learning method. The new time-use prediction for targeted zones was evaluated in an experimental study with five case studies. The grey box approach was used, integrating the ML model (time-use prediction) in the modelling process to improve the energy prediction outcome. This approach was applied to the case studies and compared with the real electricity bill to evaluate the impact of culturally influenced behaviour on energy simulation outcomes.

3.1. Data collection

This section discusses data collection and is divided into three sections. The first section discusses qualitative data collection and analysis to understand the impact of culture on occupant behaviour in Saudi homes. The second section covers data collection from the survey that developed the occupancy time-use prediction tool for guest rooms where culture greatly influenced occupant behaviour and energy use. The third section reports on five case studies used to evaluate the methodology informing the development of the time-use prediction tool.

3.1.1. Data collection for investigation

First experiment: Data collection was based on case study comparisons between five plans of Riyadh houses built from pre-1960 to the present. Data was collected from March 1–31, 2018 and gathered from two sources: Alnaim Architecture and the Riyadh municipality. The plans were collected in paper copy form then redrawn using Revit. The first case was a traditional house built using mud. While there was no exact date of the build, the house was built before concrete was commonly used in Riyadh in the 1970s (Mubarak & King 2007). The second case was a house built in 1978, the third case in 1990, the fourth in 2004, and the fifth in 2016. The researcher used observation data that enables different analyses, which can strengthen outcomes (Flick, 2018). This method of analysis will be discussed in the data analysis approaches section.

Second experiment: The second data collection was a case study. The collection used a qualitative method to improve survey questions examining the impact of culture on occupant behaviour and energy simulation outcomes. This case study was based on a standard house in Riyadh built in 2005. The house (a villa) had a basement, ground floor, first floor and second floor. Six occupants inhabited the villa. The total floor area was 500 m^2; it had double-glazed windows and 30-cm thick concrete exterior walls with 8 cm of insulation. Fig. 3.2 shows the layout of the basement where friends gather and the ground floor where guests stay. The ground floor includes male, female and family zones. The grey area highlights the male guest zone. The third plan shows the first floor, which contains all of the bedrooms. The fourth plan shows the second floor, consisting of storage and service rooms. The tools used for the experiment included Revit to draw the plans, SketchUp to design the geometry, and OpenStudio to model the case study and run the energy simulation. The landlord provided the detailed input.

The six occupants of the case study participated in one-hour interviews to understand their behaviour and improve survey questions. The focus was on four main sections. The first part of the interview explored socio-demographic questions, building and occupant details, and unique behaviour questions related to energy use. The second part explored each occupant's lifestyle and each room use in detail. The data obtained was fed into time schedules. The third part included collecting other information used to calibrate questionnaire responses and identify any necessary design readjustments. The fourth part gathered data on occupants' opinions about survey questions and whether other questions should be included to improve the room use predictions. Approval from the householder to participate in this experiment was secured before May 2018. However, data collected from interviews and additional information gathered from occupants was repeatedly performed during the survey collection period.

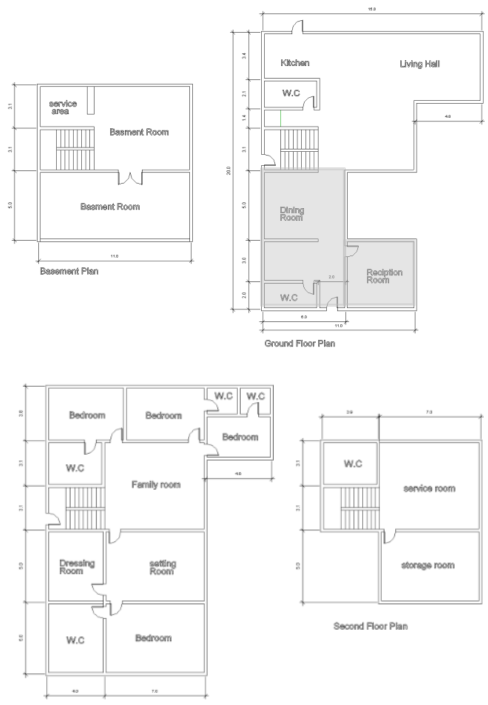


Fig. 3.2. Layout of the house in Riyadh used in the second pilot study.

3.1.2. Data collection for the occupancy time-use prediction tool

Primary survey: The time-use prediction model was informed by data collected from two surveys. The primary survey was designed after analysing the first experiment (comparing the five case studies). The goal of the survey was to gain in-depth insight into how spaces of Saudi homes are divided, how time is spent in those homes, and how different rooms are used (Table 3.1). The survey was multiple choice and consisted of twenty-three questions. Data were collected in two ways. First, the researcher attended a conference in the Architecture School in King Saudi University, where eligible participants were asked to fill in the survey via an iPad. Eligibility for study participation required individuals to be living in a house in Riyadh and married. In addition to the data collected via iPad, fifty survey responses were collected person to person. The researcher also gathered survey data via digital means (emails, social media, etc.). He advertised the study online and instructed eligible individuals to participate using Google Survey. Four hundred participants completed the survey online. Table 3.1 shows the four main sections of the survey. The last section of the survey included questions to respond to the second sub-question for this research: Does considering the Islamic culture in the energy simulation close the energy performance gap?

Case study findings show there are rooms in homes only used occasionally but nevertheless necessary because of cultural norms related to privacy. Therefore, the time-use question was divided into two sections: 1) daily time-use questions for the male guest zone, female guest zone, and living hall; and 2) hourly time-use for the kitchen, guest room, dining room, living hall, family room, bedroom, friends' room, and the outside of the house.

Table 3.1

Hierarchical structure survey questions.

Socio-demographic questions	
	What is your gender?
2.	What is your age?
3.	What is the occupation of the homeowner?
4.	Which part of Saudi Arabia do you live in?
Building and occupant details	
5.	How many men live in the home including you?
6.	How many females live in the home including you?
7.	How old the person who responsible for the family?
8.	Where do you live?
9.	When was your home built?
10.	What is the total built area of the ground floor?
11.	How many rooms are there in your home for visitors?
12.	Which room do you use for the daily family gathering?
13.	Do you only use the dining room for guests?
14.	Do you have a room in your home for gatherings with friends?
Unique behaviour related to energy use	
15.	If you travel every year for more than two weeks, please identify which month
16.	Do you use the HVAC system for heating in the winter? Tick on the rooms you use the HVAC to heat it
17.	How much approximately is your monthly electricity bill in the summer? "Riyal
18.	How much approximately is your monthly electricity bill in the winter? "Riyal"
Time use	
19.	How many times a year do you use the male guest zone?
20.	How many times a year do you use the female guest zone?
21.	How many times a year do you use the living hall?
22.	Tick the rooms that you or any of your family used during the time below (you can choose more than one or keep it empty if you do not use it) During the weekdays.

	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11
	AM												PM											
Kitchen																								
Guest room																								
Dining room																								
Living room																								
Family room																								
Bedroom																								
Friends room																								
No peaple in the house																								

23. During the weekend

	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11
	AM												PM											
Kitchen																								
Guest room																								
Dining room																								
Living room																								
Family room																								
Bedroom																								
Friends room																								
No peaple in the house																								

Main survey: The second main survey was in an open answer format, so participants were able to enter any number. Numerical responses were used to allow the model to train and validate the model accurately. The survey had 11 questions: 7 of them were the parameters (X) and 4 of them were the time-use for targeted zones (Y). Eight questions were selected from the primary survey and 3 questions were added. The first question added to the second survey was informed from the interview in the previous experiment, as the researcher learned that the friend gathering room might have a unique daily use schedule influenced by Islamic culture. The question "(Y4):

How many days do you use the friends' room in one year?" was added to the survey for further investigation to explore whether there was unique time-use behaviour that should be considered.

In addition, before publishing the full open answer survey (Table 3.2), one question was posed to 25 architects: "What do you think are the most important elements to predict the time-use of the guest zones and the living hall other than these five?' Based on architect responses, two additional questions were added. Questions were related to house location (i.e., north, south, east, west, or middle of Riyadh) and whether the occupant had siblings (as well as birth order).

The first added question asked about the location of the home in Riyadh. As mentioned in the second chapter, when Doxiadis designed Riyadh, he divided the city into these five main areas. In addition, he divided areas based on people's financial situation, which was determined by land size and street length. Areas for low-income citizens consisted of small land sizes and street lengths. The interviewed architects believed that a house's location in Riyadh could reflect on the number of people living in the house compared to the size of the house and the use of the guest zones.

The second added question asked whether the householder had siblings (particularly brothers), and if so, the birth order of the respondent (i.e., whether the respondent was the oldest, youngest, or middle sibling). This question was added to investigate whether the eldest brothers used the guest zone more than younger siblings or householders with no brothers. These additional questions were used to help evaluate the parameter impact on time-use predictions. After adjustments, the survey was published using Google Forms. Participants were recruited via Twitter, WhatsApp and email. A total of 1,231 responses were recorded. The author collected the main survey in May 2019.

Table 3.2
Structure of the survey questions.

Parameters	
X1	Location of the house (north, south, east, west, or in the middle of the city).
X2	Does the householder have brothers and, if yes, is he the oldest, middle, or youngest?
X3	Age of the householder?
X4	When was the house built?
X5	How many male members live in the house?
X6	How many female members live in the house?
X7	The size of the land
The predicted time use	
Y1	How many days do you use the male guest zone in one year?
Y2	How many days do you use the female guest zone in one year?
Y3	How many days do you use the living hall in one year?
Y4	How many days do you use the friends' room in one year? (This was a test/calibration question).

3.1.3. Data collection for the model evaluation

An experiment was conducted to evaluate the impact of survey data on energy performance gaps. Five case studies were utilised, each representing one of the main locations in Riyadh (north, south, east, west, and central). The case studies were single-family houses and had parameter inputs close to the mean value of the parameter (Table 3.3). All data were collected from householders (plans, building details, occupant behaviour, survey questions) and electricity bills. Data collection for the model evaluation experiment started in May 2019 and ended in February 2020.

Table 3.3
The five parameters for each case study.

Location	Age of the holder (Years old)	Number of the male members (People)	Number of the female members (People)	Land size (m^2)
North	43	3	2	250
East	55	5	4	754
South	56	4	4	528
West	32	3	1	360
Centre	61	3	3	500

3.2. Methods of analysis

A mixed data analysis approach was used to address the research questions. Analyses of different types of data (quantitative or qualitative) have different approaches based on each kind of data type and goal.

3.2.1. Qualitative analyses

This section discusses the qualitative analysis approach. This data analysis type allows for a deeper investigation of the impact of Saudi culture on occupant behaviour and energy performance. Qualitative analyses were used to answer the first sub-question of the research ("How does the Saudi culture impact on the occupant behaviour?"). The experiments consisted of two pilot studies (covered in Chapter 4). A brief overview of each pilot is provided below.

First pilot study: The first pilot study used five case studies to understand the impact of culture on occupant behaviour. Analyses were based on observations and their relationship to the social context. Analyses included the design layout of the five case studies, privacy and circulation, and energy use. By comparing the five case studies and culturally influenced occupant behaviour, we can gain greater insight into how house designs have changed over time and gauge occupant behaviour patterns related to the use and separation of house spaces to maintain privacy. The tools used for the pilot studies included Revit for space design, SketchUp for geometrical layout, OpenStudio for energy use, and Space Syntax for visibility analyses. The Akbar (1980) diagram method was used to define privacy levels for each of the five case studies.

Second pilot study: The second pilot study experiment tested the impact of culture on occupant behaviour for energy simulation predictions. This was achieved by comparing outcomes for three experiment stages, which included defining the problem, testing the methodology, and improving the survey. The first stage of the experiment included performing normal energy simulation analyses for case studies using OpenStudio software. In this experiment, all inputs (geometry and climate, system and building material) were provided except for load schedules, which were replaced by default schedules for a residential building (as provided by the software). The energy prediction outcome from the software was compared with real energy use from electricity bills provided by the house owner. The second stage was to test and improve the impact of survey answers on energy simulation outcomes. Here, the stage 1 method was repeated using survey data as the time-use schedules for targeted spaces instead of software defaults schedules. The third stage involved improving the survey questions which was achieved by interviewing occupants of case study houses. The outcome of this stage was compared with the outcome of the previous iteration of the model. The energy simulation software provided standard fixed time-use schedules for different space types of residential buildings. Comparing the time-use survey data outcome with the default schedules will demonstrate if culture impacts time-use schedules. The approach was iterated several times (simulation, occupant interview, survey improvement, reflection on answers on the simulation input, etc.) until no more improvements to the outcome were possible using study data and resources.

3.2.2. *Quantitative analyses*

Qualitative analyses were conducted with two survey types (multiple choice and open answers survey designs). The first survey was a primary survey, made to understand critical cultural factors that influence energy performance and definitions of the most effective variables on time-use prediction. The second survey was created to understand and prepare data used to train and validate the time-use prediction model.

Primary survey analyses: Data was collected from 395 participants. Data analysis consisted of three stages: 1) cleaning and processing the data 2) reviewing the distribution, mean and standard deviation for variables and 3) statistical analyses using chi-square tests to identify correlations between variables. The third stage was conducted to remove variables that would have a low impact on time-use predictions and also reduce the number of parameters.

Second survey analyses: The second survey (open answers design) data was used to train and validate the time-use model. Data from 1,231 participants were manually cleaned and processed, including removing outliers and identifying missing data. After cleaning and removing missing data, data from 861 participants remained. The last step consisted of statistical analysis, including correlations, curve fitting, and bootstrapping. After this stage, the data were ready to use to develop the occupancy time-use prediction tool.

3.3. Evaluating the impact of the model in the energy simulation

This section provides an answer to the second research sub-question, "How much does consider cultural behaviour improve energy simulation outcomes for houses in Riyadh?" The response to this question included three parts. First, the researcher used a statistical model to predict time-use for targeted zones. Second, he implemented the model in an energy simulation process. Third, he evaluated the process. The following tools were used for the evaluation studies.

SPSS is a statistical analysis software delivered in a user-friendly interface with a set of features that allows the user to extract actionable insights from data. The software is very commonly used by researchers, especially in the social sciences, to analyse survey data and mine test data (SPSS Statistics, 2021).

OpenStudio is a graphical interface application. It is an open source (LGPL) collection of tools to run building energy modelling. The model includes input such as envelope, loads, schedules and HVAC. The output allows browsing, plotting and comparing the data, especially time series (OpenStudio, 2021).

Energyplus is a whole building energy simulation program used by architects, engineers, researchers, and users who have average knowledge of the energy industry. The program was used to model the energy consumption for a building. The program is DOE's open source and supports many applications such as OpenStudio (EnergyPlus, 2021).

Jupyter Notebook is a server-client application that can create and share live codes, equations, visualizations, and narrative text via a web browser. It is a graphical interface application that nonprogrammers can use for data cleaning and transformation, numerical simulation, statistical modelling, data visualization, and machine learning (Project Jupyter, 2021).

PyCharm is a dedicated Python Integrated Development Environment. The application allows the user to use python with access to many tools to develop an efficient environment for web development and data science development (JetBrains, 2021).

3.3.1. Statistical analyses model development

The researchers used a statistical model approach to predict cultural behaviour patterns for the zones with unique culture time-use behaviour. Data analyses suggest there is unique culturally influenced time-use behaviour in male guest zones, female guest zones, and living halls for the single-family homes in Riyadh. The researcher used machine learning modelling as a tool to train and validate the data to predict occupant behaviour patterns. Eighteen regression algorithms (OLS, Ridge, Lasso, ElasticNet, LARS, LassoLars, OMP, Bayes, GLR, SGD, PAR, Poly, KNN, Decision Tree, RF, SVR, XGB, and MLP) were compared using the scikit-learn library (Scikit-Learn: Machine Learning in Python n.d.). A cross-validation approach was used to compare performance between prediction models. Cross-validation is one of the most common metrics used to validate and compare time series prediction models (Hyndman & Koehler 2006). 80% of the data was utilised for training and 20% for validation. Algorithm performance outcomes were compared using the mean absolute error (MAE).

Statistical analyses model evaluation: First, models with significant-high error predictions (compared to the rest of the models) were eliminated to identify the most suited ML algorithm of the eighteen algorithms. The researcher then set the maximum error range for the model to 0.7. After setting this threshold, 8 ML algorithms remained. The researcher then compared algorithm features that could help improve the model prediction, especially those which would most significantly contribute to understanding time-use behaviour for the selected rooms (male guest zone, female guest zone, and living hall).

In previous evaluation experiments, different models were developed to predict the time-use for each zone separately rather than using one model to predict the three zones together. Database inputs (location, age of the householder, number of male members, number of female members, and land size) have different impact values on predictions. To illustrate this, another experiment was conducted to compare input variable performance in each model. The researcher used the feature importance model from the scikit-learn library (Géron 2019), which is useful for non-linear estimators. The procedure breaks the relationship between the feature and the target, so the model score is indicative of how much the model depends on each feature.

3.3.2. Time-use prediction tool development

The second step to evaluate the impact of cultural behaviour in energy simulation outcomes was to integrate the statistical model in an energy simulation process. The grey box approach was used to achieve this goal. The grey box approach integrated the ML model (black box) with energy simulation software (white box) and used the ML model as one of the parameter inputs (Fig. 3.3). This research will use the grey box approach to replace the default time-use schedules with new predicted time-use schedules for male guest zones, female guest zones, and living halls. The tool uses fixed hourly schedules generated from the primary survey and predicted daily time-use resolutions for the whole year from the main survey.

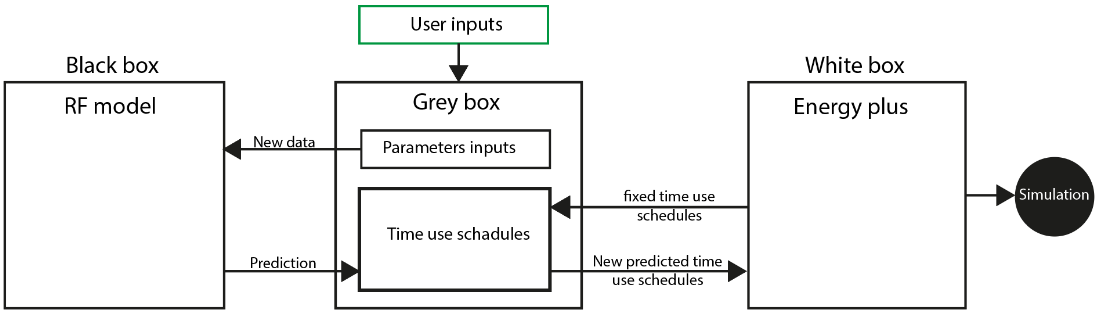


Fig. 3.3. Framework of the approach.

3.3.3. Evaluation methodology

This section includes information about the researcher's process to evaluate the methodology used to improve the energy performance gap in houses in Riyadh. The evaluation included an experiment using case studies. Five case studies were used, representing the main areas in Riyadh (North, south, east, west, middle of Riyadh). Table 3.4 shows the demographic information for the case studies. The evaluation experiment process went through three main stages. First, the researcher ran an energy simulation for each case study using default time-use schedules provided by the software. Second, he ran an additional simulation using the grey box to improve time-use schedules for targeted zones. Lastly, he compared outcomes from the first and second simulations with actual electricity bills provided by the householder. The researcher’s objective was to leverage ML technology to estimate the impact of culturally influenced occupant behaviour on energy simulation performance.

Table 3.4
Demographic information for the case studies.

	Case study # 1	Case study # 2	Case study # 3	Case study # 4	Case study # 5
Location	North	East	South	West	Middle
Age of the householder	43 years	55 years	56 years	32 years	61 years
Number of male members	3	5	4	3	3
Number of female members	2	4	4	1	3
Land size	250 m^2	754m^2	528m^2	360m^2	500m^2
Façade materials and colour	concrete block covered with façade stone (light brown)	concrete block covered with façade paint (beige)	concrete block covered with façade paint (brown)	concrete block covered with façade stone (brown)	concrete block covered with façade stone (beige)
Orientation	Male entrance facing south; Female entrance facing south	Male entrance facing west; Female entrance facing south	Male entrance facing north; Female entrance facing east	Male entrance facing west; Female entrance facing north	Male entrance facing south; Female entrance facing east
Aspect ratio between long and short sides of the structures	18:8	19:16	18:12	22:8	20:11

4. Pilot studies

The goal of this chapter is to report on pilot studies performed in the research. Pilot studies were conducted to develop the foundation of the survey questions. Study survey data were used to train and validate the data for the occupancy time-use prediction tool. This chapter describes two pilot studies. The first pilot study examines five case studies of houses in Riyadh from 1960 to 2020, as assessed based on design, privacy, visibility and energy. The second pilot study tests and improves the survey questions prior to data collection. In the second pilot study, data from one house was used to run the simulation with three different inputs. The outcome was then compared with the real energy use. The experiment first examined the differences between energy use predictions (using the default schedule) with real energy use. The default time-use schedules were then replaced with the survey answers. Following this, an interview with the house occupant was conducted to improve the survey questions further.

4.1. First pilot study

4.1.1. Method

At the first stage, five representative case studies were selected—each case study represented houses built within the span of a decade. The first is a traditional house built before oil was discovered in Saudi Arabia; the second house was built in 1978; the third house in 1990, the fourth in 2004; and the fifth in 2016. The pilot study collected data on aspects of design, privacy, space, vision and energy use.

4.1.2. Typology of the selected case study

Design: The house plans were collected from Alnaim Architectures Engineers Urban Planners and the Riyadh Municipality. Alnaim has a large database and is considered one of the biggest architectural offices in Riyadh. The plans for older houses, those built between the 1970s and 1990s, were retrieved from the municipality. The villas had two floors with a total floor area of 250 to 600 m^2. After the acquisition of the data, the plans were re-modelled using Revit. The house plans were highlighted in three different colours to represent the three main targeted zones in Saudi houses. The male guest zone was highlighted in green, the female zone in yellow, and the family zone in blue.

Privacy: The spaces were divided into multiple levels ranging from public to very private. Akbar used the same method in his 1997 research when he defined the privacy level of traditional and modern houses at that time (Akbar 1980). The main principles used to identify privacy levels were based on the accessibility of each space and the visibility between all spaces. Users were divided into six categories: 1) male inhabitant, 2) male relative, 3) male visitor, 4) female inhabitant, 5) female relative, and 6) female visitor.

Space and Vision: It is first necessary to study the spaces required and their line of sight (vision) within the residences. In the past, Saudi culture encouraged substantial privacy, for which the architectural designs of houses included spaces that hindered guests' views. Space Syntax software was used to simulate the relationship between spaces and visibility. The concept of the Space Syntax software is based on the configurational theory of space and attempts to decode spatial formations and how this impacts human activity in a city or building (Dursun 2007). The simulation was first executed in Revit, where the data of the villas were uploaded. The program mimics theoretical water flow in the space, where Space Syntax corresponds to the plan, which is divided into 100,000 points. The walls are demonstrated as objects that would stop the water flow (visibility volume) on the plan. The program defined the volume of visibility by the number of points directly connected to another point in the plan. Every small point is a square shape, and the size of each point is based on the number of points the user set the grid to. The greater number of points touch another point from any side with no objects (wall) blocking the sequence of the points, which indicates higher visibility. Each floor plan in the case studies was divided into 100,000 points (default number of points provided by the software) (Space Syntax | Thriving Life in Buildings & Urban Places n.d.). For example, if the size of the case study is 400 m² and the user chooses 100,000 points, then the size for each point is 400/100000 = 0.004 m².

Energy Use: OpenStudio software was used to model and predict energy consumption for each of the five buildings. The simulation analysis focused on electricity use. One of the advantages of OpenStudio is that the software is open source for EnergyPlus (LGPL) (GNU Lesser General Public License v3.0 - GNU Project - Free Software Foundation n.d.). The program supports the import of gbXML and IFC for geometric design. It further includes an envelope, schedules, loads, and HVAC usage. The program gives the user the ability to browse, plot, and compare the output data and time series.

4.1.3. Results of pilot study 1

First case study—traditional house: Before oil was discovered in Saudi Arabia, resources in the country were low. This was reflected in the lifestyle of the Saudi people, as reviewed in the second chapter. For example, mud was the main material for structure and façade. Because of this, one of the houses included in the pilot study was a traditional house where electricity was not used. It represents the behaviour, culture, and resources of the time, which can be used to compare with newer buildings. Dwellings in Riyadh at the time were built with no guidelines or codes. The focal house measured (225 m^2) and was depicted using Revit (Fig. 4.1).

A significant difference from modern houses was that the traditional house had a courtyard in the middle of the house, and all windows faced the courtyard (except the men's guest room which faced outside). It is to be noted that in big cities in Saudi Arabia, houses used to utilise the courtyard, but in villages and small towns they did not use courtyards as much. There they typically had a farm and so did not need a courtyard area. The windows faced the farm, not the street, which kept privacy high. The male guest room was separated from the rest of the house. This separation between the male guest room and the rest of the house reflected the occupants' behaviour. The occupants preferred high privacy for their family in the house. This pattern was likely for houses in Riyadh City, as well as the surrounding areas. The total size of Riyadh is 404,240 km².

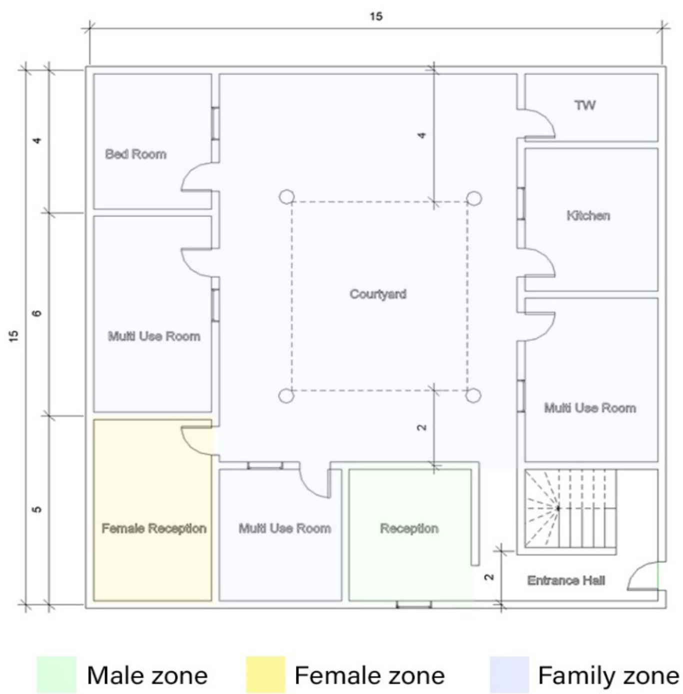


Fig. 4.1. Plan for a traditional house in Riyadh.

Since electricity was not available, windows played an essential role in improving the indoor environment. To have windows and still maintain privacy, windows were high, starting at 180 cm above the floor. Since the city is in the desert, external views were not the main consideration. Fig. 4.2 illustrates the privacy level from the public (1) to the very private (5). In such traditional houses, male guests were not allowed to pass level two. A female guest had a greater ability to move inside the house. The diagram shows that access to all of the rooms except the male guest room was from the courtyard. There was an overlap between family members and female guest circulation. This shows that there was no problem with integration between the family and female guests. However, the male guests were not allowed to go to the courtyard (level three), which shows that the male guest room had very limited accessibility compared to other spaces in the house.

Fig. 4.2. Privacy level for the traditional house.

This type of traditional house had either one entrance for guests and family or two entrances, one for male guests and one for the family. The male guest room was still separated visually when the house had one entrance. Although the male guest room was the main part of the house, it was visually separated, and there was no overlap in circulation (i.e., one can go to the room without moving through another room) between the male guest and the family members.

Houses with two entrances were often seen for households who frequently invited guests over. Fig. 4.3 is a visibility analysis using Space Syntax software. The bigger number of points means higher visibility in the house. In this case study, the average number of continuous points in the plan was 5,378. The lowest number of connected points in this case study was 406 points (Fig. 4.3, coloured dark blue). The highest number of connectivity points was 10,631 (Fig. 4.3, coloured red) located in the courtyard (in the middle of the house). The standard deviation in this case study was 3,487. The analyses show that the men's guest room circle (R) in Fig. 4.3 had a very low visibility rate (around 600 points).

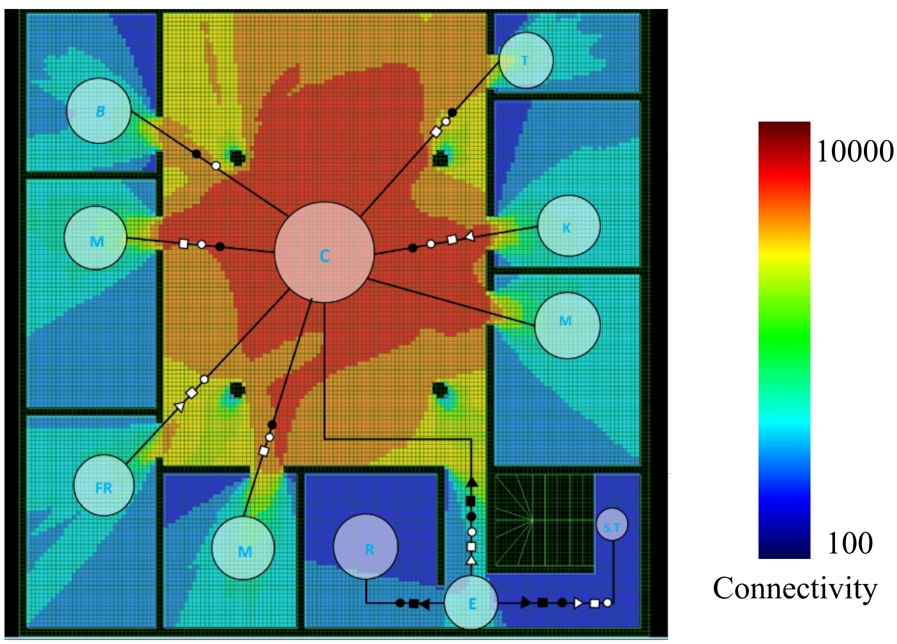


Fig. 4.3. Visibility analyses integrated with privacy level. The colour code indicates the number of points in the house from which a focal point can be viewed. Darker red indicates higher visibility (10,631 points maximum); blue represents lower (100 points maximum).

In this traditional house, electricity was not available, so the house was designed to get as much sunlight and air movement as possible, while respecting the occupants' privacy as influenced by their culture. However, to experiment with the impact of the shape on the electricity use of the building, this case was modelled using OpenStudio, and an energy simulation experiment was run to see how much this house would consume energy if there were electricity. The outcome of the net site energy cost was calculated as follows: 73570 KBTu = 73570000 Btu; 73570000 Btu × 0.000293 (to convert power to kilowatt = 21556 kW; 21556 kW / 12 = 1796 kW; 1796 kW × 0.10 SR (Saudi Riyal) = 179 Riyal/month. The annual energy use per m^2 is 73570 Kbtu / 225 = 326 Kbtu/m^2.

Second case study—house built in 1978: The second case study was a house built in early 1978. The total area in this case study is 225 m^2. This house represents an early version of Riyadh's modern houses. After the discovery of oil, traditional houses were replaced as more housing construction options became available. The main changes were in terms of material and design. Mud as a structural material was replaced by reinforced concrete, and the courtyard was replaced with a living hall (Fig. 4.4). At that time, new building regulations were introduced. This rule meant that the building must leave 20% of the street width as a front yard, and there must be a 2-meter minimum between the house and the border of the property for each neighbour—these new rules were referred to as the "setback technique." The setback technique allowed for more sunlight through the windows. Besides, the front and back yards have more exposure to the sun compared to the courtyard technique (Asfour 2020).

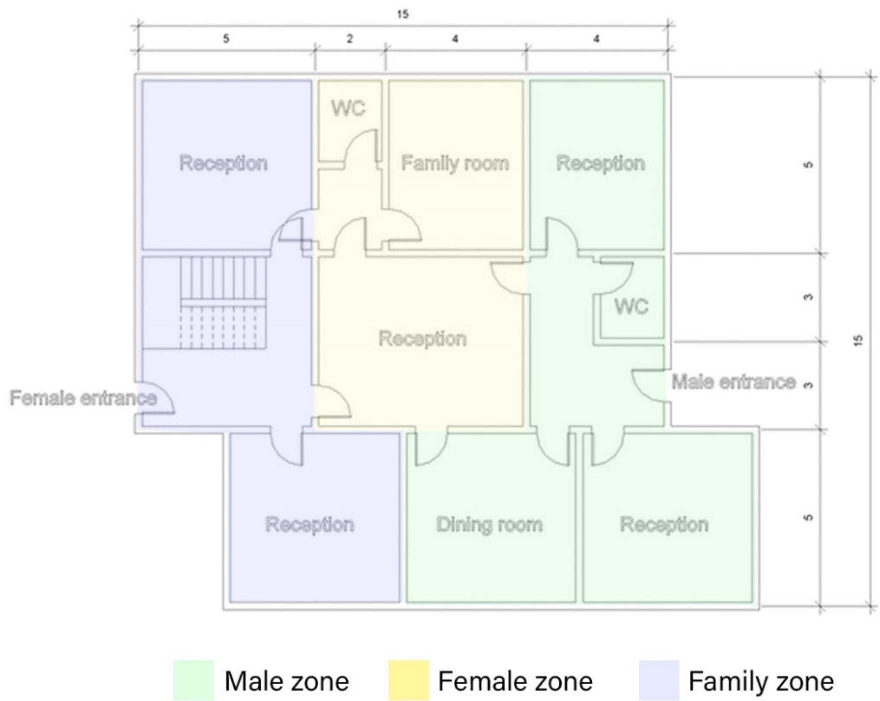


Fig. 4.4. Floor plan for modern house design built in 1978.

The main guest zone in modern houses received more attention than the men's guest zone in traditional houses. In this case, the men's zone was separated. It had two guest rooms and a dining room with a bathroom. The total male guest zone size was 37% of the total space. In the previous case study (traditional), the male guest zone was around 10%. This case had two entrances, one for male guests and the other for female guests and family members (Fig. 4.4). From this period until now, all the Saudi houses had at least two entrances, which were on different sides of the building. The guest zone was separated from the female and family zones. The female zone had two guest rooms, and they shared the same entrance with the family. In total, 70% of the total ground floor could be used by guests.

The internal circulation, in this case, was more complicated than in the traditional house, partly because replacing the courtyard through the setback technique made it harder to maintain the level of privacy that Saudi people desired. Fig. 4.5 shows the privacy level diagram for this case study. It illustrates that male visitors have some access to spaces at the third level of privacy (dining room). Family members were also using level three but for a different room (living hall). Walls separated these rooms of the house and prevented a line of sight between them. An important difference in the modern houses built in the 1970s compared to the earlier traditional houses is that the female zone and the family zone in the modern houses were separated. In traditional houses, there was no circulation separation between the female guests and the family members. To avoid overlap and block the visibility from the male visitors, female visitors, and family members, more walls were built. Compared to the traditional houses, family and female guests share the same zone.

The visibility analysis using Space Syntax (Fig. 4.6) was compared to the traditional house to examine how the modern style impacted the relationship between the spaces. The outcome shows that the average number of connected points was 3,029, the minimum number of connections was 417, the maximum number of points connected was 6,535, and the standard deviation of the points was 963. Compared to the traditional house, the visibility analysis shows fewer connected points in the modern house compared to the traditional house in the first case study—even though the house is the same size. The average number of connected points was 43.68% lower, and the maximum number of connected points in this case study was 6,535 compared to the traditional house with 10,631 points.

Fig. 4.5. Privacy level for a house built in 1978.

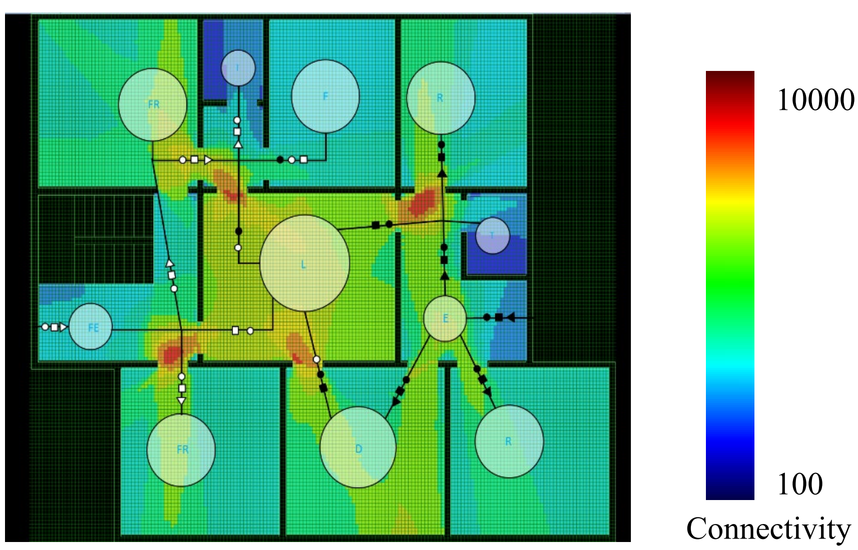


Fig. 4.6. Visibility analyses integrated with privacy level for a house built in 1978.

The last step was to run the energy simulation software (OpenStudio) to examine energy use. The net site energy cost was calculated as follows: 86981 KBTu = 86981000 Btu; 86981000 Btu × 0.000293 (to convert power to kilowatt = 25485.433 kW; 25485.433 kW / 12 = 2124 kW; 2124 kW × 0.10 SR (Saudi Riyal) = 212 Riyal/month. The annual energy use per m^2 is 86981 Kbtu / 225 m^2 = 386 Kbtu/m^2.

Third case study—house built in 1990: The third case study was a house built in 1990. At that time, oil prices were increasing, so the country's economy was thriving. This house was built as a modern house with a setback technique and a total built area of 266 m^2. There is a clear separation between the male guest zone, female guest zone, and family zone (Fig. 4.7). The previous modern case study had two entrances, one for male guests and the other for the family and female guests. In this house, there are three entrances, and the third is for services and family. As shown in this case and the second case, the male guest zone was visibly separated, as was the female zone. The female zone had an entrance hall and two rooms with a bathroom. The dining room had two doors, one open to the family zone (for preparing food and organising the dining service for visitors) and one open to the guest zone.

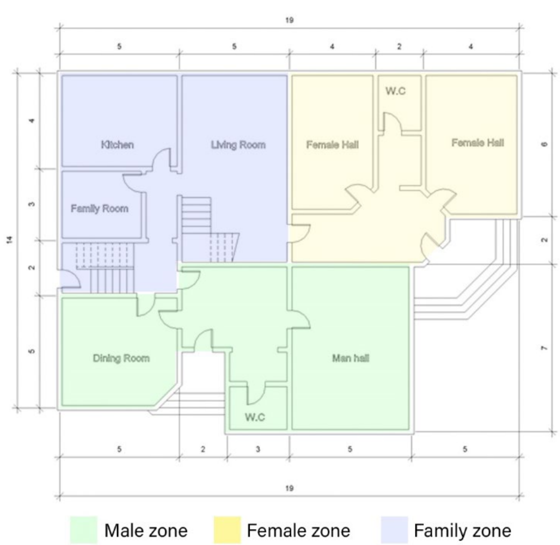


Fig. 4.7. Floor plan for a modern design house built in 1990.

This case study shows that the privacy level in Riyadh in the 1990s was still high. Internal circulation was more complicated to achieve the privacy level desired (Fig. 4.8). For example, any level of privacy in the diagram had a category of people, such as a male relative, which were allowed to go to a specific room and not allowed to go to another room at the same level. This meant that the designer must separate rooms to avoid circulation overlap. On the privacy diagram, level 3 includes three rooms, but each room is for different people (a dining room for all males, a living hall for male and female inhabitants, and a female reception area for all females).

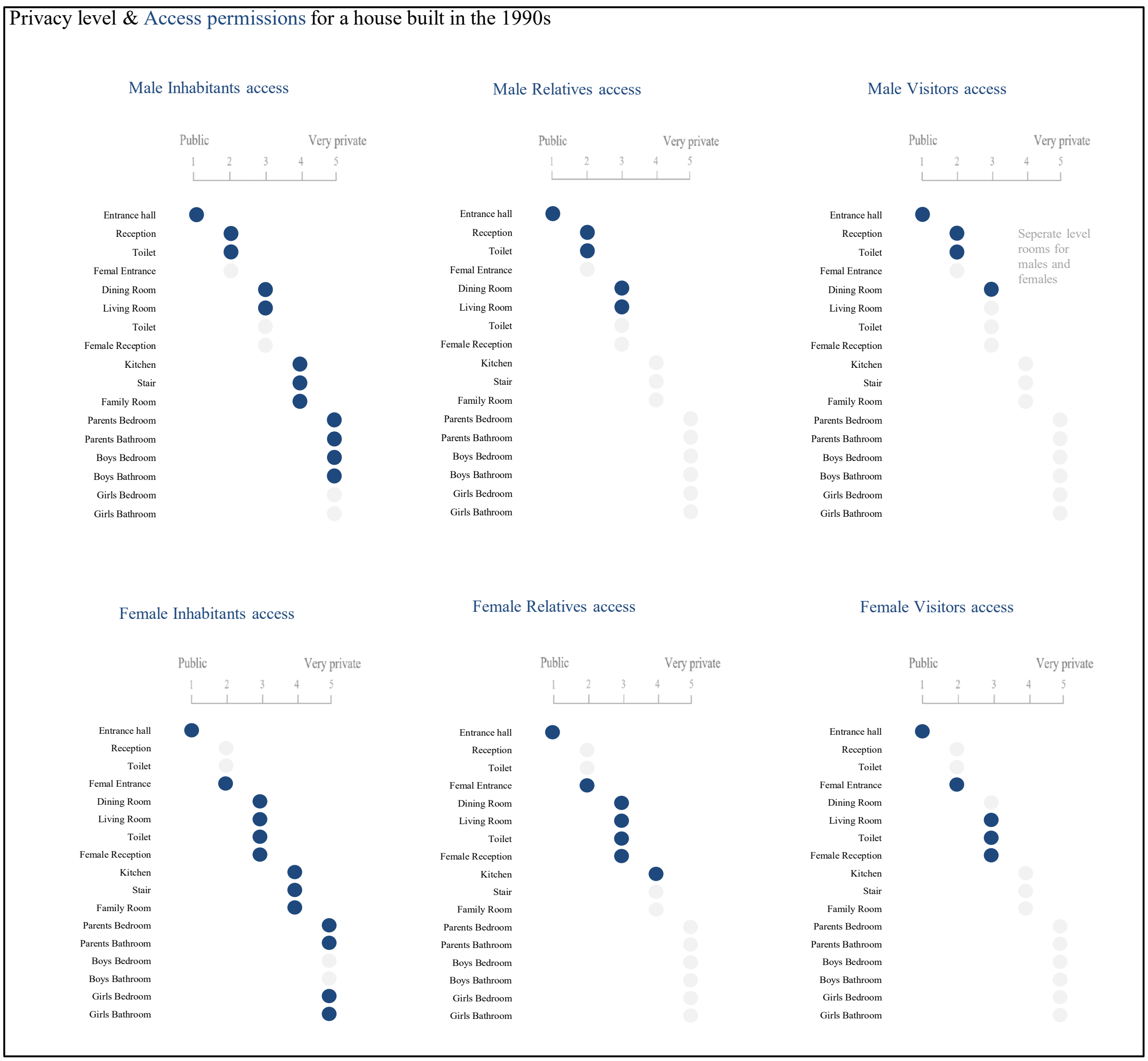


Fig. 4.8. Privacy levels for a house built in the 1990s.

As noted in the diagram, this house had more a greater number of smaller rooms compared to the previous case study, resulting in fewer connecting points. The average number of connecting points was 2,755, the minimum was 419, and the maximum was 5,737 (Fig. 4.9). The main difference was the third entry for the family, involving more complicated internal circulation and lower visibility. Here, the male guest zone size was the most significant, and the female guest and family zones were equal in size. Compared to the house built in 1978, there was not a big difference between the sizes of the male and female guest zones and the living hall. The separation between the rooms using walls for privacy was still clear in this case as the householder's choice.

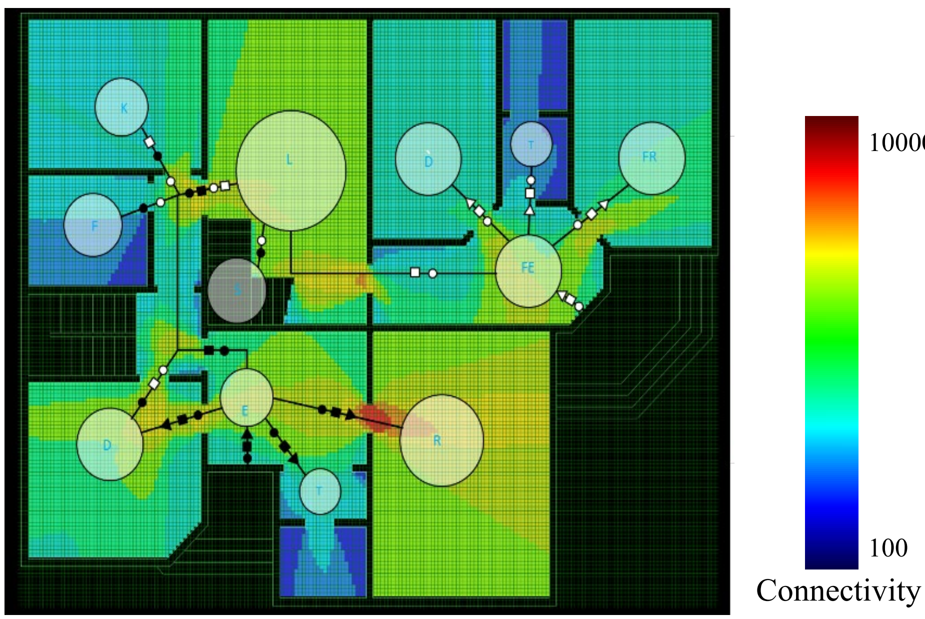


Fig. 4.9. Visibility analyses integrated with privacy level for a house built in 1990.

The monthly energy cost calculation was based on the same method discussed in the previous section: 90820000 Btu × 0.000293 = 26610.26 kW; 26610.26 kW / 12= 2218 kW; 2218 kW × 0.10 SR = 221.8 Riyal/month. The outcome compared with the previous modern house shows similar energy simulation prediction outcomes. House users of that time were not concerned with energy consumption levels due to low energy prices (one kW/ h cost was 0.022 Pound;

monthly electricity cost was 44 pounds). The annual energy use per m^2 is 90820 Kbtu / 266 m^2 = 341 Kbtu/ m^2.

Fourth case study—house built in 2004: The fourth case study was a modern house, built in 2004 with a total floor-built area of 272 m^2. In 1991, Saudi Arabia was involved in the Gulf War with 33 other countries, including the USA and the UK. This war led to changes in oil prices, the overall economy, and people's behaviour. These changes were reflected in Saudi houses. Fig. 4.10 shows fewer rooms in the house built in 2004 compared to previous houses. The female guest zone and the family zone were not separated; the family and female guests shared the same entrance. The male guest zone was still separated and was considered just as important as earlier; it had two guest rooms and an entrance hall with a dining room.

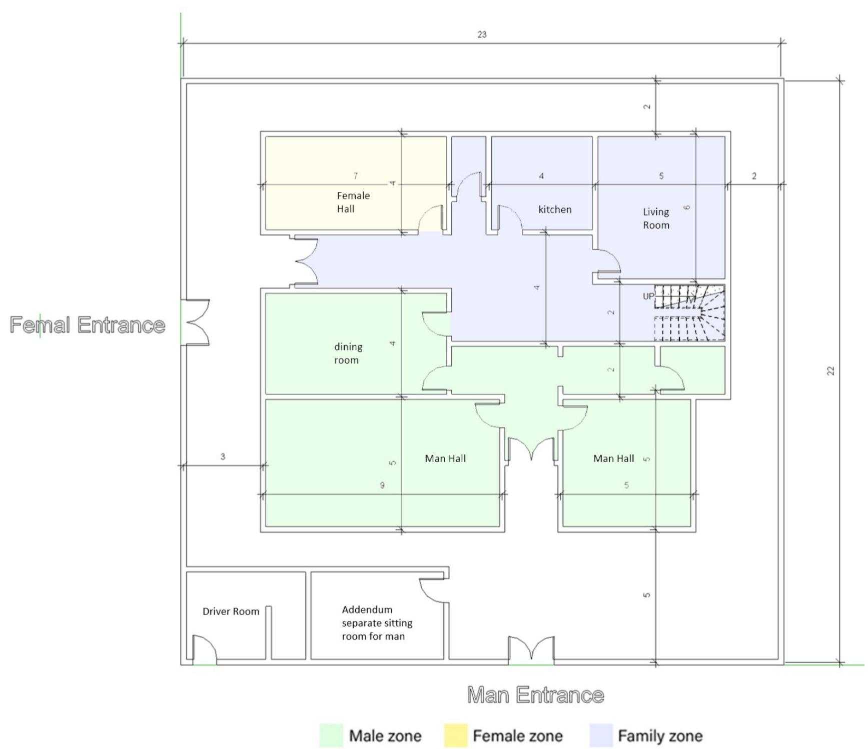


Fig. 4.10. Floor plan for a modern design built in 2004.

The privacy diagram (Fig. 4.11) shows that the level of privacy was reduced slightly by integrating the female visitors and the family members' zones. It was a significant change because this meant that people were more open and allowed female visitors to use some of the family rooms. The previous case study showed that the circulation of female visitors started from the entrance to the guest room, and the bathroom was separated, not overlapping with inhabitant male or female rooms. The present case study shows that female visitors shared most of the rooms with the female inhabitant, quite similar to the traditional house in the first case study.

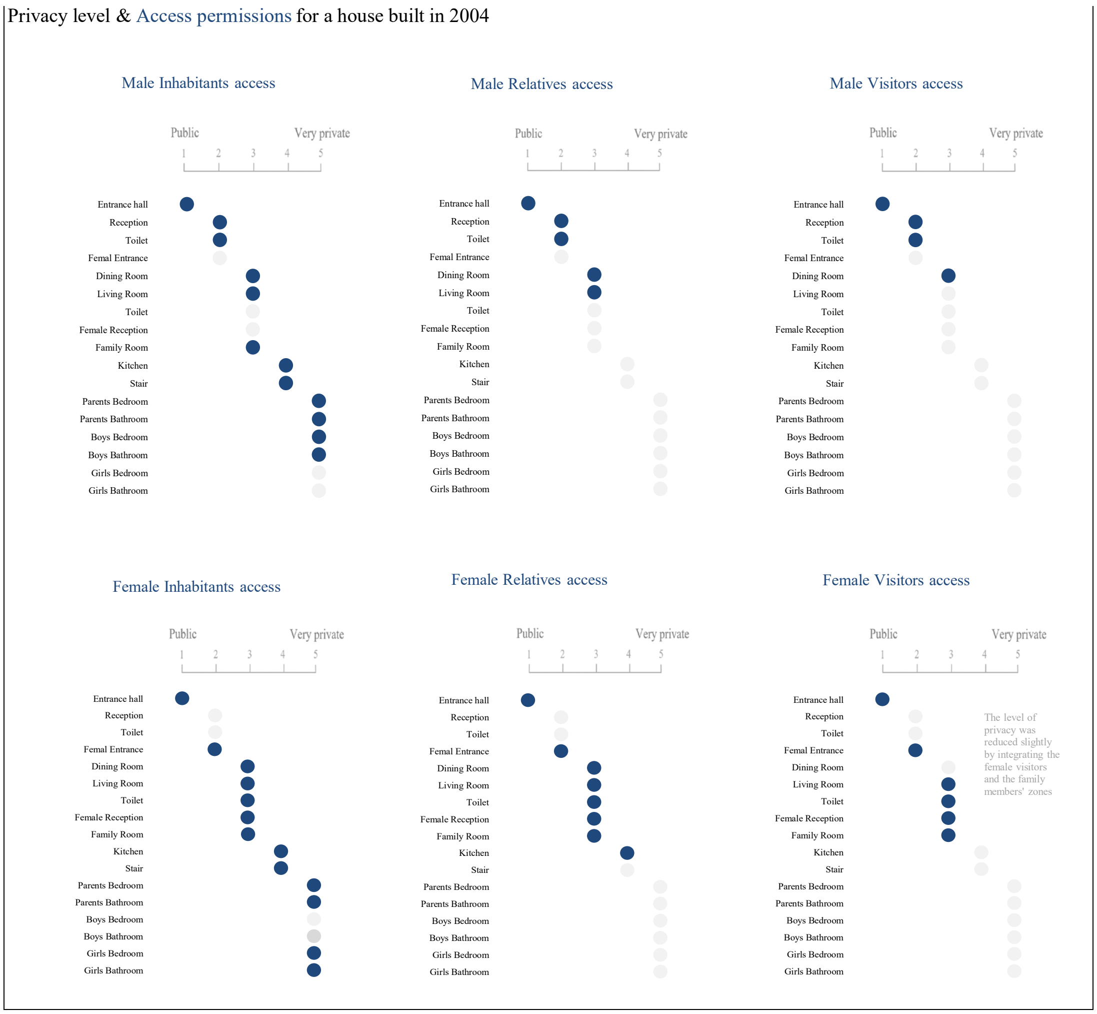


Fig. 4.11. Privacy level for a house built in 2004.

The visibility analysis shows that the connection number of points increased compared to the second and third case studies. The average number of points was 3,658. This is 900 points higher than the third case study. The maximum connected number of points was 6,875; this area was in the living hall with higher visibility and fewer walls (Fig. 4.12); the standard deviation was

1,260. This case study shows that the level of privacy was reduced, reflecting the house design and occupant behaviour, even though the male guest zone had the same privacy and visibility level as in the previous case studies.

The estimated net energy cost was: 145774000 Btu × 0.000293 = 42711.782 kW; 42711.782 kW / 12 = 3560 kW; 3560 kW × 0.10 SR = 356 Riyal/month (equal to 71 Pound). This shows a 58% higher energy cost compared to the previous study. The significant increase in energy cost was due to the big load and reliance on the HVAC system to cool big spaces like the living hall, which was used most of the day. The annual energy use per m^2 is 145774 Kbtu / 272 m^2 = 535 Kbtu/ m^2.

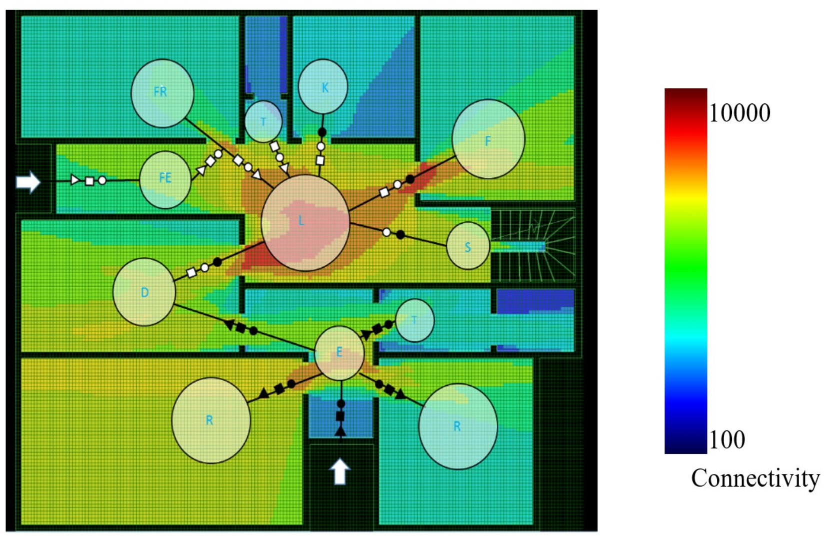


Fig. 4.12. Visibility analyses integrated with privacy level for a house built in 2004.

Fifth case study—house built in 2016: The built area of this house was 192 m^2. The design (Fig. 4.13) is a representative case for the current vision of Saudi houses. The male guest zone in this case study is in the range of the common proportion (30–40%) of the total built area. The zone has one large male guest room and dining room instead of two male guest rooms and one dining room, as in the previous cases. The female guest zone and living room are together, and the living hall is open to another room, which can be used as a female guest room and family room. There is a third entrance direct to the service stairs, which does not have access to the ground floor. There are fewer rooms overall as the plan is more open.

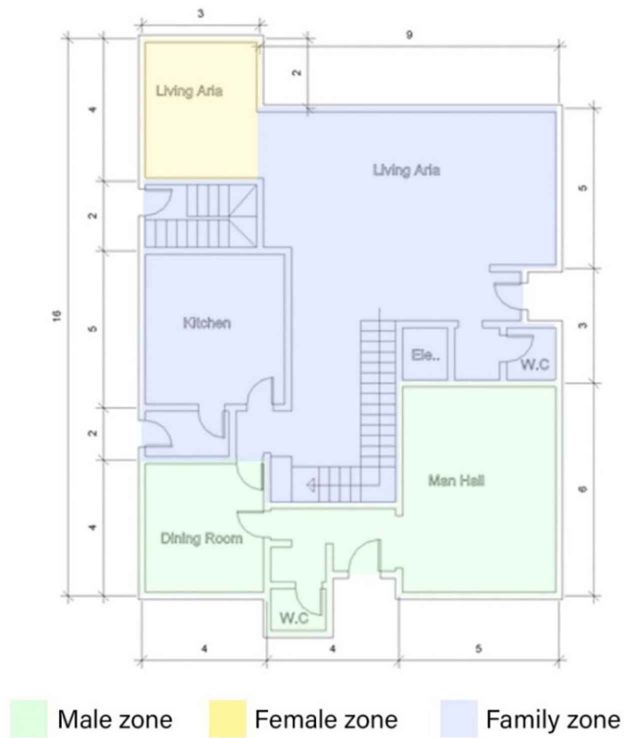


Fig. 4.13. Floor plan for a modern house built in 2016.

The privacy level is lower compared with previous studies. But still, the female visitors are separated from the male members; and the male guests are separated from the female members. The plan is similar to the traditional house in terms of privacy levels, but it is still more complicated than the traditional house (Fig. 4.14). In the last ten years, one of the changes in Saudi culture that has influenced behaviour and building design is the decreased number of visitors and guests. Therefore, the visitor areas are less important compared to previous houses, and the female and family share the same rooms with an open plan.

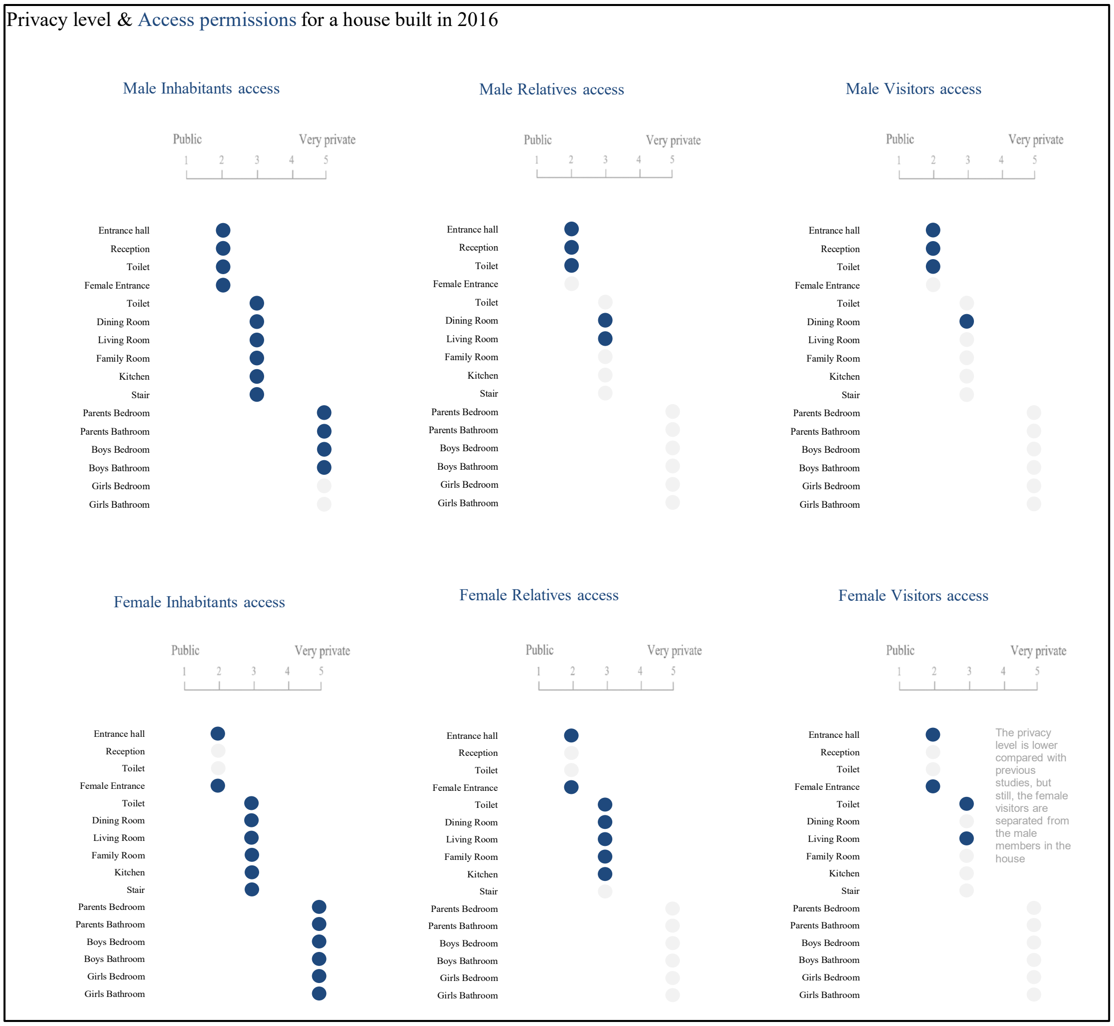


Fig. 4.14. Privacy level for a house built in 2016.

The visibility analysis in this case study shows that the average number of connected points is 3,636, almost equal to the fourth case study. The minimum number of connected points is 127 and the maximum is 6,745. The maximum connected points are fewer than the previous house design (6,875 points). However, the standard deviation here is 1,802 points, which is higher than the fourth case study (1,260 points). A high standard deviation means that the data points have greater variability. This means that the number of high connective points, in this case, is higher due to the big open house plan in the female and family zones. As Fig. 4.15 shows, high visibility points are more common in this house than in other modern house designs.

The energy use calculation was: 81522 000 Btu × 0.000293 = 23885.946 kW; 23885.946 kW / 12 = 1990 kW. The electricity use was lower than in case study four, but the cost is greater: 1990 kW × 0.30 SR = 597 Riyal/month. This was the result of increased electricity prices from 0.1 to 0.3 Saudi Riyal in 2016. Increased electricity prices reflect the Saudi government's intention of increasing energy costs to encourage residents to reduce energy consumption. This increase will likely continue to 2030 as a directive to reduce energy use in residential and commercial buildings. The annual energy use per m^2 is 81522 Kbtu / 192 m^2 = 424 Kbtu/ m^2.

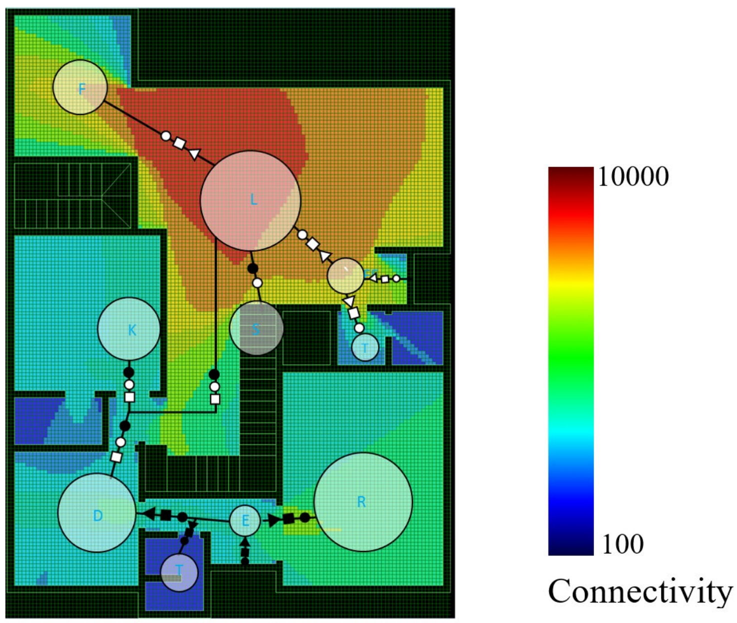


Fig. 4.15. Visibility analysis integrated with privacy level for a house built in 2016.

4.1.4. Summary

The case studies began with a traditional house. The second case study was a modern design built in 1978. The third house was built in 1990, the fourth in 2004, and the last in 2016. The most significant change in terms of building design between the traditional houses and the modern ones was the layout of the house. The traditional houses had a courtyard, but modern houses followed the setback technique.

Table 4.1 summarises the five case studies, starting with the plans to show the design changes from the traditional house to the modern. The plan highlighted in green represents the male guest zone. The yellow colour represents the female guest zone, and the blue colour represents the family zone. The guest zone became more important over time, which is reflected in the size of the zone compared to the other case studies. The size range of the total guest zone was between 40% to 70% of the total build area.

The second step integrates the privacy level diagram with the circulation analysis. The traditional house shows the courtyard located in the central area of the house (red colour). But in the second case study, the courtyard is replaced with the living hall. In the third case study, the living hall was moved from a central position to one side of the house. The fourth case study shows that the living hall became more of a path to the other rooms. In the last case study, the layout was more open, and the living hall opened to the female guest zone, kitchen, and stairs. In this house, there was no separated family room, which used to be the living room in the other case studies. The reason could be because of the house size or because people now prefer to use the living hall for family gatherings as well.

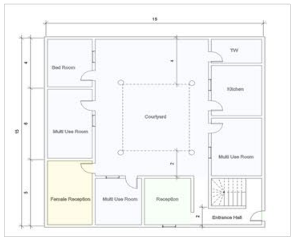
Comparison of visibility analysis and privacy level from the oldest modern house in the case studies to the newest (1978–2016), shows that the average visibility in the newer house is greater (Table 4.1). However, there was no significant change in privacy level. The different house occupants shared similar cultural principles, but the evolution in their house designs indicates that occupants of newer houses have fewer visitors.

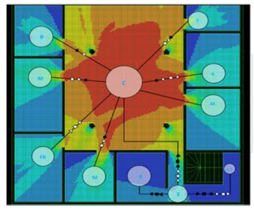
The standard deviation in the traditional house was the highest (3,487) compared to the other case studies. This indicates an open plan, which resulted in a high visibility number (maximum 10,631). On the other hand, private spaces for guests had a low visibility number (406). The newest case study (open plan layout) had the highest standard deviation (1,802) compared to the other modern houses. This also indicates an open plan with high privacy for guests, which accounts for the lower standard deviation as shown in the traditional case study. This is because old houses had courtyards in the middle and modern houses had front and back yards instead. Only internal spaces are used in simulation software. Therefore, interactions in front and backyards are not accounted for related to circulation considerations, as is the case with the traditional house.

Table 4.1

Summary of the experiment (design, privacy, and visibility).

Traditional House

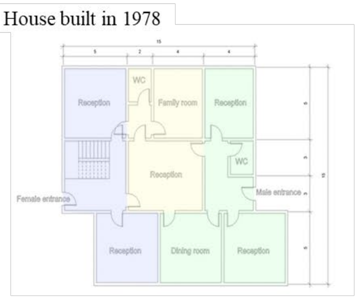


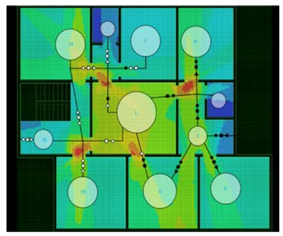


Average: 5378
Minimum: 406
Maximum: 10631
Standard Deviation : 3487

Connectivity 100 10000

House built in 1978

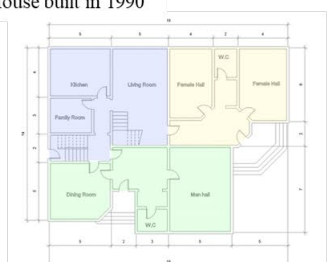


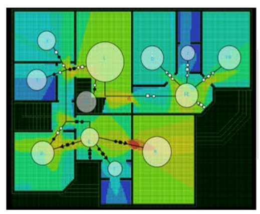


Average : 3029
Minimum: 417
Maximum: 6535
Standard Deviation: 963

Connectivity 100 10000

House built in 1990

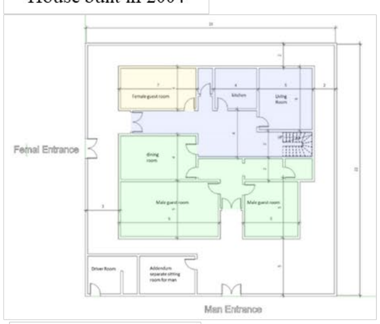


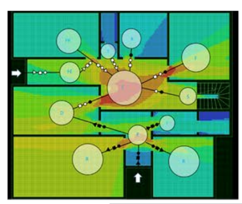


Average : 2755
Minimum: 419
Maximum: 5737
Standard Deviation: 880

Connectivity 100 10000

House built in 2004

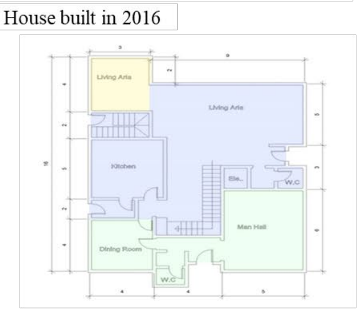


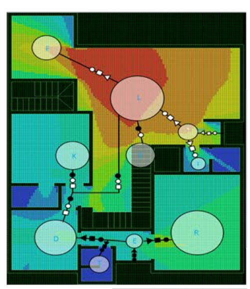


Average : 3658
Minimum: 303
Maximum: 6875
Standard Deviation: 1260

Connectivity 100 10000

House built in 2016





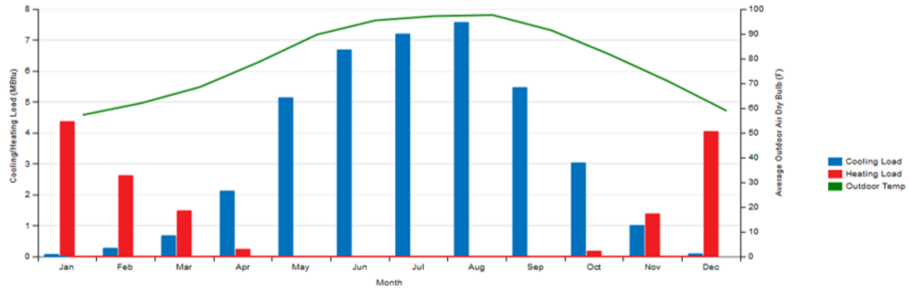
Average : 3636
Minimum: 127
Maximum: 6745
Standard Deviation: 1802

Connectivity 100 10000

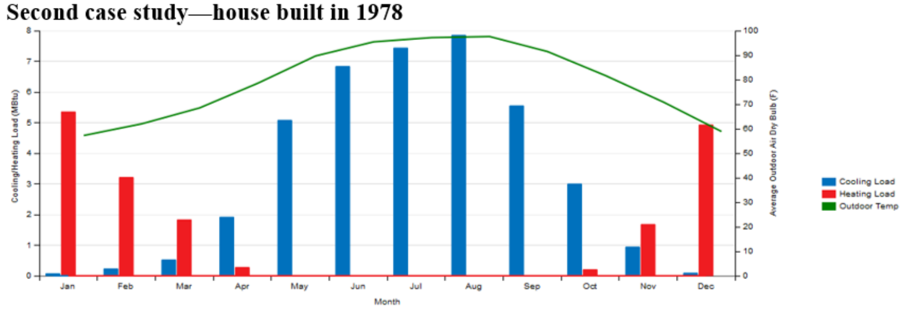
The privacy levels analysis shows complicated internal circulation patterns in modern houses compared to the traditional house. Saudi houses require three separate zones (male guest zone, female guest zone, and family zone), each zone used for a particular purpose. The guest zones are only closed and occupied when guests are present. However, there are no time-use schedules available to evaluate the energy performance for these specific zones, which cover up to 70% of the total building area. To close the energy performance gap, the lack of information on the time-use data of these zones needs to be addressed. This requires the development of three separate time-use schedules that can be applied for energy performance estimations. Considering these unique end-use energy patterns in energy simulations is envisaged to greatly improve energy prediction accuracy.

The electricity consumption experiment for the case studies shows that the traditional house (courtyard) consumes the lowest (326 kbtu per m^2). On the other hand, the new open plans layouts, found in houses built after 2000, consume the highest (535 and 424 kbtu per m^2). Because of the location of Riyadh (desert climate), the HVAC system uses the most electricity. Fig. 4.16 illustrates the cooling and heating loads with the outside temperature. It also shows that the traditional house uses less energy for cooling and heating compared to the open plan, which is the highest.

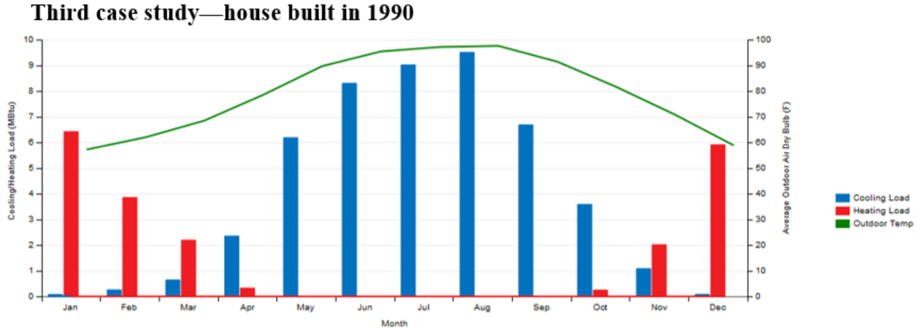
First case study—traditional house



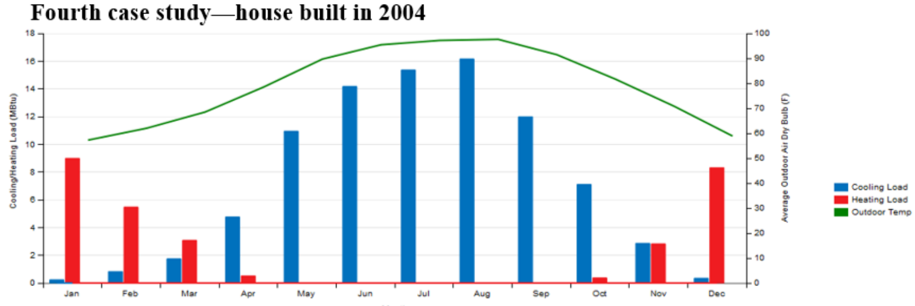
Second case study—house built in 1978



Third case study—house built in 1990



Fourth case study—house built in 2004



Fifth case study—house built in 2016

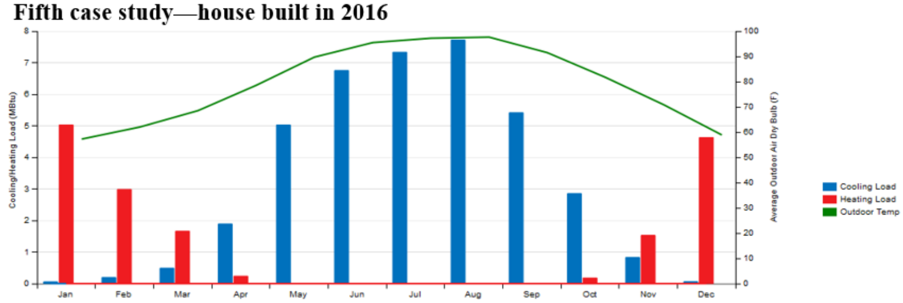


Fig. 4.16. HVAC load profiles and the outside temperature for the case studies.

4.2. Second pilot study

The goal of the first experiment was to understand the behaviour of Saudi house occupants and identify what factors influenced their behaviour. The first experiment demonstrated that spaces in Saudi houses could be grouped into three different user behaviour zones (male guest zone, female guest zone, and family zone). The lack of information about the energy use of these zones causes a gap between the energy simulation predictions and real energy use. The purpose of this research is to fill this lack of information by considering behaviour driven by cultural values. This second pilot study aimed to test and evaluate how considering cultural behaviour data in energy simulations can impact the outcome of energy simulations.

The second pilot study was divided into three stages. First, the house was modelled using SketchUp and OpenStudio with the data provided by the householder as input. The researcher used the default time-use load schedule provided by the Building Component Library (BCL) in OpenStudio. In the library, the most suitable schedule was the multiple-use room schedule. The hourly operation for this room is from 03:00 to 10:00 and from 14:00 to 23:00. The daily operation schedule is every day for the whole year. The goal of the first stage was to identify the energy performance gap that results from using default schedules. The second stage involved using the occupant answers to survey questions as input for the load schedule instead of the default schedules, to explore how the new model of using real energy use information can close the energy performance gap. The third stage aimed to improve the survey questions by interviewing the occupant and updating the actual load schedules in the energy simulation.

4.2.1. The first stage of data collection

The focal house was built in 2002 with 350 m^2 of ground floor space (Fig. 4.17). The family included two parents and four children aged between 6–15 years old (two boys and two girls). The house was modelled using Revit for the plans, SketchUp for the geometry, and OpenStudio for the energy simulation. The electricity consumption estimate was compared to actual electricity bills for 2016 and 2017. The gap reflected an approximate 40% higher predicted energy use compared to the actual energy use.

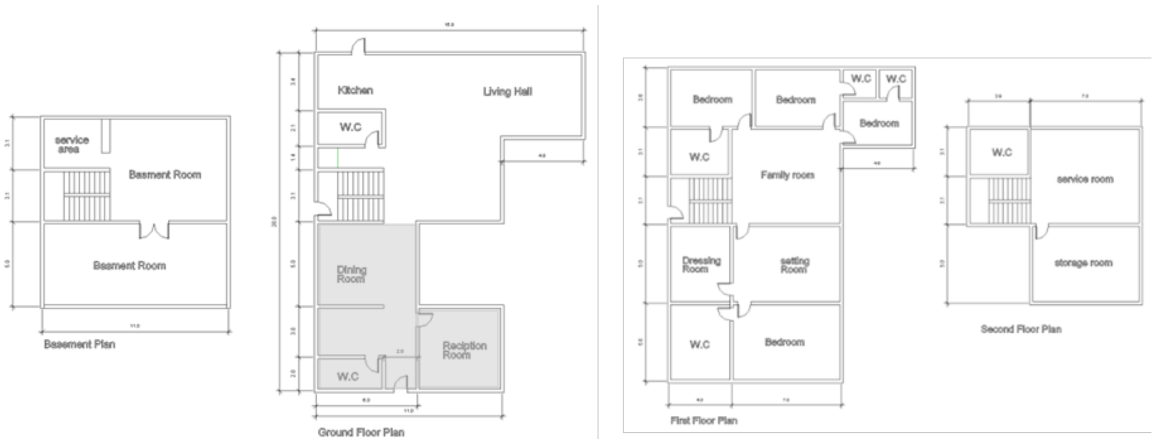


Fig. 4.17. Layout of the house in Riyadh used in the second pilot study.

4.2.2. The second stage of data collection

The second stage used occupant survey answers as load schedule input in the energy software. Occupants stated that they use the guest zone 90–100 times a year. The living hall was rarely used by the family, usually not more than once a month (because they had a family room for daily gatherings). Therefore it could be used by guests, retaining appropriate levels of privacy from the rest of the house. The basement of the house was mostly used during the weekend for gatherings of friends at night. As such, the guest room, dining room, living hall, and basement were not used daily. The kitchen, bedrooms, and family room were typically used daily.

The input for the energy simulation software using the survey answers went in a loop experiment. Using the time-use input from the occupants reduced the energy performance gap from 40% to 30%. Survey questions were re-examined to close the gap even further. The researcher discovered that the survey questions were confusing, and some occupants misunderstood questions. The survey questions were improved, tested for comprehension, and then data were collected again from the same occupants. After the survey questions were updated and the results were reanalysed using the new input, the energy performance gap reduced to 25%. The author decided to improve this further by adding hourly resolution time-use for the targeted zones to the survey. After reanalysis, this reduced the gap between predicted energy use and real energy use to 15%.

4.2.3. Interviews

The third stage was to conduct interviews using an open-answer method and then use the answers as input for the energy simulation prediction. Before conducting the interviews, visual observations were conducted of the house to understand the levels of privacy and to help identify questions for further investigation. Space Syntax software was used to analyse the visual data, which was applied to the circulation diagram of the house plan (Fig. 4.18). The program analysed flows through the house space.

The grid point spacing was set to 100,000 points and the walls were defined as solid to block vision. Space Syntax defined the visibility volume from the number of points that touch other points in the space with no wall to block the connection (Fig. 4.18). The highest connection in the plan, which means the highest visibility volume, was 2,850 points (the red colour on the plan) in the living hall and the lowest connectivity was 173 in the bathrooms and kitchen (dark blue on the plan). The main guest room showed the second-lowest level of connectivity on the plan. In this case, the owners have one main guest room for both male and female occupants. The female and family zones are joined together without walls, and therefore this area displays the highest levels of visibility. This led to further investigation of the circulation between the kitchen and the dining room through the female zone when guests are in the house.

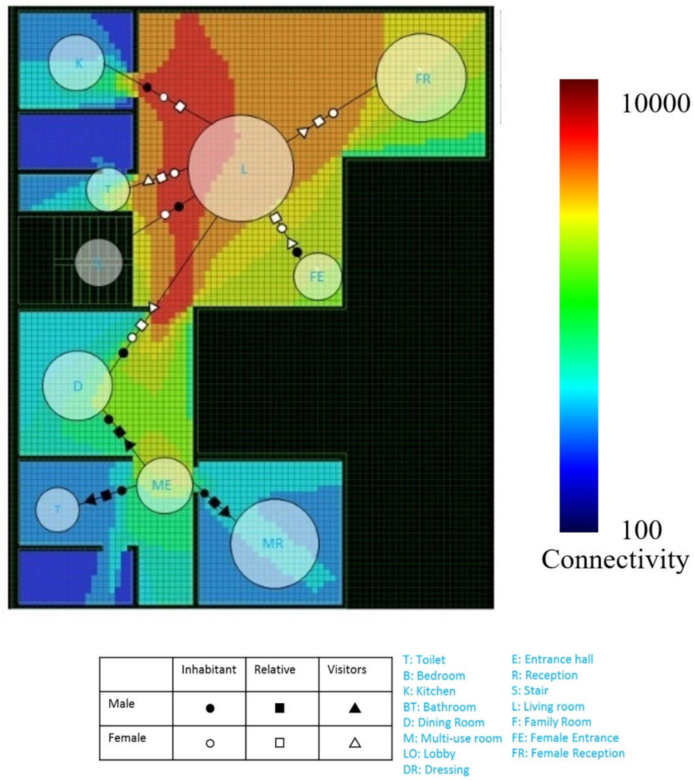


Fig. 4.18. Space syntax analyses for the case study.

The interviews took approximately one hour to complete for parents and less for the children. The first part of the interviews explored socio-demographic questions, building and occupant details, and unique behaviour questions related to energy use. The second part explored the occupant's lifestyle and how they use each room in detail. The data obtained was fed into time schedules. Other information collected was used to calibrate the questionnaire responses and identify any necessary design readjustments.

During the interviews, the occupants explained how the HVAC system was used. It was discovered that most people use the HVAC system mostly for cooling in summer and only for a few days of heating in winter. Additionally, the family was normally absent for three to four weeks each August, which lowered energy use. Considering these unique behaviours in the energy use prediction input reduced the gap between predicted energy use and real energy use to 5%.

4.2.4. Second pilot study summary

This pilot study showed how consideration of culture and behaviour in load schedules for energy simulations could achieve a more accurate prediction of energy use. This is particularly important in Saudi Arabia as the cultural use of living space differs from other cultures.

The survey and the interviews revealed the complexities of how Saudi Arabian families use spaces, especially when guests are present. Based on expected standards of privacy, guests must have rooms that are separated visually from other rooms in the house. Based on the initial case study in Riyadh, appropriate data collection methods were designed to reduce the energy performance gap from 40% to 15%. After the load schedules were updated with information from interviews with the occupant, the gap reduced even further to 5%.

In the experiment, a consideration of the new time-use on the load schedules (occupation, electricity and lighting), as well as for the HVAC time-use schedules, reduced the gap 10% more than considering the new time-use just on the load schedules.

The findings of this experiment can serve as a foundation for large-scale surveys that intend to look at cultural and behavioural factors in greater depth. To help refine and fine-tune the accuracy of energy performance evaluation further, the survey questions in this stage are ready to publish. The next chapter presents the data and analyses.

5. Data analysis

The goal of the current study is to address the gap in information about zone-specific occupant behaviour influenced by Saudi culture. In order to achieve this goal, the researcher collected occupant behaviour data from study participants residing in single-family homes in Riyadh, Saudi Arabia. Earlier chapters focused on the research articulating the impact of culture on occupant behaviour in their houses and highlighted specific guest zones contributing to energy performance gaps. Chapter 4 identified the lack of information on time-use for guest rooms, which is directly influenced by unique cultural behaviour.

Data and analyses from two surveys are reviewed in this chapter. The first survey consisted of 23 questions in a multiple-choice closed answer format. Questions were carefully chosen after studying the Saudi culture, as reported in Chapter 2. This first survey aimed to understand Saudi behaviour and defined the rooms in guest zones, which have unique time-use behaviour (hourly and daily resolution). The second survey consisted of eleven open answer questions and focused on the rooms that have unique time-use behaviour (daily resolution), as well as the parameters that influenced the time-use of these rooms. Previous research has not examined the daily time-use of guest rooms. Data from the second survey were used to generate the machine learning algorithm that was applied to case studies and study experiments, as discussed in Chapter 4. The algorithm is further discussed in Chapter 6.

5.1. Survey 1 (closed answers)

Initial survey data gathered information from 215 participants. However, study inclusion criteria only allowed for participants living in single-family homes in Riyadh, Saudi Arabia. After excluding the data from participants who do not live in single-family homes in Riyadh, a total of 215 responses were used from the 23 items multiple-choice question survey. Survey questions were divided into four main categories: 1) socio-demographic, 2) building characteristics and people who live in the house, 3) household behaviour influencing energy use, and 4) the time-use schedule for selected rooms. Table 5.1 shows questions from the survey. Google Forms was used to design and collect survey data. Microsoft Excel was used to clean and process the data, and SPSS software was used for data analyses.

Table 5.1
Survey questions.

Section 1	Socio-demographic characteristics
X1	What is your gender?
X2	What is your age?
X3	What is the occupation of the homeowner?
X4	Which part of Saudi Arabia do you live in?
Section 2	People who live in the house and building characteristics
X5	How many men live in the home including you?
X6	How many females live in the home including you?
X7	How old the person who responsible for the family?
X8	Where do you live?
X9	When was your home built?
X10	What is the total built area on the ground floor?
Section 3	Household behaviour influencing energy use
X11	How many rooms are there in your home for visitors?
X12	Which room do you use for the daily family gathering?
X13	Do you only use the dining room for guests?
X14	Do you have a room in your home for gatherings with friends?
X15	If you travel every year for more than two weeks, please identify.
X16	Do you use the HVAC system for heating in the winter?
X17	How much approximately is your monthly electricity bill in the summer? "Saudi riyal"
X18	How much approximately is your monthly electricity bill in the winter? "Saudi riyal"
Section 4	Time-use schedules
X19	How many times a year do you use the male guest room?
X20	How many times a year do you use the female guest room?
X21	How many times a year do you use the living hall?
X22	Check the rooms that you or any of your family used during the time below (during weekdays)
X23	Check the rooms that you or any of your family used during the time below (during weekends)

5.1.1. *Survey 1 findings data analysis*

Section 1—socio-demographic characteristics: Section 1 of the survey focused on the socio-demographics of research participants. Table 5.2 shows the number of participants by gender. Survey findings show that 62% of the study sample were female. Participants were aged between 18 to 74 years old, with 38% falling into the 25-34 years old category, followed by 27% in the 35-44 years old category (Table 5.3). Thirty-eight per cent of participants worked in government, 16.7% reported their occupation as working in a company, 18.1% reported working elsewhere, and 16.3% reported their occupation as retired (Table 5.4).

Table 5.2
What is your gender?

Gender	Frequency	Per cent
Female	134	62.3%
Male	81	37.7%
Total	215	100.0%

Table 5.3
What is your age?

Age	Frequency	Per cent
18 - 24	49	22.8%
25 - 34	82	38.1%
35 - 44	58	27.0%
45 - 54	18	8.4%
55 - 64	7	3.2%
65 - 74	1	0.5%
Total	215	100.0%

Table 5.4
What is the occupation of the homeowner?

Occupation	Frequency	Per cent
Working in the government	76	35.3%
Others	39	18.1%
Working for a company	36	16.8%
Retired	35	16.3%
Businessman	29	13.5%
Total	215	100.0%

Section 2—house occupants and characteristics: Section 2 of the survey focused on information about house and occupant characteristics. Initial questions in the section examined the number of people living in a house, which of course has a big influence on energy use. To understand the relationship between the number of people and energy use, the survey questions about the number of occupants were divided into two separate questions examining occupant characteristics by gender. Findings show that there were 711 males (Table 5.5) and 688 females (Table 5.6) in the 215 participant houses. Fifty-three per cent of householders were aged 50 or younger (Table 5.7).

Table 5.5
How many men live in the home including you?

Number of men	Frequency	Per cent
1	26	12.1%
2	46	21.4%
3	51	23.7%
4	20	9.3%
5 or more	72	33.5%
Total	215	100.0%

Table 5.6
How many females live in the home including you?

Number of females	Frequency	Per cent
1	27	12.5%
2	49	22.8%
3	44	20.5%
4	44	20.5%
5 or more	51	23.7%
Total	215	100.0%

Table 5.7
How old is the person who is responsible for the family?

Age	Frequency	Per cent
30 - 40	58	27.0%
41 - 50	56	26.0%
51 - 60	53	24.7%
61 - 70	35	16.3%
Over 70	13	6.0%
Total	215	100.0%

Householder age had a big impact on the time-use of guest zones. For example, the behaviour of a young single family is different compared to an older family. Comparisons between age and time-use of guest zones are further investigated in the discussion section. Additional questions in this section of the survey examined home characteristics, including the time the house was built (Table 5.8) and the size of homes (Table 5.9).

Table 5.8
When was your home built?

Year	Frequency	Per cent
1970s or before	1	0.5%
1980s	25	11.6%
1990s	39	18.1%
2000 – 2010	74	34.4%
2010 – 2018	76	35.4%
Total	215	100.0%

Table 5.9
What is the total built area of the ground floor?

Area	Frequency	Per cent
$100m^2$ or less	5	2.3%
$101m^2$ - $200m^2$	14	6.6%
$201m^2$ - $300m^2$	37	17.2%
$301m^2$ - $400m^2$	45	20.9%
$401m^2$ - $500m^2$	45	20.9%
more than $500m^2$	69	32.1%
Total	215	100.0%

These questions were important in order to understand the architectural layout designs and changes that have occurred over time. For example, in the early 2000s, open-plan designs became more popular; and this design had a direct influence on the time-use of guest rooms. Chapter 4 described the evolution and changes in house layouts and occupant behaviour of Saudi houses over the past 50 years. For example, courtyards used to be in the middle of the house; however, this has changed to the setback technique. The newest layout currently in use is the setback technique with an open-plan design. In this dataset, 70% of participants had their house built after 2000 (using the setback technique with an open-plan design), and 32.1% of participants live in houses 500 m^2 or bigger. Table 5.9 shows the results of the various sizes of homes in the study sample. Findings show that approximately 21% of participants live in homes ranging in size from 301 m^2 to 400 m^2 and an additional approximate 21% live in houses ranging from 401 m^2 and 500 m^2 in size. These findings indicate a high average house size in Riyadh city.

Section 3—occupant behaviour and energy use: Section 3 of the survey focused on occupant behaviour and energy use in guest zones. Guest zones consisted of at least one reception room, dining room, and bathroom. In terms of household visitors, 68% of respondents had two separate guest zones, one for males and one for females; 32% of respondents had one guest zone for both genders (Table 5.10).

Table 5.10
How many rooms are there in your home for visitors?

Rooms	Frequency	Per cent
One for both genders	67	31.2%
Two separate guest rooms	148	68.8%
Total	215	100.0%

House size influences the number and use of gender-specific guest zones. Smaller homes (less than 400 m^2) typically have one guest zone for both males and females. Guest zones in Saudi homes from this study sample accounted for approximately 30 to 40% of the total built home area. Results from survey findings indicate the size of the home did not correlate with lower use of guest zones; smaller homes had comparable use of guest zones to larger homes, suggesting 30% of the total built area is sufficient for just one guest zone.

In addition, survey data show that 67% of participants use the living hall for daily family gatherings, and 23% use the living hall for occasions only. Homes that only occasionally use the living hall have a separate daily family gathering room (Table 5.11).

Table 5.11
Which room do you use for the daily family gathering?

Rooms	Frequency	Per cent
Living Hall	146	67.9%
The Family room	62	28.8%
Guest Room	6	2.8%
No	1	0.5%
Total	215	100.0%

The data also shows that almost 50% of participants use the dining room daily, and the other half of participants only use the dining room when guests are in the house (Table 5.12). Furthermore, in addition to the guest zone, 71% of participants stated that they have a friend gathering room in their house (Table 5.13).

Table 5.12
Do you only use the dining room for guests?

Yes/No	Frequency	Per cent
No	107	49.8%
Yes	108	50.2%
Total	215	100.0%

Table 5.13
Do you have a room in your home for gatherings with friends?

Yes/No	Frequency	Per cent
No	62	28.8%
Yes	153	71.2%
Total	215	100.0%

Section 4—room time-use data: The last section of the survey explored the time-use of rooms in Saudi houses. This was divided into two parts: 1) daily and yearly time-use for rooms where time-use is influenced by culture and 2) hourly use for all rooms. A pilot study was

conducted (as described in Chapter 4). The pilot informed the survey design and consisted of an experiment using energy simulation software (OpenStudio) to simulate predicted energy use as compared to the real energy use of a case study for a house in Riyadh. Results of the simulation indicated a mismatch in predicted versus real energy use. The main contributor to the observed mismatch was attributed to the lack of information about the time-use of some rooms.

The room categories in this survey grouped the spaces into a male zone, a female zone, and a living hall. The first part explored how many days participants used rooms representing unique cultural use behaviour. Survey results indicate that male zones, female zones and living halls make up around 70% of the total area of Saudi homes. To examine daily time-use, the survey included a multiple-choice response set with options of ten-day increments until days reached a threshold of 140 days or more. If a participant chose 140 days or more, the researcher inferred this to mean the zone was used every week for more than two days (weekend) and therefore interpreted this to mean daily use (no unique behaviour).

The first question explored male guest zone time use. Survey data for the male guest zone (Table 5.14) showed that the greatest proportion of study participants (86%) use this zone between 1 and 40 times a year.

The second question examined female guest zone time-use. As mentioned before, smaller houses typically do not have a separate female guest zone. With this, 28% of participants chose zero, meaning they do not have a female guest zone in the house. Twenty-seven per cent of participants indicated they used the female guest zone between 1 and 10 times a year (Table 5.15).

Table 5.14
Time use of the male guest zone.

Number of times	Frequency	Per cent
1 – 10	103	47.9%
11- 20	43	20.0%
21 - 30	28	13.0%
31 - 40	11	5.1%
41 - 50	3	1.4%
51 - 60	6	2.8%
61 - 70	2	0.9%
91 - 100	5	2.3%
Between 131 - 140	1	0.5%
More than 140	13	6.1%
Total	215	100.0%

Table 5.15
Time use of the female guest zone.

Number of times	Frequency	Per cent
0	59	27.4%
1 - 10	58	27.0%
11 - 20	27	12.6%
21 - 30	20	9.3%
31 - 40	8	3.7%
41 - 50	6	2.8%
51 - 60	7	3.2%
61 - 70	7	3.2%
71 - 80	3	1.4%
81 - 90	1	0.5%
91 - 100	4	1.8%
101 - 110	1	0.5%
Between 121 - 130	1	0.5%
Between 131 - 140	1	0.5%
More than 140	12	5.6%
Total	215	100.0%

For the living hall time-use, data distribution showed two different types of behaviour (Table 5.16). Sixty-four per cent of participants use the living hall for more than 140 days a year, which was considered daily use. The remaining participants (36%) occasionally use the living hall (1 to 140 days). The greatest proportion of respondents indicating occasional living hall room use consisted of approximately 9% of the sample reporting use between 1 and 10 days a year.

Table 5.16
Time use of the living hall.

Number of times	Frequency	Per cent
1 - 10	19	8.8%
11 - 20	10	4.7%
21 - 30	11	5.1%
31 - 40	2	0.8%
41 - 50	4	1.9%
51 - 60	3	1.4%
61 - 70	4	1.9%
71 - 80	1	0.5%
81 - 90	4	1.9%
91 - 100	2	0.9%
101 - 110	2	0.9%
Between 111 - 120	4	1.9%
Between 121 - 130	2	0.9%
Between 131 - 140	8	3.7%
More than 140	139	64.7%
Total	215	100.0%

The next question explored how many hours participants used each room daily during the weekday and weekends (Fig. 5.1). These data were compared to average time-use data utilized in energy simulation software (in EnergyPlus). Results showed low time-use for guest zones, the living hall, and gathering rooms for friends, which is discrepant from average time-use data typically used in energy simulation software. Other rooms used daily, such as the kitchen, family room and bedroom, showed little discrepancies between default schedules for rooms' daily use provided by the energy simulation software and the average time-use from survey responses. Time-use data for targeted zones indicated higher usage during weekends. Average hourly time-use for the male guest zone, female guest zone, and living hall were used as fixed time-use schedules for the energy simulation input in the next chapter.

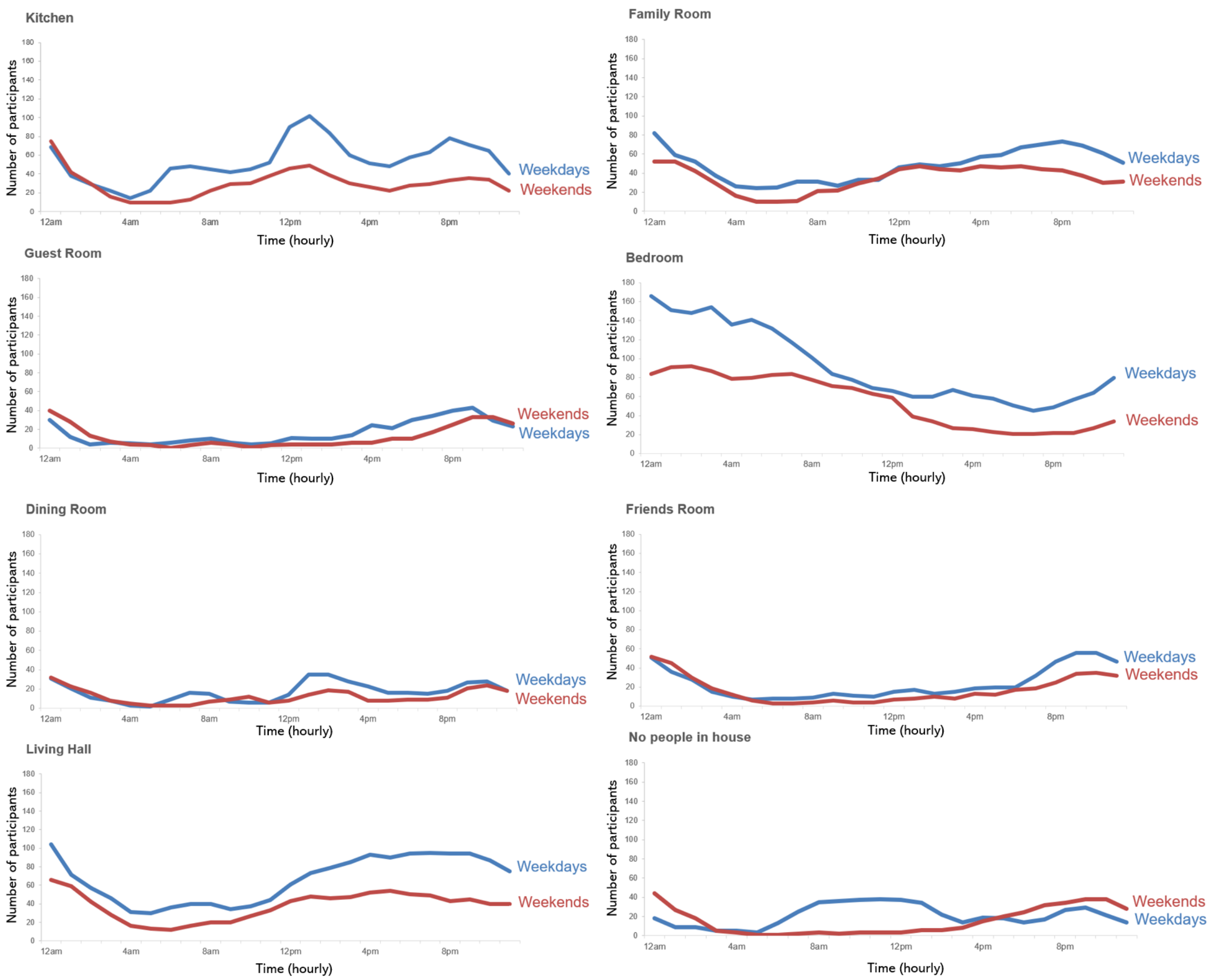


Fig. 5.1. Time-use of the rooms by frequency (weekdays and weekends).

5.1.2. Analysis discussion

Further analysis showed a relationship between householder age and guest zone time-use. Findings indicate that younger householders use guest zones less frequently than older householders. Figs. 5.2 and 5.3 show mean plots for householder age compared to the use of the male guest zone and female guest zone, respectively. Results show that time-use for both zones increases with the age of the householder until they reach 60 years old.

Small houses have one guest zone, which accounts for approximately 30% to 40% of the total built area. Guest rooms in small houses are used by either male or female visitors as long as their visits do not coincide. In other words, in small houses where there is just one guest zone for both genders, the male and female guest facilities were separated by time of occupation rather than by physical space.

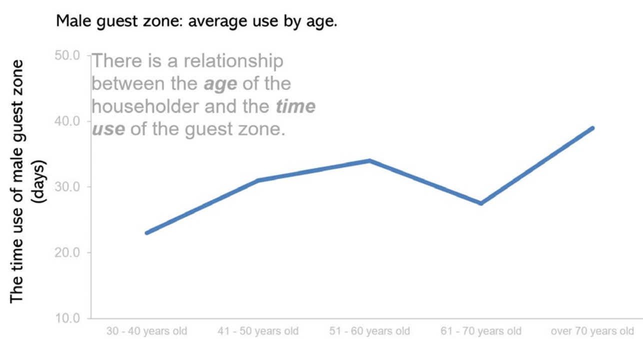


Fig. 5.2. Mean plot showing the relationship between the mean of the householder's age and the time-use in the male guest zone per year.

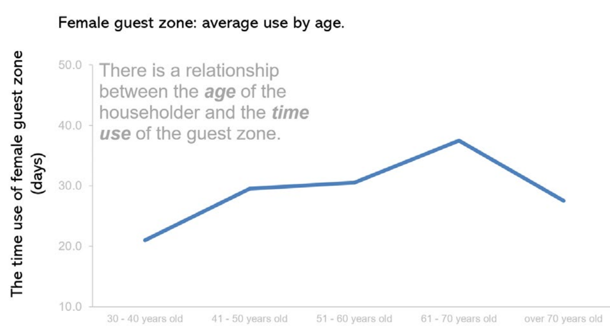


Fig. 5.3. Mean plot showing the relationship between the mean of the householder's age and the time-use of the female guest zone per year.

Survey questions also explored the time-use of guest rooms compared to the number of occupants. Fig. 5.4 illustrates the mean of the number of times male guest rooms are used annually compared to the number of males per household. The trend shows an inverse relationship between the number of male occupants and guest room use. The greater the number of male occupants in the house, the less the guest zone is used, until male occupancy reaches five or more male members. When male occupancy reaches five or more, guest zone use increases. Homes with five or more male members accounted for approximately 33% of respondents. These findings indicate a significant difference in the normal daily usage schedule included in energy simulation software. Failure to recognise this variation would lead to inaccurate energy estimates.

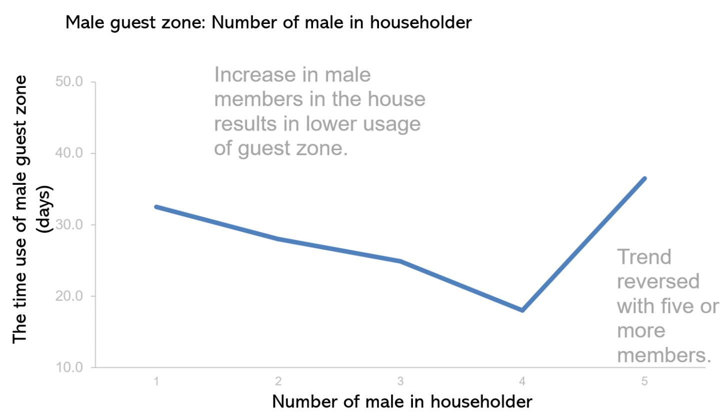


Fig. 5.4. Mean plot showing the relationship between the mean for the number of male members and the time-use for the female guest zone per year.

Traditional Saudi views require that female zones be visibly separated from male zones. Female zones have an entrance, reception room and a bathroom. Presently, some houses only have one zone used by both males and females. Fig. 5.5 shows the relationship between the number of females living in homes and the use of female guest rooms. Findings indicate time-use of female guest zones increases with the number of females living in homes. The lowest average use for one female member is 20 times per year. For homes with five or more female members, guest zones are used an average of 45 days.

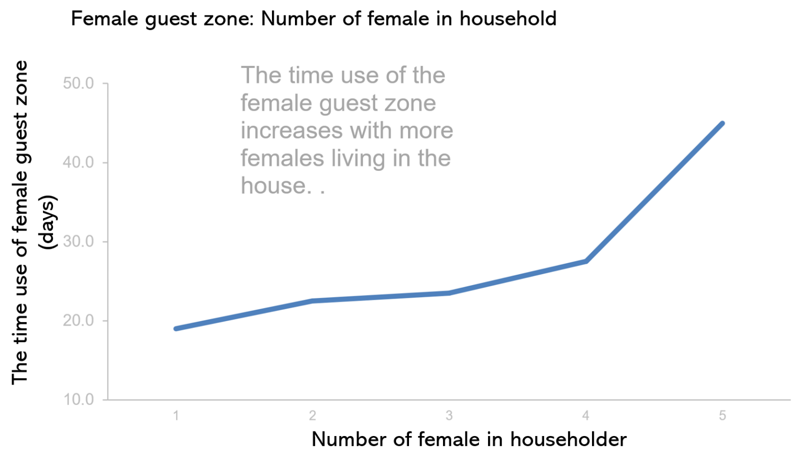


Fig. 5.5. Mean plot showing the relationship between the mean number of female members and the time-use for the female guest zone per year.

The space used for daily family gatherings varied between houses. For instance, some people used the living hall for daily family gatherings, while others had a separate family room and only used their living hall occasionally. Survey data showed a relationship between the size of the house and how spaces were used for family gatherings (Fig. 5.6). Families living in houses smaller than 500 m^2 tended to use the living hall for family gatherings, while people with larger houses tended to have a separate family room for daily use and a living hall only used occasionally.

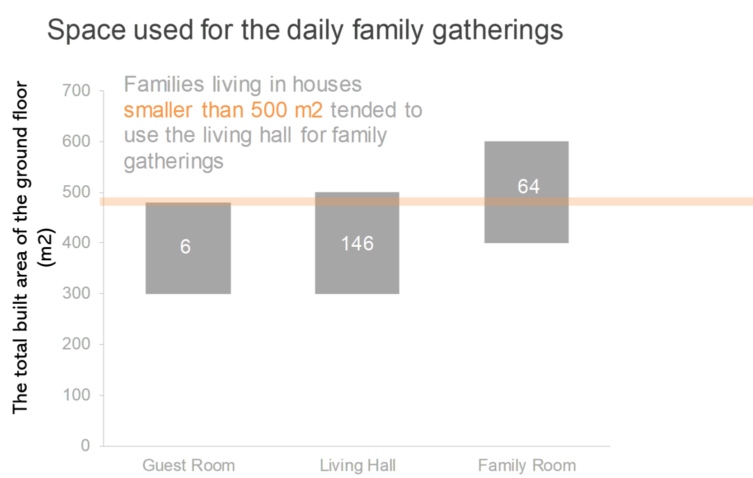


Fig. 5.6. Total built area of the ground floor and the space used for the family's daily gathering.

Correlation analyses: Correlational analyses were conducted to examine relationships between variables used in the study. Table 5.17 illustrate the results of correlation analyses.

Table 5.17
Correlation analyses.

	Gender	Age	Occupation	Male members	Female members	Householder age	House Built	House Size	Visitors Room	Room Gathering	Dining Guests	Room Friends	Travel	HVAC	Bill Summer	Bill Winter	TU male GZ	TU Female GZ	TU Living Hall
Gender	1	-0.07	.199**	-0.014	0.122	.167*	-0.029	.211**	0.005	.209**	-0.064	0.035	0.054	0.056	0.054	-0.006	-0.041	0.061	0.051
Age	-0.07	1	-.145*	0.033	-.180**	-0.102	0.071	-.193**	0.117	-0.013	.150*	0.11	0.049	-0.022	.205**	.139*	-0.002	0.031	-0.028
Occupation	.199**	-.145*	1	-0.098	0.077	.176**	-0.086	0.056	-0.079	0.072	0.077	-0.11	-.226**	-0.113	-0.074	-0.114	0.123	0.099	0.097
Male members	-0.014	0.033	-0.098	1	.165*	.196**	-0.087	0.052	-0.081	0.084	-0.04	0.015	-0.045	0.043	.193**	0.129	0.06	0.006	-0.072
Female members	0.122	-.180**	0.077	.165*	1	.364**	-0.12	.161*	-.254**	.159*	-0.038	0.132	-0.03	0.106	.141*	.175*	0.102	.213**	0.111
Householder age	.167*	-0.102	.176**	.196**	.364**	1	-.302**	.279**	-.276**	0.095	-0.04	.177**	-.173*	0.111	.145*	0.074	0.087	0.106	0.037
House built	-0.029	0.071	-0.086	-0.087	-0.12	-.302**	1	-0.1	0.01	0.035	0.128	0.085	0.07	-0.02	0.006	0.045	-0.067	-0.044	-0.045
House size	.211**	-.193**	0.056	0.052	.161*	.279**	-0.1	1	-.163*	.220**	-0.02	.166*	0.007	0.05	0.089	0.076	0.012	0.109	-0.081
Room for Visitors	0.005	0.117	-0.079	-0.081	-.254**	-.276**	0.01	-.163*	1	-0.132	-0.114	-.192**	-0.026	-0.045	-0.038	-0.094	-0.063	-.353**	0.1
Room Gathering	.209**	-0.013	0.072	0.084	.159*	0.095	0.035	.220**	-0.132	1	-0.002	0.07	0.082	0.024	-0.015	0.016	0.054	0.051	-.333**
Dining Guests	-0.064	.150*	0.077	-0.04	-0.038	-0.04	0.128	-0.02	-0.114	-0.002	1	.208**	-0.019	-0.081	0.101	.163*	-0.008	0.023	-0.022
Room Friends	0.035	0.11	-0.11	0.015	0.132	.177**	0.085	.166*	-.192**	0.07	.208**	1	0.051	0.026	.176**	.163*	-0.103	0.083	0.06
Travel	0.054	0.049	-.226**	-0.045	-0.03	-.173*	0.07	0.007	-0.026	0.082	-0.019	0.051	1	0.104	.138*	0.126	-0.11	-0.047	-0.021
HVAC	0.056	-0.022	-0.113	0.043	0.106	0.111	-0.02	0.05	-0.045	0.024	-0.081	0.026	0.104	1	-0.009	0.047	0.008	-0.002	0.037
Bill Summer	0.054	.205**	-0.074	.193**	.141*	.145*	0.006	0.089	-0.038	-0.015	0.101	.176**	.138*	-0.009	1	.733**	-0.022	0.033	0.091
Bill Winter	-0.006	.139*	-0.114	0.129	.175*	0.074	0.045	0.076	-0.094	0.016	.163*	.163*	0.126	0.047	.733**	1	0.027	-0.004	0.072
TU male GZ	-0.041	-0.002	0.123	0.06	0.102	0.087	-0.067	0.012	-0.063	0.054	-0.008	-0.103	-0.11	0.008	-0.022	0.027	1	.528**	0.101
TU Female GZ	0.061	0.031	0.099	0.006	.213**	0.106	-0.044	0.109	-.353**	0.051	0.023	0.083	-0.047	-0.002	0.033	-0.004	.528**	1	0.133
TU Living Hall	0.051	-0.028	0.097	-0.072	0.111	0.037	-0.045	-0.081	0.1	-.333**	-0.022	0.06	-0.021	0.037	0.091	0.072	0.101	0.133	1

Findings show a significant correlation between the number of male members and participant age. However, this is considered less important in this study because this study focused on householders who live in the house with their family but not the age of non-householder participants. Most of the survey data were collected in-person, and some of the participants were unmarried students. In Saudi culture, males typically live with their family until marriage.

This section will focus on correlations of important variables in the study:

- There is a significant relationship between householder age and the variables related to the number of male members and female members living in homes Table 5.17 illustrates the statistically significant relationship between the variables. The p-value for the number of male members and the age of the householder is 0.004, and the correlation coefficient is 0.196. For the number of female members and the age of the householder, the P-value is $p<0.001$, and the correlation coefficient is 0.364.
- A strong correlation was found between householder age and occupation (Table 5.17). Results show that younger householders work more for companies, and older householders work more in government. Most householders older than 50 years old either work as businessmen or are retired.
- Furthermore, "when homes were built" and "house size" were strongly correlated with householder age (Table 5.17). Older and bigger houses were owned by older occupants and newer smaller houses by younger families. Fig. 5.7 illustrates a positive trend between householder age and house size. Experimental findings from Chapter 4 showed older houses have more guest rooms compared to new houses. Further, younger householders preferred to have one guest zone for both genders. The householder age

is significantly correlated with seven variables, and thus it is one of the most significant variables.

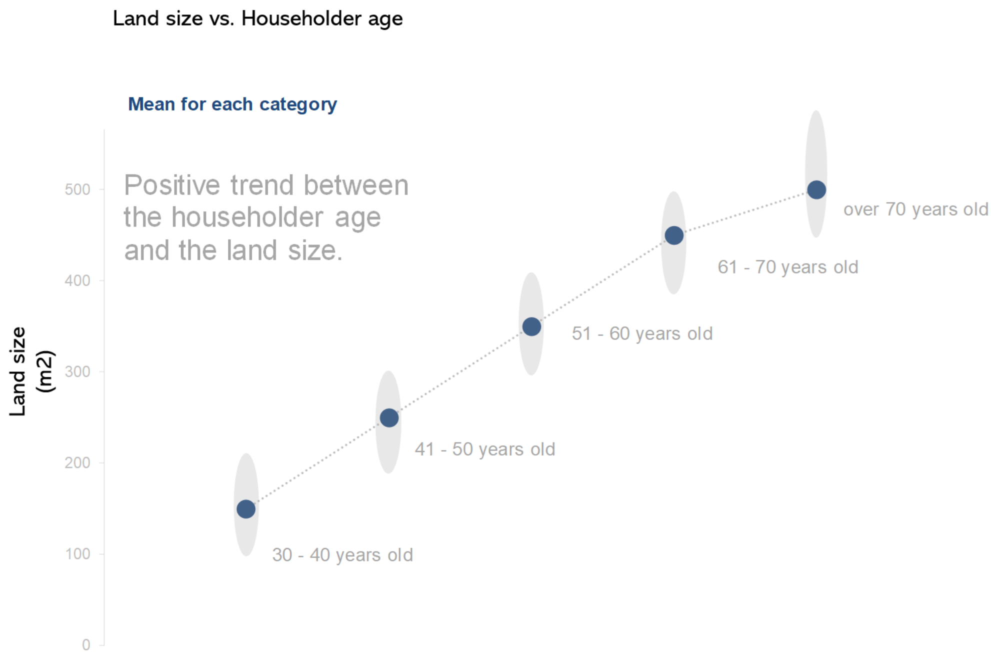


Fig. 5.7. Mean plot shows the relationship between the land size and the householder age.

- Findings also show a strong correlation between the number of guest zones and house size and householder age variables (Table 5.17). As noted earlier, most smaller homes have one guest zone for both genders. Fig. 5.7 illustrates that older householders typically have bigger homes.
- The relationship between female members and the number of guest zones is also significantly correlated (Table 5.17). Forty-five per cent of participants reported having four or more female members in their homes. Eighteen per cent reported having one guest zone for both genders. The remaining 82% reported having two separate guest zones, one for males and the other for females. The analysis shows that the number of

female members has a greater effect on house design and behaviour than the number of male members.

- Analysis indicates a significant correlation between the number of female members and female guest zone use (Table 5.17). Findings show a positive trend between the two variables. The number of female members is also strongly correlated with other variables such as householder age, house size, number of guest zones, room use for gathering, and summer electricity bills. Therefore, it is also a significant variable.
- Findings show gathering rooms have a strong correlation with the size of the house (Fig. 5.6). Small houses use the living hall for daily family gatherings. Big houses typically have a separate family gathering room. In addition, there is a correlation between daily family gathering room use and the number of female members. The data show that families who have more than three female members in the house intend to have a separate family gathering room for daily use.
- Results also show a significant correlation between householder age and the time they travel (Table 5.17). The data show that most people travel in July and August. However, 82% of the retired participants do not travel for more than two weeks a year.
- The correlation analysis also shows a significant correlation between the summer electricity bill and the total number of household occupants, meaning the greater number of occupants, the greater the electricity bill (Table 5.17). The number of household occupants has less impact on the winter electricity bill. This finding may be because the HVAC system is off most of the time in the winter as there are only a few days cold enough where people use their HVAC for heating.

- Findings also show female guest zone use is significantly correlated with the number of guest zones in the house (Table 5.17). People with two guest zones use them more when more than three female members are living in the house, suggesting this positive correlation between female occupancy and guest zone use.
- Living room time-use was found to be correlated with family gathering room type and house size (Table 5.17). Bigger houses that have separate family gathering areas use the living hall less.

5.1.3. Summary

The outcomes from the first survey analyses are as follows:

- Among rooms in Saudi houses, the study shows that three zones have unique time-use behaviour influenced by culture. These rooms are the male guest zone Y1 (Table 5.14), female guest zone Y2 (Table 5.15), and living hall Y3 (Table 5.16).
- Based on the studies in Chapter 4, the total size of the area for guests is 30–40% of the total built area. The full guest zone has a reception room, dining room, and bathroom. Forty percent of the total built area smaller than 400 m^2 is not large enough for two guest zones.
- Table 5.12 shows that 50% of the participants use the dining room only when there are guests in the house. The average time-use for people using the dining room just for guests is 24 days a year.
- Younger families use the male and female guest zones less than older ones (Figs. 5.2 and 5.3).
- Houses that have more male members tend to use the male guest zone more when there are five male members or more in the house (Fig. 5.4).

- There is a positive trend between the number of female members living in the house and the time-use of the female guest zone (Fig. 5.5).
- People with small houses use the living hall for daily family gatherings. People with larger houses have a separate room for daily gatherings and use the living hall occasionally (Fig. 5.6).
- Friend rooms are separated from the house, located next to the male outside entrance. Saudi homes have these rooms separated because of the cultural need for privacy, especially when male friends visit. This room is for casual male friend gatherings. The data show that 71% of participants have a room called a “friends’ room” (Table 5.13) which suggests that adding a daily time-use question for friends' rooms in the second survey was important to understand if a daily unique time-use behaviour needed to be added to the model.

5.2. Survey 2 (open answers)

The data collected from the second open-ended survey will be used as input to train and validate the time-use prediction model discussed in the next chapter. This model will predict the time used for the rooms with unique types of behaviour (male guest zone, female guest zone, and living hall). The questions were chosen, as mentioned before, from the closed answers survey. Twenty-six experts examined the data from the first survey to identify what could be added to improve the accuracy of the prediction. From this, two pieces of information were added: (X1) the location of the house; and (X2) whether the householder had brothers and, if yes, if he was the oldest, middle, or youngest. These two questions were multiple-choice (Table 5.18). The next five questions were from the first survey: (X3) age of the householder; (X4) when the house was built; (X5) how many male members live in the house; (X6) how many female members live in the house; and (X7) the size of the land. These seven questions (parameters) form the first section of this survey. The second section focused on the predicted use of rooms based on the (Y1) male zone, (Y2) female zone, and (Y3) living hall. One additional question was added: (Y4), which examined how many times participants used the outside room (Table 5.18). The framework used for the survey data collection, processing and analysis was the same as in the first survey. The total number of respondents was 1,230 people living in homes in Riyadh, Saudi Arabia.

Table 5.18
Survey questions.

X1	Location of the house (north, south, east, west, or in the middle of the city).
X2	Does the householder have brothers and, if yes, is he the oldest, middle, or youngest?
X3	Age of the householder?
X4	When was the house built?
X5	How many male members live in the house?
X6	How many female members live in the house?
X7	The size of the land
Y1	How many days do you use the male guest zone in one year?
Y2	How many days do you use the female guest zone in one year?
Y3	How many days do you use the living hall in one year?
Y4	How many days do you use the outside room in one year?

5.2.1. Data processing

The open-ended survey questions were included so that every participant could write without limit. Like the first survey, Google Forms was used to design and collect the survey data. Data were organised manually by reading through all 1,230 participants' answers and cleaning the data accordingly. The answers were revised to match one format. For the survey question related to when the home was built, responses were the date the house was built. Answers using the Hijri calendar, or the number of years since the house was built, were converted to the Gregorian calendar. Data from participants who left answers empty were removed.

For the data to fit the statistical analysis requirements, two corrections were made. First, the data were adjusted for each question to follow one style (e.g., changing answers from Arabic to English and changing the building age from date to the number of years). Second, participants who were not from the target region (Riyadh) were excluded. The sample included participants who live in a single-family detached house in Riyadh, Saudi Arabia. Participants not fitting these criteria were excluded. The total number of eligible respondents after data cleaning was 870. The last step in the processing stage was to label the answers for the multiple-choice items. Then at this stage, the data were cleaned, organised and prepared for analysis. The researcher used SPSS to conduct analyses examining data distribution, density, mean, and standard deviation. The scikit learn library was used for analyses related to correlation and distribution of means. More details about the workflow can be found in Chapter 3 on methodology.

5.2.2. Data analyses

5.2.2.1. Question 1: Where do you live?

The first question, related to home location, covered the first parameter for the model. Labels for responses are noted below (Fig. 5.8). If the house were in the

- middle of Riyadh, then it was labelled number (1),
- north side of the city, it was labelled number (2),
- south side of the city, it was labelled number (3),
- east side of the city, it was labelled number (4), and
- west side, it was labelled number (5)

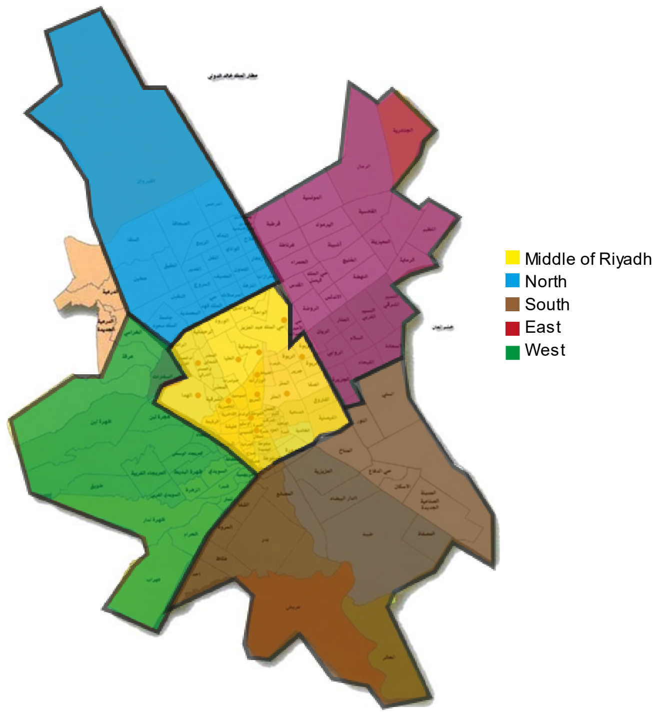


Fig. 5.8. Map of Riyadh highlighting the five areas mentioned in question one.

Table 5.19
Number of participants in each area.

Location	Frequency	Per cent	Cumulative per cent
North	360	41.5	54.8
South	75	8.6	63.5
East	201	23.2	86.6
West	116	13.4	100.0
Total	868	100.0	

Table 5.19 shows the number of participants in each area. Forty-one per cent of participants lived in the north of the city, the newest part. North and middle areas are new compared to the houses south of the city. The data show that each location of these five areas has different behaviour in terms of guest zone usage. The greatest expansion of Riyadh city has been to the north.

Fig. 5.9 shows the relationship between the location of the house and the land size. The chart shows the mean for each category using a bootstrap approach (Horowitz 2001). The mean for houses in Riyadh's middle area is 920 m^2, and 800 m^2. for the north. The middle and north areas are the largest in the city. The smallest average land size is in the south, with a mean of 600 m^2. The number of people who live in houses in the south area is large compared to the house size. The houses are bigger in the north, but the number of occupants and the time-use of the guest zones are less than in houses of the same size in the southern area.

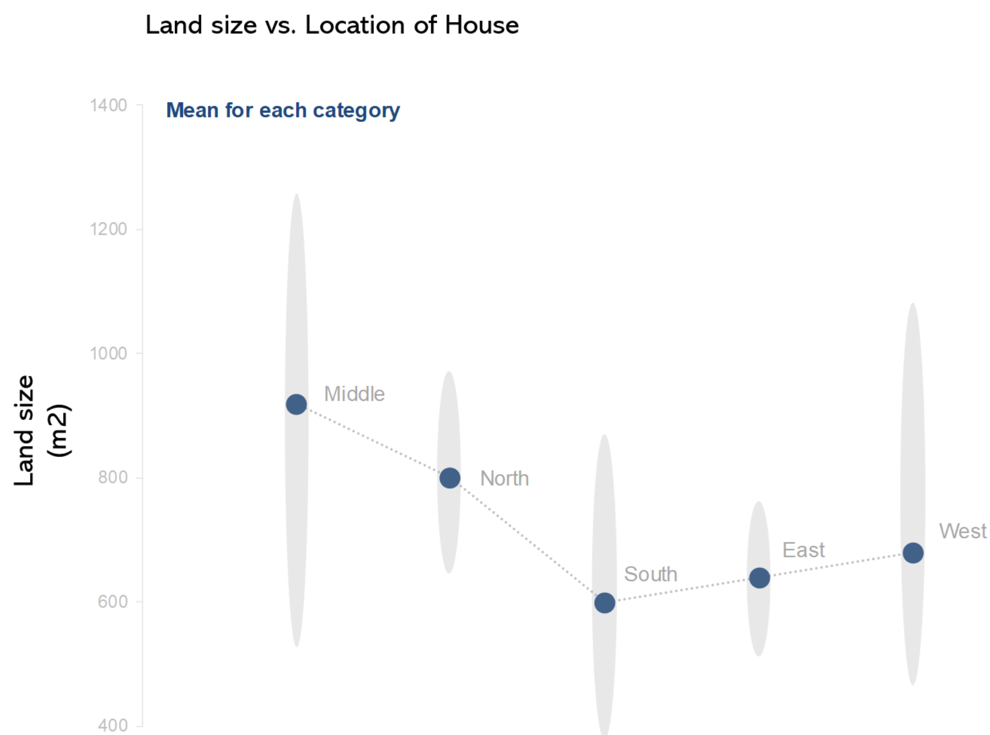


Fig. 5.9. The relationship between the location and the size of the house.

5.2.2.2. Question 2: Birth order

For question two, which explored whether respondents had brothers, participants provided responses related to their birth order (i.e., whether they were the oldest, youngest, or in the middle). If participants indicated they were the oldest brother in his family, their responses were labelled number 1; if they indicated they were in the middle, their response was labelled number 2; if they indicated they were the youngest, their response was labelled number 3; and if they indicated having no brothers, the responses were labelled number 4. Table 5.20 shows that most participants in this dataset (470 of 868) were labelled as the eldest. The second highest participants were in the middle, followed by the youngest; the smallest number of participants were those who did not have brothers.

Table 5.20
Family member.

	Frequency	Percent	Cumulative percent
Biggest brother	481	55.0	55.4
Middle	278	31.8	87.4
Younger brother	83	9.5	97.0
No brothers	26	3.0	100.0
Total	868	100	

Fig. 5.10 shows the mean use of the male guest room for each answer. The analysis shows similar behaviour in terms of using the zones for each label. The oldest brother used the male and female guest zones the least (38 days per year), whereas the participants without brothers used the guest room the most (54 days). The findings indicate that the eldest brother uses guest zones more than his younger brothers. This is consistent with Saudi culture, where it is typical for brothers to gather in the oldest brother's house. However, recall the privacy level diagram discussed in Chapter 4. The household brothers are considered inhabitants, so in the Saudi culture, they are allowed to go inside the house and not just in the guest zone. The data show that the living hall is used most by the oldest householder and least by people with no brothers. This is because when brothers visit, they gather in the living hall, not the guest zone.

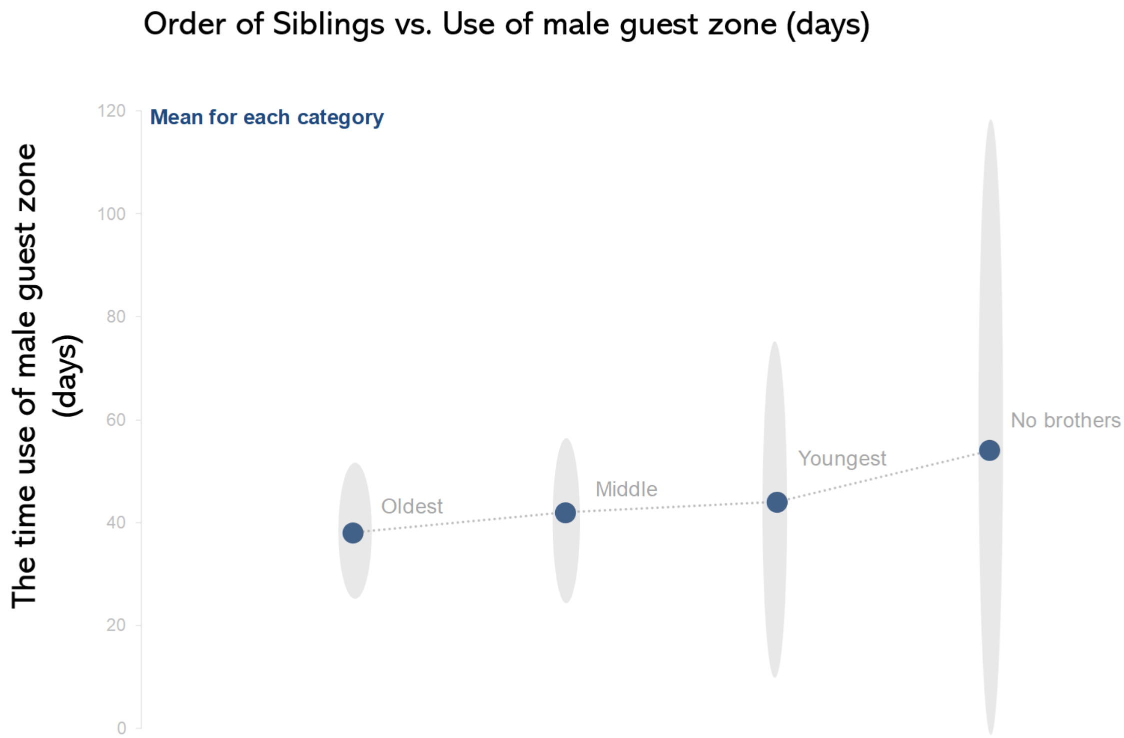


Fig. 5.10. The relationship between the order among the siblings for the householder and the use of the male guest zone.

5.2.2.3. Question 3: The age of the householder.

Fig. 5.11 shows the age distribution of participants. The youngest participant was 20 years old, and the oldest was 89. The mean age of participants was 53 years old (yellow line in the bar chart), and the standard deviation was 12.2. The first standard deviation (blue line) covers 68% of the data and is between the ages 41 and 65. The second standard deviation (red line) covers 95% of participants and includes respondents' ages between 29 and 77 years old.

Fig. 5.12 is the distribution of means showing the confidence intervals of the data. The figure helps to understand the relationship between the householder's age and the use of the guest zones. The plot defines the range of the mean for each age group using a bootstrap technique. The age data were grouped by decade—20, 30, 40, 50, 60, 70, and 80—and means were calculated by each ten-year grouping. The outcome for this analysis shows an increasing trend with age (except for the oldest age group).

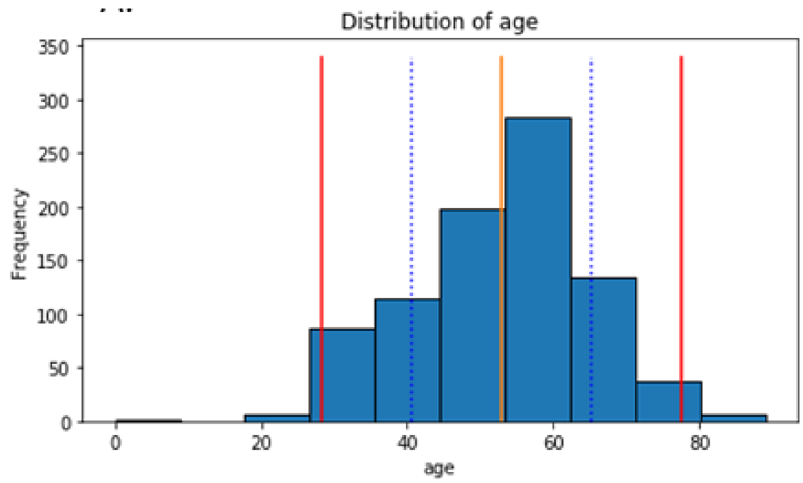


Fig. 5.11. The frequency distribution of the number of the respondents' age, including the mean (yellow line) and the first (dotted blue line) and second (red line) standard deviation.

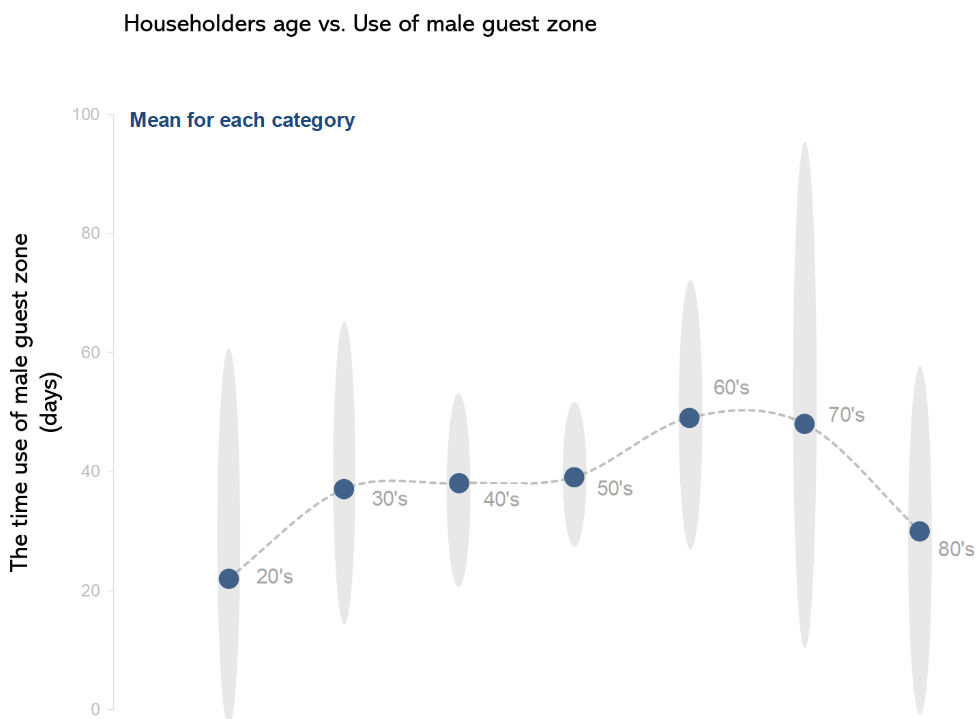


Fig. 5.12. The relationship between the householder's age and the use of the male guest zone.

5.2.2.4. *Question 4: How many men live in the house?*

This is the fourth parameter for the occupant model. Based on the case study analysis in Chapter 4, the number of male members living in the house has a big impact on the size of and time use of male guest zones. Fig. 5.13 shows that the highest proportion of participants (n = 221) had three male members living in the house, followed by homes having four male members (n = 79). The average number of men living in the homes was 3.56. The first standard deviation that covers 68% of the data was 1.8 from the average number, between 1.76 to 5.36 people. The second standard deviation included a range from zero to 7.16 people.

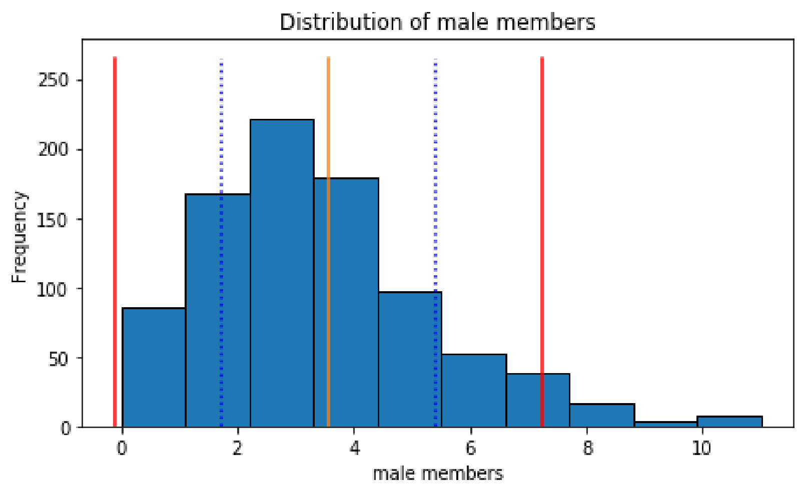


Fig. 5.13. The frequency distribution of the number of male members living in the house, including the mean (yellow line) and the first (dotted blue line) and second (red line) standard deviation.

Fig. 5.14 shows the relationship between the number of men living in the house and the use of the male guest zone. The middle chart is the mean for every bin group (2, 4, 6, 8, 10, 12). There is a hump-shaped (unimodal) trend between the number of men and guest zone use. This suggests decreasing guest zone usage in houses with more than six male occupants. The reason for this behaviour is usually because households with more than six men living in the house are not single-family. However, in this research, we are focusing on single-family house behaviour.

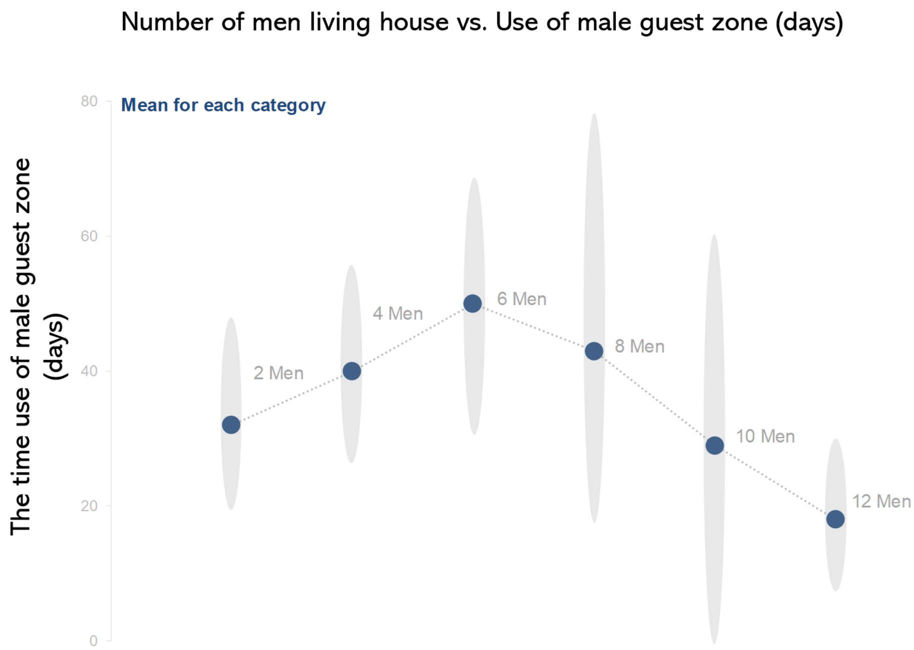


Fig. 5.14. Distribution of bootstrapped means of the number of men living in the house and the use of the male guest room.

5.2.2.5. Question 5: How many female members live in the house?

The first survey showed that the number of females living in the homes had a big impact on the use of the female guest zone and living hall. Fig. 5.15 shows 220 participants stated that three female members were living in the house, compared to 183 participants who stated that two to four females live in the household.

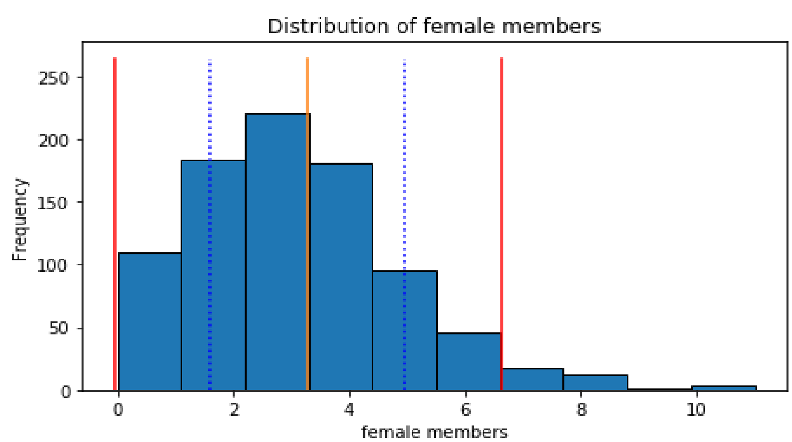


Fig. 5.15. Panel A shows the frequency distribution of the number of female members living in the house, including the mean (yellow line) and the first (dotted blue line) and second (red line) standard deviation.

In 11 households, no females lived in the house. The mean is 3.29, and the standard deviation is 1.67. The first standard deviation covers 68% of participants, having a range between 1.6 and 5 female members in the house. The second standard deviation covers 95% of the dataset, ranging from 0 to 6.7 female occupants.

Fig. 5.16 shows the relationship between the number of females living in the house and the use of female guest zones. Homes with up to two female members used the female house zone on average 25 days per year. Houses with six to eight female members used the female guest zone for an average of approximately 60 days per year. There is an increasing trend from zero to eight female members, which covers the first and second standard deviation (95% of the data). The use decreases with more than eight female members, suggesting different behaviour of using female guest zones when more than one family lives in a house.

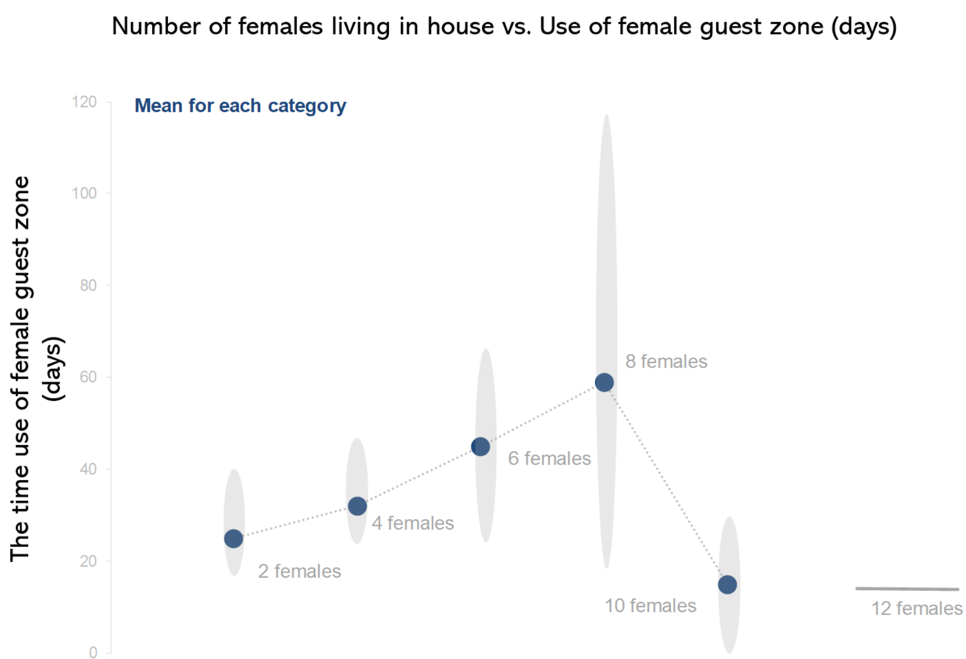


Fig. 5.16. The relationship between the number of females living in the house and the use of the female guest zone.

Fig. 5.17 shows the relationship between the female members living in the house and the use of living halls. The findings suggest an increasing trend for living hall use when households had up to eight female members. The greater the number of female occupants, the greater the use of the living halls. This room use is also related to house size, which is discussed in a later section. In smaller houses, living halls are also used as a family gathering room (daily use). People with bigger houses have separate family gathering rooms for daily use and use living halls only occasionally.

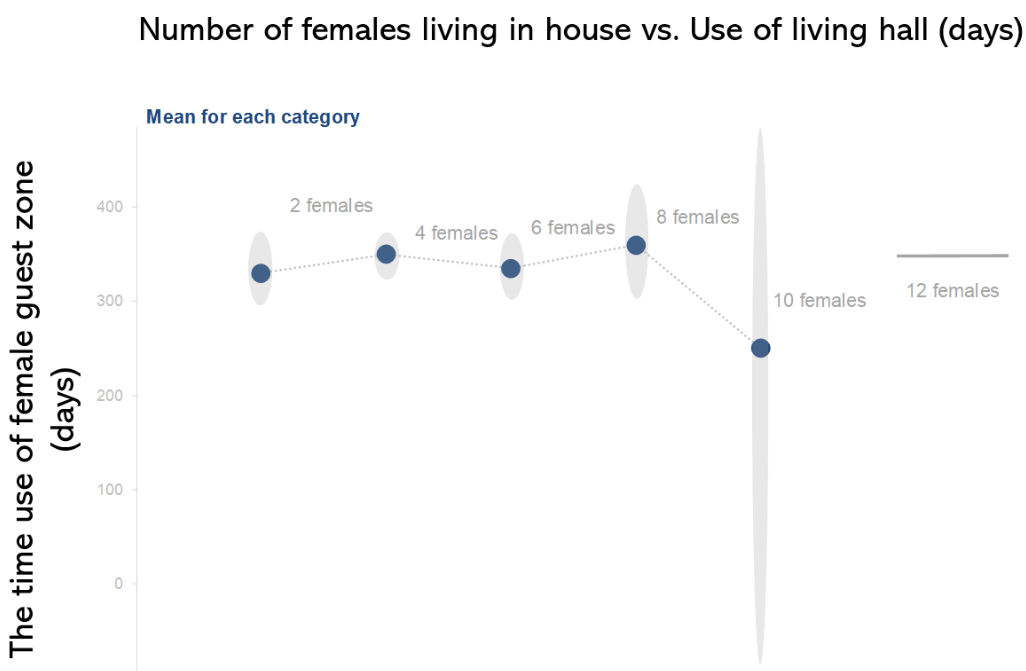


Fig. 5.17. The relationship between the number of female members and the living hall.

5.2.2.6. Question 6: When was the house built?

This question helped to understand changes in people's behaviour in terms of the use of the male zone, female zone, and living hall. The oldest house in this dataset was built in 1969 and the newest house in 2019. Fig. 5.18 shows that newer houses are more common; 68% of the houses were built between 1992 and 2014, and 95% between 1981 and 2019.

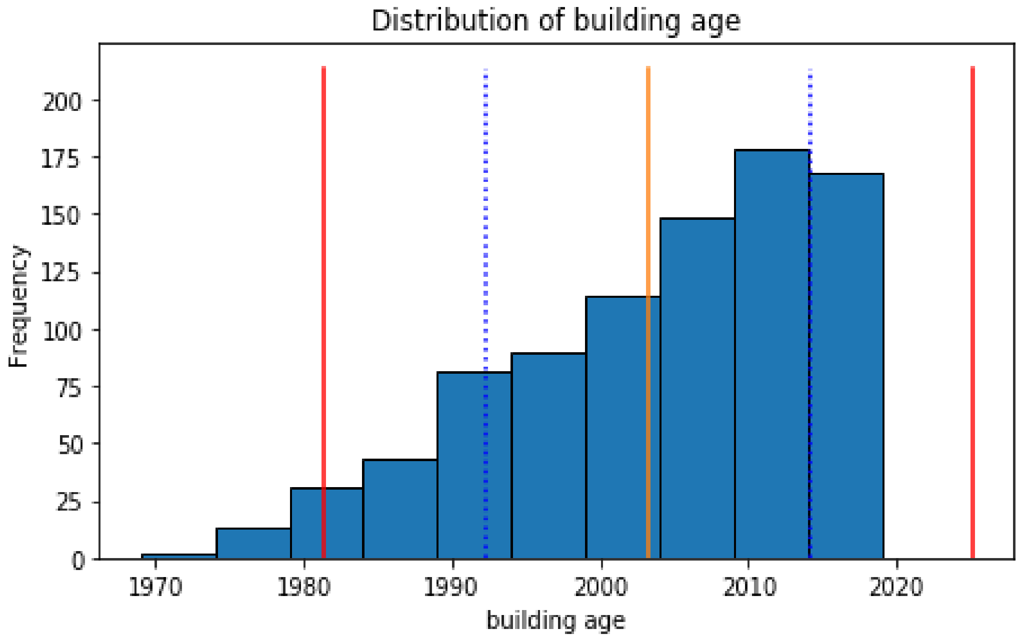


Fig. 5.18. Panel A is the frequency distribution of the building age, including the mean (yellow line) and the first (dotted blue line) and second (red line) standard deviation. Panel B is a cumulative frequency plot.

Fig. 5.19 is a chart that shows the mean for each 10-year bin. The relationship between the use of the male guest zone and the year that the house was built trends downward. Previous experiment results from case studies in Chapter 4 show that people living in homes built in the 1990s and earlier used guest zones more than they do now. People prefer to invite a friend out, perhaps for coffee or to a restaurant. Despite this, the size of guest zones was still comparable to older houses. The use of the female guest zone has a similar trend (Fig. 5.20).

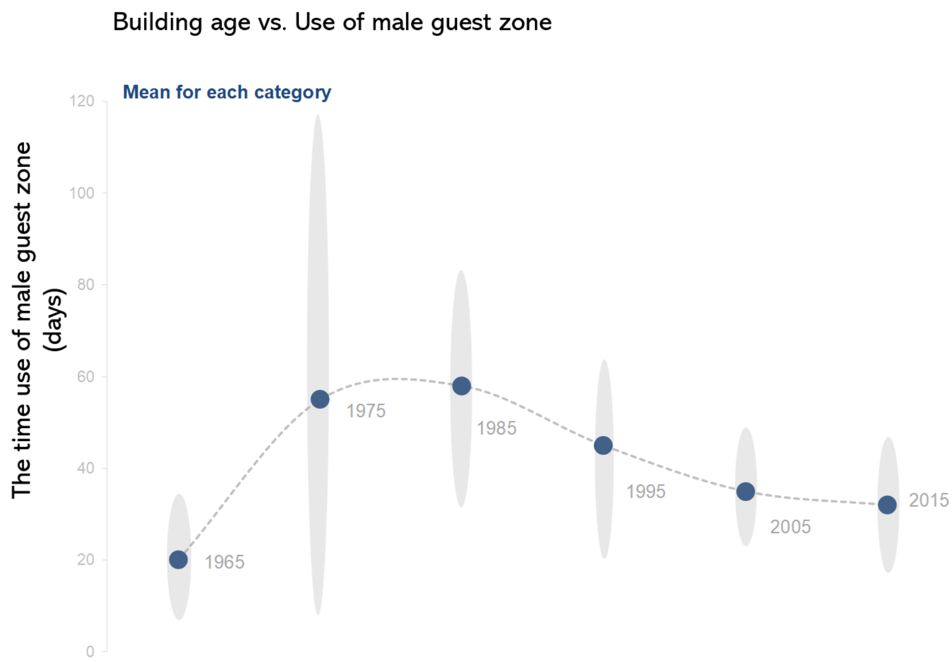


Fig. 5.19. The relationship between the building age and use of the male guest zone.

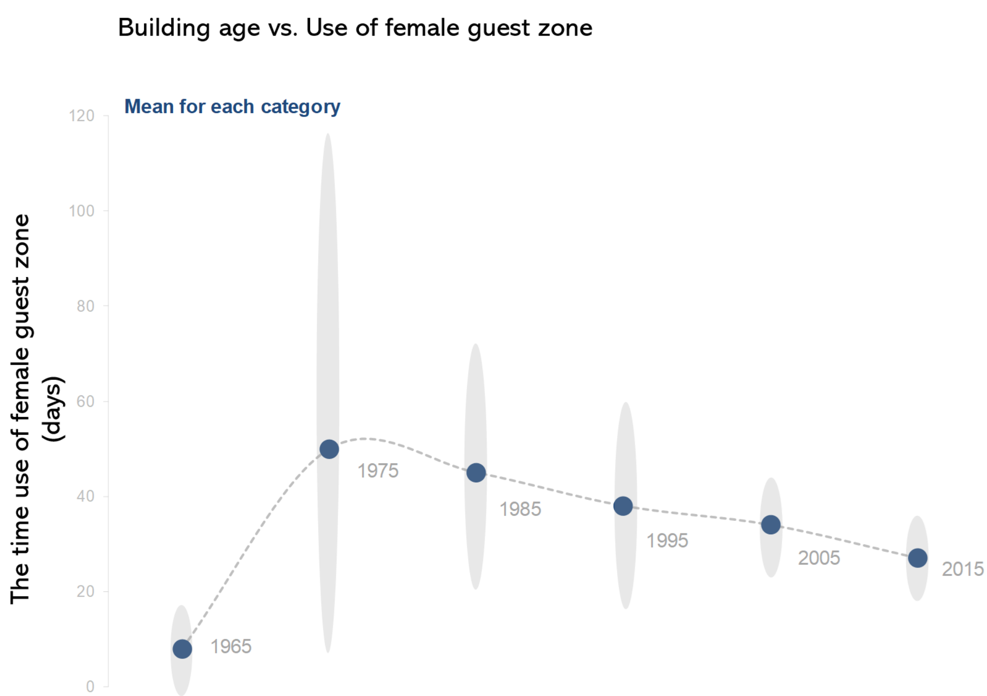


Fig. 5.20. The relationship between the building age and the use of the female guest room.

People who lived in homes built in 1985 used their guest zone on average 58 days a year, and people who lived in homes built in 2015 used it on average 32 days per year. The chart for the living hall (Fig. 5.21) shows slightly greater use of the living halls in newer buildings.

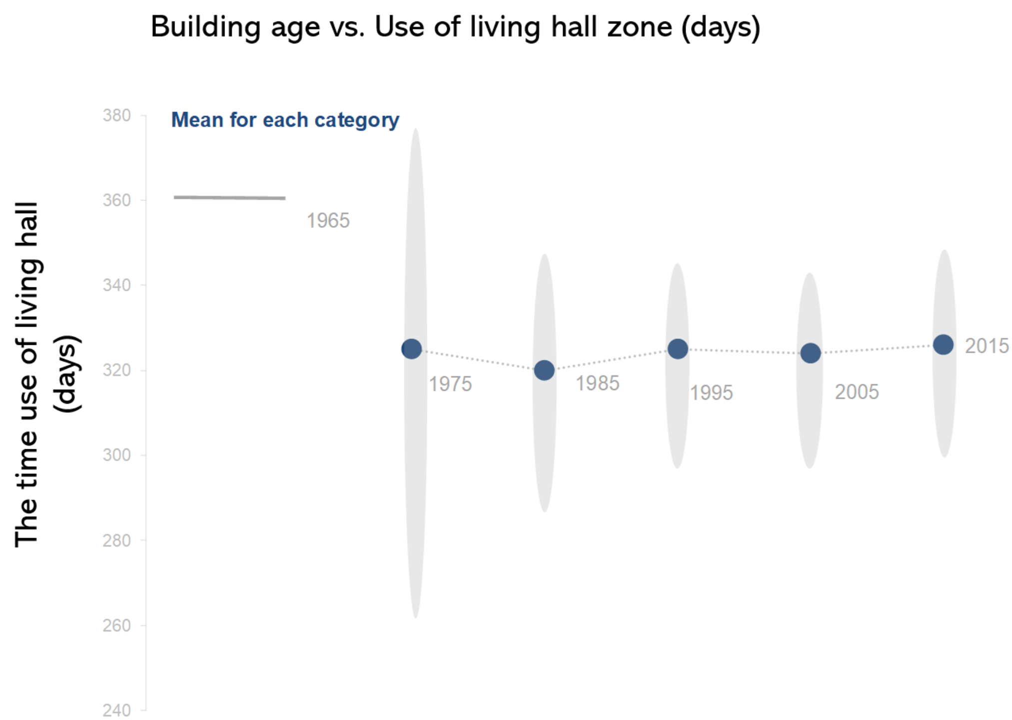


Fig. 5.21. The relationship between the building age and the living hall use.

This analysis answered some questions raised from the first pilot study (Chapter 4). In the pilot study, the comparison between houses based on 10-year bins shows that male guest zones are the same size compared to the total house size; however, the house had fewer rooms. These findings suggest guest zones were still important, but the zone's use was reduced, which is why homes more commonly had one big room instead of two smaller rooms.

5.2.2.7. Question 7: What is the land size for your house?

Fig. 5.22 shows the distribution of the land size data. The chart shows the highest frequency density to the left of the axis, suggesting smaller land parcels are more common. The largest is 10,000 m^2. The mean is 750 m^2 and the standard deviation is 939. The data range that covers 95% of the data is 120 m^2 to 2628 m^2. Fig. 5.23 shows the mean for every 400 m^2.

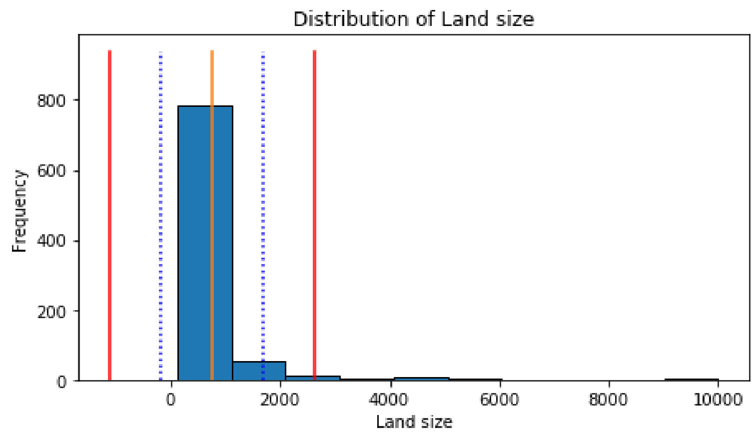


Fig. 5.22. Panel show the frequency distribution of the land size, including the mean (yellow line) and the first (dotted blue line) and second (red line) standard deviation.

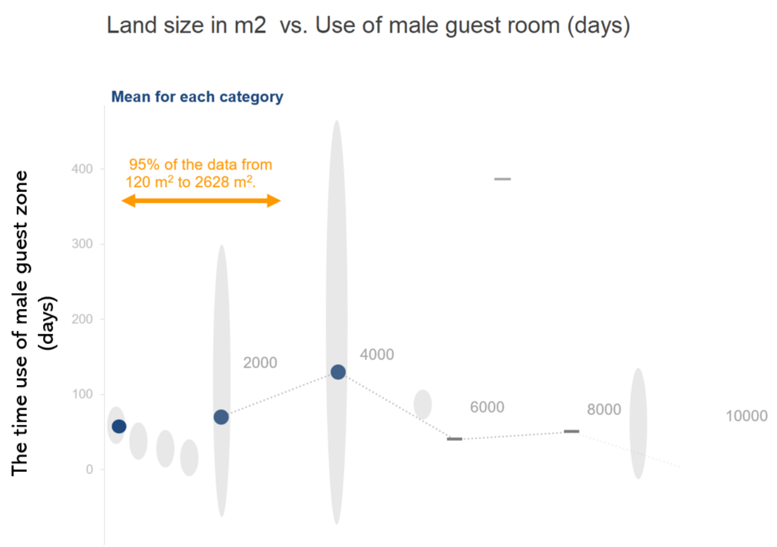


Fig. 5.23. The relationship between the land size and the use of the male guest room. Each bin is 400 m^2.

The trend for the first four bins (400, 800, 1,200, and 1,600) was decreasing in terms of zone use. For the 1600m^2 land size, behaviour showed an irregular trend. The land size had a positive relationship with the number of family members living in the house (Fig. 5.24). Big houses and fewer family members indicated lower use of guest zones. For example, a big house in the south would have higher use of the guest rooms. On the other hand, a big house in the north would have lower use of guest rooms.

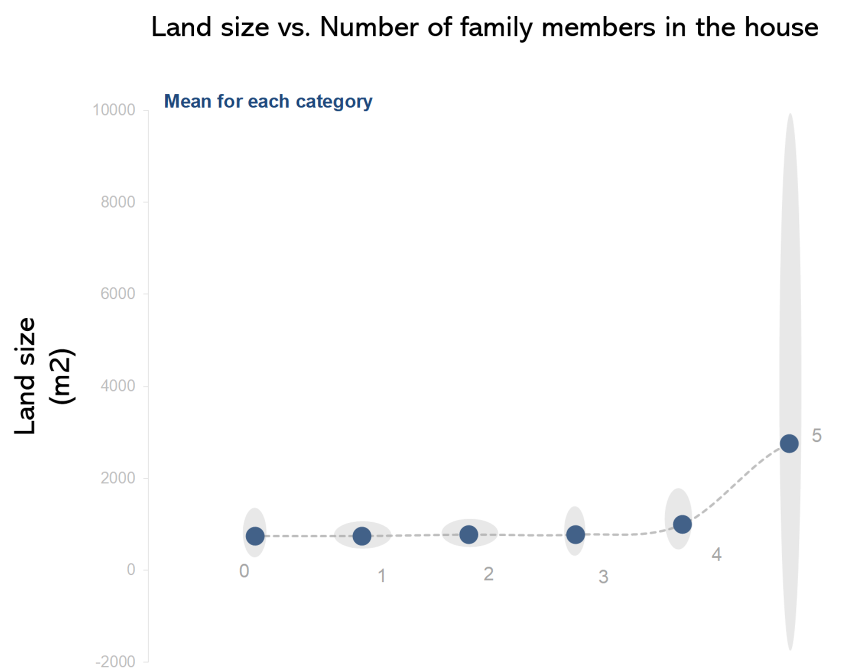


Fig. 5.24. The relationship between land size and the number of family members in the house.

These findings suggest house location had a big impact on the size and use of male guest zones (Fig. 5.25). As such, land size needs to be considered to understand the overall use patterns of guest zones.

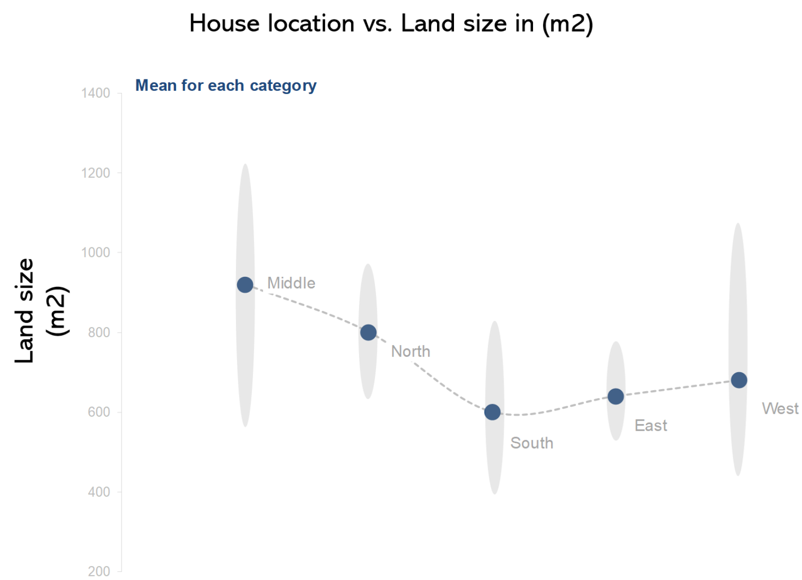


Fig. 5.25. The relationship between the house location and land size.

5.2.2.8. Question 8: How many days do you use the male guest zone each year?

The second section of the survey was the actual frequency use of targeted zones, including male guest zones, female guest zones, and living halls. Fig. 5.26 shows the data distribution of the time-use of male guest zones. Most participants used male guest zones just a few days a year. The mean use of male guest zones was 40 days a year.

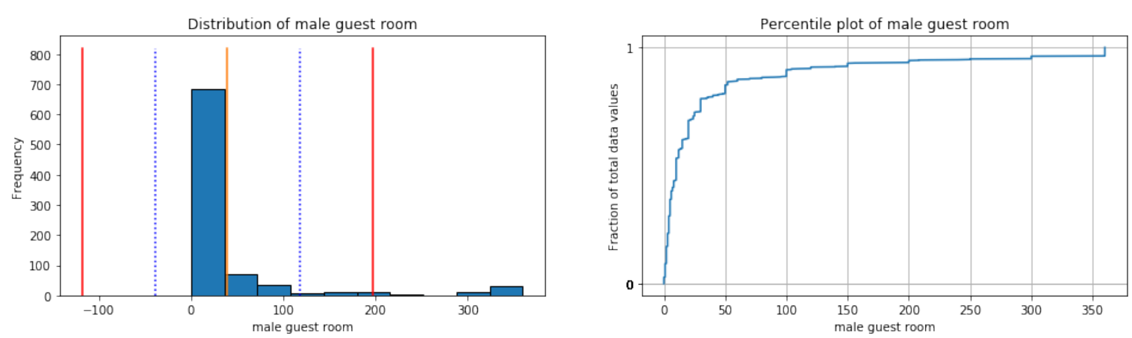


Fig. 5.26. The frequency distribution of the male guest room use, including the mean (yellow line) and the first (dotted blue line) and second (red line) standard deviation.

5.2.2.9. Question 9: How many days do you use the female guest zone each year?

Fig. 5.27 shows the data distribution for female guest rooms, which is similar to the male guest zone data. When male guests were in the house, there was a high likelihood that there were also female guests using the female zone. The data show female guest zone use was slightly lower than male guest zone use. The mean for the female guest zone was 33.7 times a year.

Female guest zone use was lower than the male guest zone use, with 198 out of 870 participants reporting no female guest zone use (i.e., they do not have a female guest zone). A total of 120 participants who did not have a female guest zone lived in a house smaller than 500 m^2. Thus, the size of the house had a big impact on occupant behaviour in the home. Houses that had one guest zone led to hosting male and female guests at different times, which resulted in more time-use for the guest zone because they were used for each gender on different days.

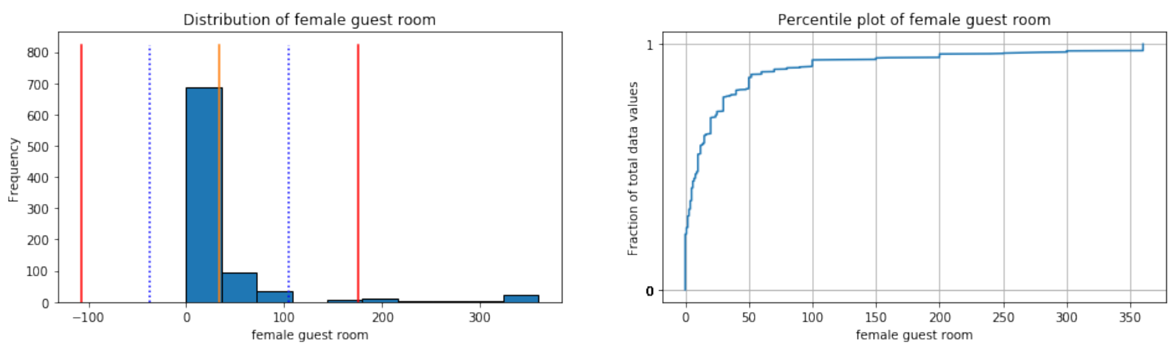


Fig. 5.27. The frequency distribution of the female guest room use, including the mean (yellow line) and the first (dotted blue line) and second (red line) standard deviation.

5.2.2.10. Question 10: How many days do you use the living hall each year?

Fig. 5.28 shows the frequency distribution of living hall use. The mean was 323 days a year. The data distribution shows two different types of living hall behaviour. Small houses did not tend to have a family room, so they used the living hall as a family gathering area (i.e., for daily use). Big houses had a separate family room and only used the living hall occasionally.

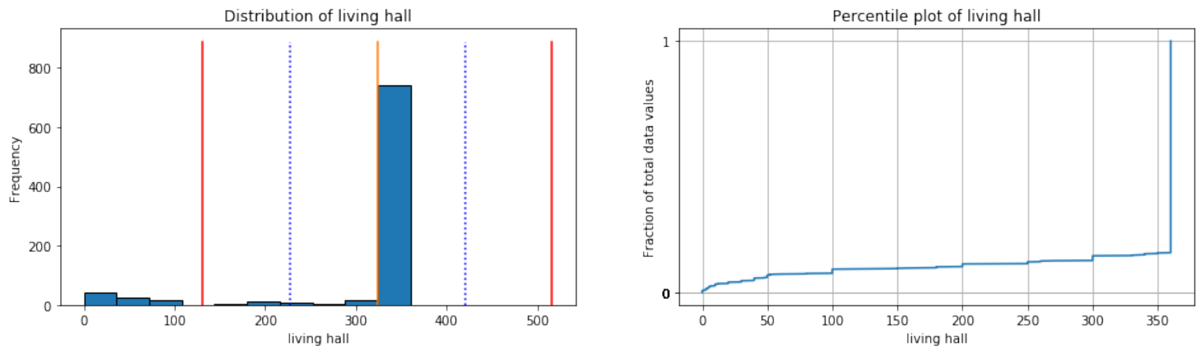


Fig. 5.28. The frequency distribution of the living hall use, including the mean (yellow line) and the first (dotted blue line) and second (red line) standard deviation. Panel B is a cumulative frequency plot.

5.2.2.11. Question 11: How many times you use the outside room?

The outside room was a room for gathering friends and is popular in Saudi houses. This question was added to this survey to identify if there was unique behaviour for this room. The average time-use for this room was 168 days and the standard deviation was 160 (Fig. 5.29). The first standard deviation encompasses all the days in the year. 321 participants stated they did not have an outside room, and 303 participants said they use their outdoor room for 360 days per year (daily use). The analysis shows that 193 participants do not have an outdoor room. Those who did not have an outside room tended to live in homes 500 m^2 or smaller, suggesting smaller houses do not have enough space for outside rooms, and big houses have outside rooms and use them daily.

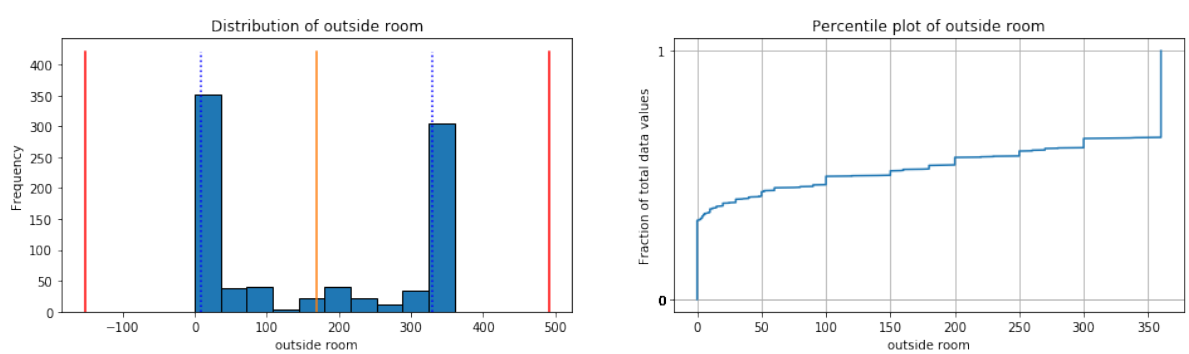


Fig. 5.29. The frequency distribution of the outside room, including the mean (yellow line) and the first (dotted blue line) and second (red line) standard deviation.

5.2.3. Summary

This chapter covered data and analyses for surveys conducted as part of this research. These data were used for the occupancy prediction model development. Table 5.21 shows a summary of each variable, including the minimum, maximum, mean, and standard deviation.

Table 5.21
Descriptive statistics.

	N	Minimum	Maximum	Mean	Std. deviation
Age	868	20	89	53.03	12.177
Male members	868	0	11	3.56	1.841
Female members	868	0	11	3.29	1.667
Building age	868	1969	2019	2003.25	10.918
Land size	868	120	10000	750.42	939.815
Male guest room	868	0	360	39.86	79.100
Female guest room	868	0	360	33.71	70.700
Living room	868	0	360	323.63	96.483
Outside room	868	0	360	168.94	160.947
Total family members	868	0	18	6.85	2.617
Valid N (listwise)	868				

Key findings from the survey data are as follows:

- The use of male guest zones is negatively correlated to home size (Fig. 5.24). For example, people in smaller houses in the southern part of Riyadh use guest zones more. The average size of houses in each location is, in descending order, the middle of Riyadh, north, west, east, and south.
- There is a positive correlation between householder age and male guest zone use. The youngest group (20s) used male guest zones on average ten times a year, and the oldest group (80s) used male guest zones on average 50 times a year.
- The data show that the more male members there are in a house, the greater the use of the male guest zone. However, this trend decreases with six or more male members in the house. These data suggest that homes with more than six people mean more than one family

live in the house. Multiple families in one house are typical in Saudi Arabia. For example, the father and his children live in the same house. However, multiple family households are not the focus of this research. People who live in multiple family houses were excluded from this study.

- The analysis also shows a positive trend for female occupants and time-use for female guest zones. However, the trend decreases in homes with eight or more female members.
- People who live in older houses tend to use guest zones more than people who live in newer houses. However, the size of the guest zone in new houses could still be the same compared to the older houses, regardless of house age. This shows that guest zones are still considered an important element in home design regardless of how often they will be used.
- Findings suggest that big houses do not necessarily mean a greater number of occupants or greater guest zone use. On the contrary, data show that larger houses tend to use guest zones less than smaller houses.
- For the friends' room or outside room, analyses show there is no unique behaviour for the time-use of these rooms. Homes either do not have these rooms, or they use them for normal daily use.

6. Prediction model for occupancy behaviour

Several studies have focused on occupants' behaviour using a time-use data (TUD) survey method for houses in different countries. However, no previous research has utilised the TUD method to predict time-use while also considering the social factors (culture and privacy) that affect the behaviour of house occupants in Saudi Arabia. For this reason, the data structure in this study was established as a database for other researchers, such that more data can be added in the future to predict time-use for different space-type structures in Saudi houses. The present study enables future researchers to improve the model by collecting more data for other areas not covered in this study, such as data for large houses or houses with multiple families. Chapter 7 discusses the limitations and boundaries of this research.

The final dataset, which was used to train and validate the model, contained five variables representing the parameters. Future researchers should follow the same data structure used in this study to replicate this study on larger sample sizes, or to investigate if the results can be replicated on different house types in Saudi Arabia. The structure for the variables was as follows:

1. (X1) Location of the house: This is a multiple-choice question; the respondent must classify the house into one of five locations (north, south, east, west or centre of Riyadh).
2. (X2) Age of the householder: The second input is open-ended and contains only numerical values.
3. (X3) Male members of the household: The input is numerical and open-ended; the householder should be included in the number of male members living in the house. Male employees or any male who does not use the male guest zone should not be included in the record of household members.

4. (X4) Female members of the household: The input is numerical and open-ended. Female employees or any female who does not use the guest zone should not be included in the record of female household members.
5. (X5) Land size: The input must include the total floor area of the house and the plot on which it is built, in numerical form, in m^2.

The last three variables (Y1, Y2, and Y3) in the database were the time-use data for the targeted spaces (male guest zone, female guest zone, and the living hall). The primary survey analysis for time-use in Saudi houses showed that the daily consumption gap in the energy software is more important to solve than the hourly resolution. Thus, this survey focused on daily use; the respondents' answers for the last three variables were between 0 to 360 days per year. The first column (Y1) was the respondent's answer for the time used in the male guest zone, the second column for the female guest zone (Y2), and the last column for the living hall (Y3). Usually, small houses in Saudi Arabia did not have a female guest zone, so they used the male guest zone for both male and female guests. In this case, when the house did not have a female guest zone, the respondent filled the column with zeros, meaning there was no female guest zone in the house. The data was exported in a CSV file format as preparation for the next step, which uses a machine-learning algorithm to train and validate the data.

6.1. An ML model to predict the time use frequency for rooms (male/female guest rooms and living hall)

6.1.1. ML model selection

Choosing the machine learning process that fits a particular data type is essential to obtain an accurate prediction model. Study data were collected from 861 respondents who answered eight questions: 1) location of the house, 2) the age of the householder, 3) the number of male members living in the house, 4) the number of female members living in the house, 5) the land size, 6) the number of days the male guest zone is used a year, 7) the number of days the female guest zone is used each year, and 8) the number of days the living hall is used each year).

A two-stage method was used to define the most suitable algorithm to predict the time-use of each of the targeted zones. The first stage consisted of using the study dataset for performance evaluation comparisons of all the regression algorithms available in the scikit-learn library (Fig. 6.1); followed by comparisons of performance outcomes for each zone (male and female zones and the living hall). The second stage excluded the models with a performance error higher than 0.7, then compared the remaining models again with a more detailed investigation of the selected algorithms.

6.1.2. Performance metric(s)

The K-fold cross-validation approach was used (Rodriguez, Perez & Lozano 2010) To compare and evaluate the performance of each algorithm. The K-fold cross divides the data into several folds, based on the number of instances contained in the dataset. The data was split into five folds (K = 5), based on the number of data points (861 respondents). The K-fold process initially used the first fold to test the model and the remaining four folds to train it. The second iteration used the second fold to test the model and the rest for training, and the same method was applied to the rest of the three folds, ultimately determining the error that represented each model's performance. In this case, 861 data points were divided into five folds: 80% to train and 20% to validate.

In total, 18 machine learning algorithms were tested from the scikit-learn library: OLS, Ridge, Lasso, ElasticNet, LARS, LassoLars, OMP, Bayes, GLR, SGD, PAR, Poly, KNN, Decision Tree, RF, SVR, XGB, and MLP. The data collected were used to predict each zone individually, using the K-fold cross-validation solution, which produced a total outcome of 54 prediction values. To measure the outcome, mean absolute error (MAE) was employed. This is a measure used to compare performance between prediction models and is one of the most common metrics used to validate and compare time series prediction models (Hyndman & Koehler 2006). The MAE measurement was applied to all 18 algorithm models for each zone. This process gave the average across the test samples of the absolute differences between outcome predictions and actual observations.

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n}$$

MAE = mean absolute error, Yi = prediction, Xi = true value, and n = total number of data points. In total, 54 regressions were executed using the 18 algorithms (OLS, Ridge, Lasso, ElasticNet, LARS, LassoLars, OMP, Bayes, GLR, SGD, PAR, Poly, KNN, Decision Tree, RF, SVR, XGB, and MLP). Each algorithm was used for the male guest zone, the female guest zone, and the living hall, respectively. Figs. 6.1, 6.2, and 6.3 compare the outcomes for each zone (male guest zone, Fig. 6.1; female guest zone, Fig. 6.2; and living hall, Fig. 6.3).

MAE outcomes – Male guest zone.

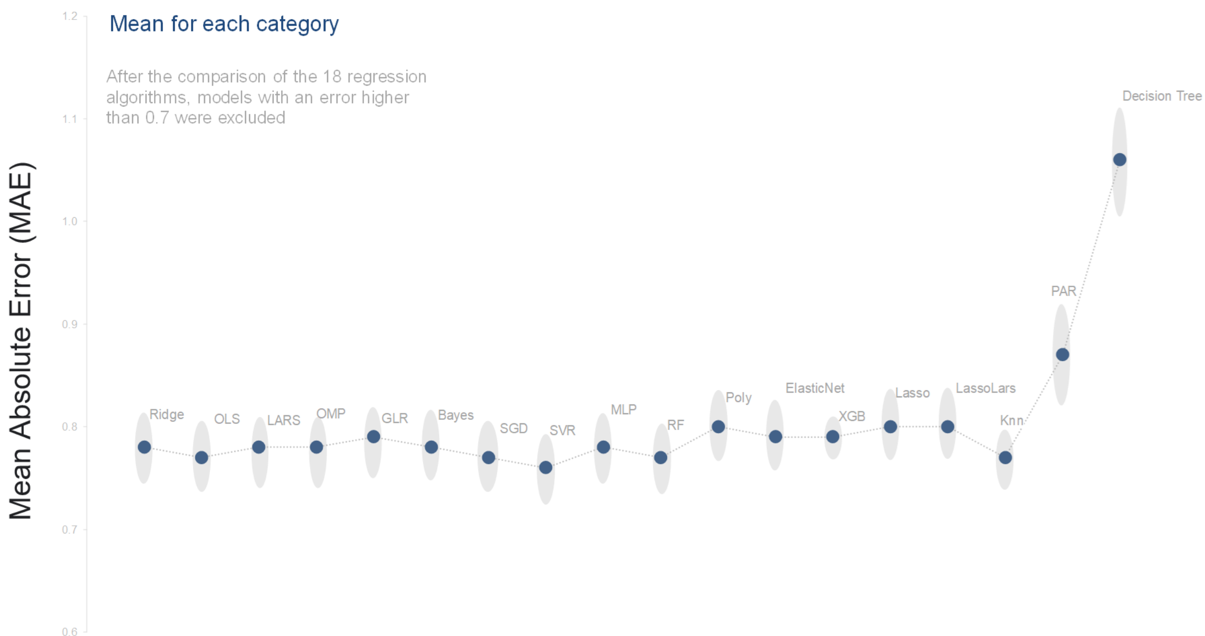


Fig. 6.1. The outcomes using the MAE for 18 algorithms for the 'people's behaviour' data type: male guest zone.

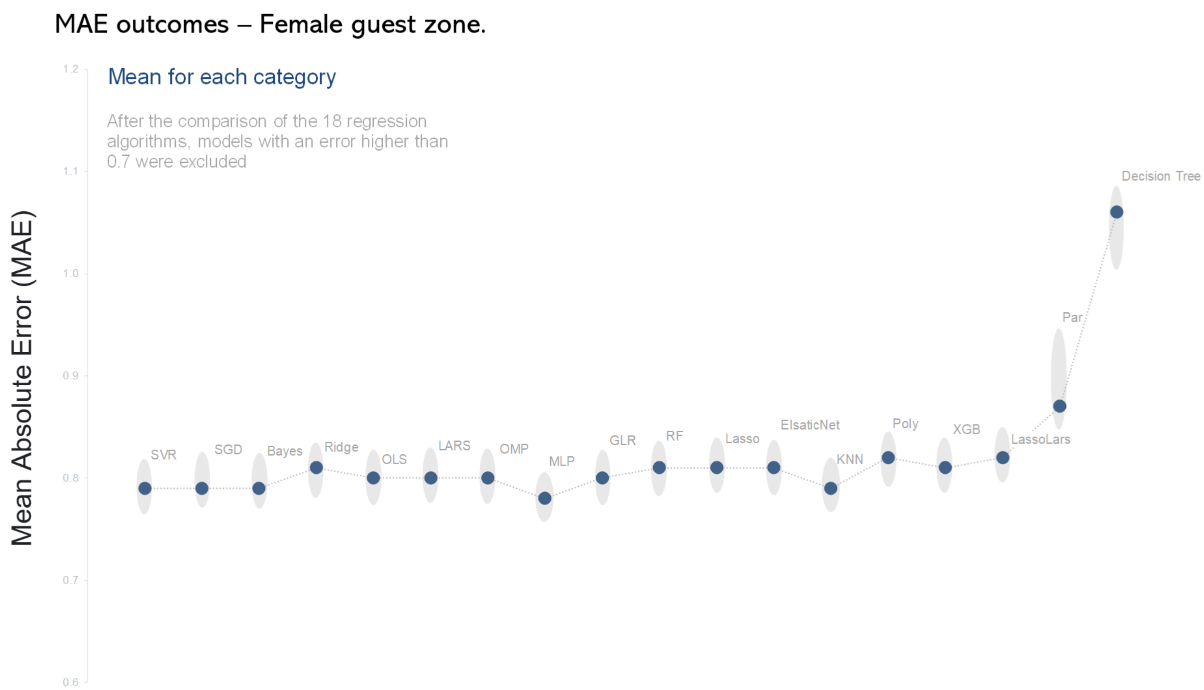


Fig. 6.2. The outcomes using the MAE for the 18 algorithms for the 'people's behaviour' data type: female guest zone.

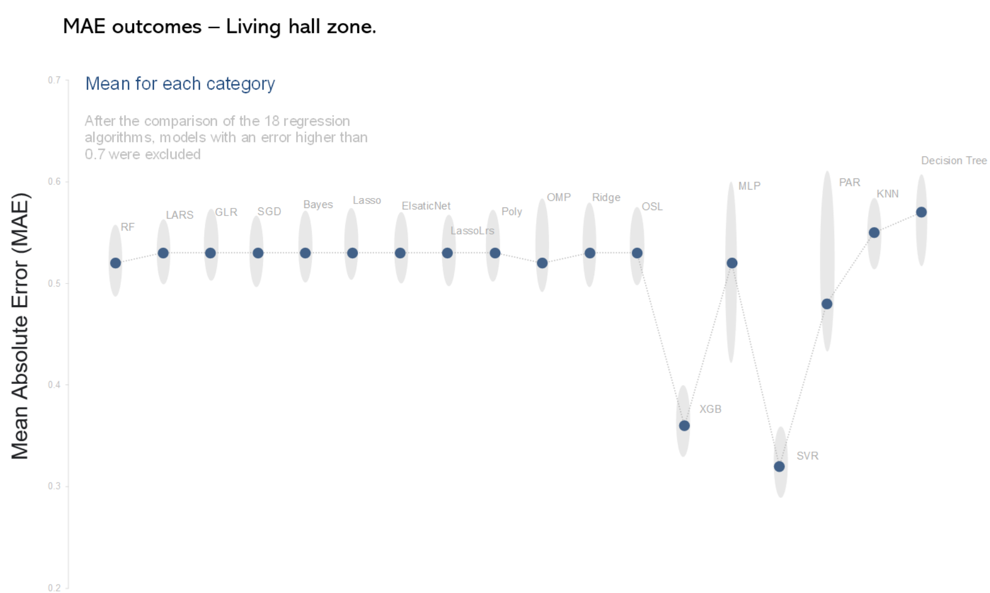


Fig. 6.3. The outcomes using the MAE for the 18 algorithms for the 'people's behavior' data type: living hall.

6.1.3. Data description and preliminary training results with multiple ML algorithms

After the comparison of the 18 regression algorithms, models with an error higher than 0.7 were excluded. The outcome using MAE demonstrated that nine algorithms were lower than the limit (Table 6.1): Ridge, Linear Regression, Lars, Orthogonal Matching Pursuit, Bayesian Ridge, SGD Regressor, SVR, Random Forest Regressor, and K-Neighbors Regressor. The researcher determined the Random Forest (RF) algorithm the most suitable to use in this study when compared to the other algorithms because of its following advantages:

- RF is a collection of decision trees that represent logical statements. Because of its learning ability, RF allows multiple predictions and a value for each prediction (Ahmad, Mourshed & Rezgui 2017). This advantage is useful for the statistical analysis model (Hong et al. 2017).
- RF gives the user the ability to adjust the prediction and see the confidence level of the new prediction. Data composed of social data (cultural behaviour) might encounter a missing input; for example, the age of the householder dataset. If the participants' ages were between 30 to 60 years old but there was no participant aged 45, the model can predict the behaviour of this age. RF can also overcome the problem of over-fitting (Ali et al. 2012).
- The data is composed of social data (age of the householder, number of male members, number of female members) and building information (building size and location). RF has the advantage of learning ability, as it is less sensitive to outliers and can define variable importance automatically. Thus, if more data were collected in the future and added to the previous dataset, the model will learn and adapt automatically (Ao et al. 2019).

Table 6.1
The outcomes using the MAE for the 9 most suitable algorithms for this data type.

Model name	MAE_Male guest zone	MAE_Female guest zone	MAE_Living hall
Ridge	0.78	0.81	0.53
OLS	0.77	0.80	0.53
Lars	0.78	0.80	0.53
OMP	0.78	0.80	0.52
Bayes	0.78	0.79	0.53
SGD	0.77	0.79	0.53
SVR	0.76	0.79	0.32
RF	0.77	0.81	0.52
KNN	0.77	0.79	0.55

6.1.4. Input variables considered for ML model training and comparison

The five variables used to run the three models (male guest zone, female guest zone, and living hall) are the location of the house, age of the householder, number of male members living in the house, number of female members living in the house, and land size. The plot-feature importance model compares the variables' performance in each model (Fig. 6.4). The feature importance model from the scikit-learn library is useful for non-linear estimators. It is a procedure that breaks the relationship between the feature and the target, so the model score is indicative of how much the model depends on each feature. For the first model, the male guest zone, the plot showed that householder age is the most important variable, followed by land size, the number of male members, and the number of female members. Location was the least important variable. For the second model, the female guest zone, the land size was the most important variable, followed by the number of female members, householder age, location, and the number of male members. The third model for the living hall showed that householder age was the most important, followed by land size, the number of female members, location, and the number of male members in the household (Fig. 6.4).

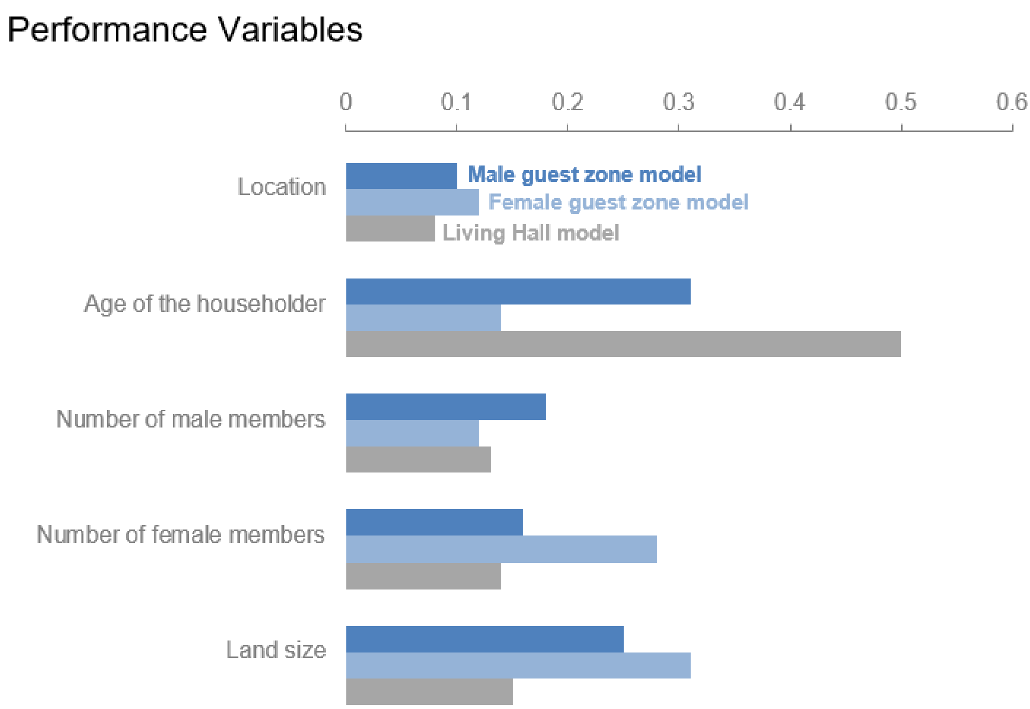


Fig. 6.4. Comparison of variables for all three models (male guest zone, female guest zone, and living hall).

6.1.5. Performance tuning for the selected algorithm (RF)

This section focuses on validating the RF algorithm performance of the three statistical analysis models: first for the male guest zone, then the female guest zone, and finally, the living hall. Random Forest is a forest consisting of many trees; the user does not need to be concerned about the back-end calculation (black box). However, there are six main parameters the user could tune to improve the predictive power of the model (Kane et al. 2014). The model runs a different value for each of the parameters and subsequently chooses the parameters that provide the best model performance prediction. The first parameter is to define the best number of trees (N-estimators; Table 6.2). The second parameter is the definition of the longest path, which is determined from the root node to the leaf node (Max depth; Table 6.3). The third parameter consists of the minimum number of samples to split an internal node (Min-samples-split; Table

6.4). The fourth parameter determines the minimum number of samples at a leaf node (Min-samples-leaf; Table 6.5). The fifth parameter consists of the maximum number of features the RF is allowed to try on an individual tree (Max-features; Table 6.6). Lastly, the sixth parameter is where the node will split if the split induces a decreased impurity which is higher or equal to this value (Min-impurity-decrease; Table 6.7) (Ullah et al. 2020; Xie et al. 2018). The Root Mean Square Error (RMSE) score is used to measure the error of the model for the evaluation (Barnston 1992), where:

$$\mathrm{RMSE}_{fo} = [\sum_{i=1}^{N} (z_{f_i} - z_{o_i})^2 / N]^{1/2}$$

Σ = summation ("add up")

(zfi – Zoi)2 = differences, squared

N = sample size.

Table 6.2
Random Forest models with a different number of trees.

n-estimators	Male guest zone model score (RMSE)	Female guest zone model score (RMSE)	Living hall model score (RMSE)
50	0.065	0.055	0.100
100	0.040	0.041	0.090
150	0.043	0.058	0.085
200	0.058	0.048	0.080
250	0.055	0.045	0.082
300	0.053	0.057	0.082
350	0.048	0.050	0.090
400	0.050	0.052	0.090
450	0.051	0.050	0.088
500	0.052	0.050	0.085

Table 6.3
Random Forest models with different max depth parameter.

Max depth	Male guest zone model score (RMSE)	Female guest zone model score (RMSE)	Living hall model score (RMSE)
RF with max depth =2	0.055	0.055	0.075
RF with max depth = 4	0.045	0.065	0.075
RF with max depth = 6	0.048	0.061	0.070
RF with max depth = 8	0.04	0.055	0.073
RF with max depth = 10	0.048	0.055	0.080
RF with max depth = 12	0.055	0.035	0.058
RF with max depth = 14	0.046	0.055	0.090
RF with max depth = 16	0.055	0.054	0.085
RF with max depth = 18	0.060	0.062	0.085
RF with max depth = 20	0.063	0.041	0.090
RF with max depth = 22	0.045	0.041	0.098

Table 6.4
Random Forest models with a different minimum-split number of internal nodes.

Min-samples-split	Male guest zone model score (RMSE)	Female guest zone model score (RMSE)	Living hall model score (RMSE)
RF with min split = 2	0.055	0.040	0.090
RF with min split = 4	0.050	0.043	0.080
RF with min split = 6	0.059	0.046	0.075
RF with min split = 8	0.050	0.050	0.070
RF with min split = 10	0.048	0.050	0.060
RF with min split = 12	0.050	0.046	0.060
RF with min split = 14	0.050	0.048	0.055
RF with min split = 16	0.040	0.057	0.058
RF with min split = 18	0.040	0.053	0.058
RF with min split = 20	0.042	0.059	0.089
RF with min split = 22	0.050	0.043	0.059

Table 6.5
Random Forest models with a different minimum-split number of leaf nodes.

Min-samples-leaf	Male guest zone model score (RMSE)	Female guest zone model score (RMSE)	Living hall model score (RMSE)
RF with min-leaf = 2	0.48	0.058	0.077
RF with min-leaf = 4	0.42	0.052	0.063
RF with min-leaf = 6	0.44	0.062	0.065
RF with min-leaf = 8	0.048	0.053	0.070
RF with min-leaf = 10	0.045	0.058	0.063
RF with min-leaf = 12	0.042	0.059	0.070
RF with min-leaf = 14	0.049	0.050	0.065
RF with min-leaf = 16	0.048	0.060	0.070
RF with min-leaf = 18	0.043	0.059	0.065
RF with min-leaf = 20	0.051	0.060	0.070
RF with min-leaf = 22	0.050	0.050	0.065

Table 6.6
The maximum number of features considered to make the split.

Max-features	Male guest zone model score (RMSE)	Female guest zone model score (RMSE)	Living hall model score (RMSE)
Auto	0.06	0.045	0.09
Sqrt	0.048	0.055	0.073
Log2	0.052	0.046	0.078

Table 6.7
Node split induces minimum impurity decrease.

Min-impurity-decrease	Male guest zone model score (RMSE)	Female guest zone model score (RMSE)	Living hall model score (RMSE)
Value 0	0.058	0.051	0.073
Value 10^{-1}	0.056	0.053	0.070
Value 10^{-2}	0.042	0.058	0.073
Value 10^{-3}	0.045	0.060	0.078
Value 10^{-4}	0.058	0.43	0.078

The best performing parameters in each model for the male guest zone, female guest zone, and living hall (Table 6.8) are as follows:

6.1.5.1. Male guest zone

The outcome showed that the best RF performance for the male guest zone was 100 estimator trees, as this gave the lowest RMSE score of 0.04. The longest path from the root node to the leaf node was 8 for the lowest error. The minimum number of samples to split an internal node was 16. The fourth parameter was the minimum number of samples at a leaf node, which was 4. The maximum number of features using square root was 0.05. The last parameter shown in Table 6.8 indicated that the best performance was 0.04 when the min-impurity decrease was 0.001.

6.1.5.2. Female guest zone

The outcome shows that the best performance of the RF for the female guest zone was 100 estimator trees, with the lowest RMSE score of 0.04. The longest path from the root node to the leaf node was 12 for the lowest error. The minimum number of samples to split an internal node was 2. The minimum number of samples at a leaf node was 22. The maximum number of features using square root was 0.045. The last parameter in Table 6.8 shows that the best performance was 0.04 when the min-impurity-decrease was 0.0001.

6.1.5.3. Living hall

The outcome of the RF model performance for the living hall showed 200 estimator trees, which had the lowest error at 0.08. The longest path from the root node to the leaf node was 6 for the lowest error. The minimum number of samples to split an internal node was 22. The minimum number of samples at a leaf node was 4. The maximum number of features using square root was 0.07. The last parameter in Table 6.8 showed that the best performance was 0.065 when the min-impurity decrease was 0.01.

Table 6.8
The pest parameters performance for the Random Forest models.

Parameters	Male guest zone model	Female guest zone model	Living hall model
n-estimators	100	100	200
Max-depth	8	12	6
Min-samples-split	16	2	22
Min-samples-leaf	4	22	4
Max-features	0.05	0.045	0.070
Min-impurity-decrease	0.001	10^-4	0.01

In this stage, the RF model was ready to be implemented in the time-use schedule tool to predict the time used for targeted zones. Table 6.8 demonstrates the best parameters for the RF model for each zone. The three models each had decision trees. Fig. 6.5 illustrates the decision tree from the RF model for predicting time-use for the male guest zone. Fig. 6.6 depicts the decision tree from the RF model for predicting time-use for the female guest zone. Fig. 6.7 shows the decision tree from the RF model for predicting time-use for the living hall. The trees shown below are only part of each decision tree for demonstration purposes, with just three estimators and maximum depths of 4. The three below show that the leaf nodes stop splitting when the MSE equals zero. The actual trees are more complex, as indicated in Table 6.8. The use of different colours represents the prediction outcome in days. Darker shades of colour mean closer to 360 days, and lighter shades mean closer to 0 days. The values were $\log_2(x+1)$ = number of days, 360 equalling approximate value 5.9 and zero value equalling 0 days.

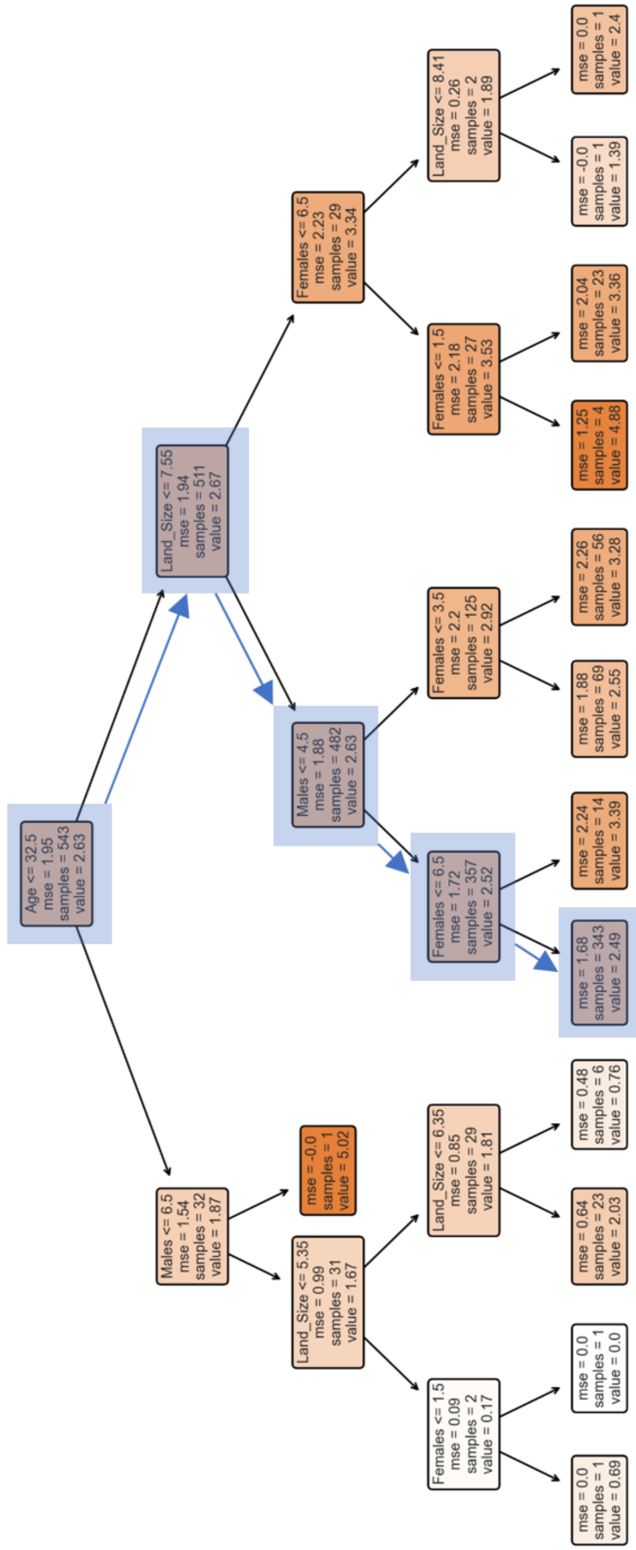


Fig. 6.5. Decision tree from a Random Forest model for predicting the time-use for the male guest zone.

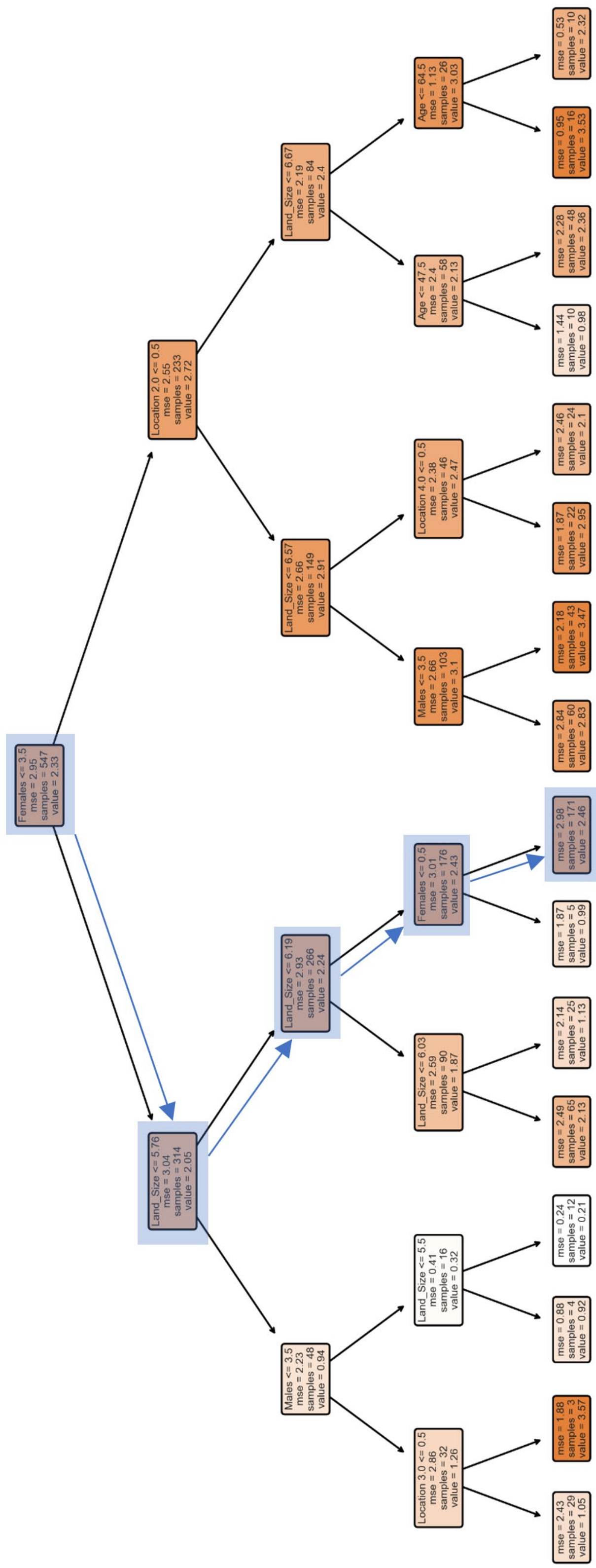


Fig. 6.6. Decision tree from a Random Forest model for predicting the time-use for the female guest zone.

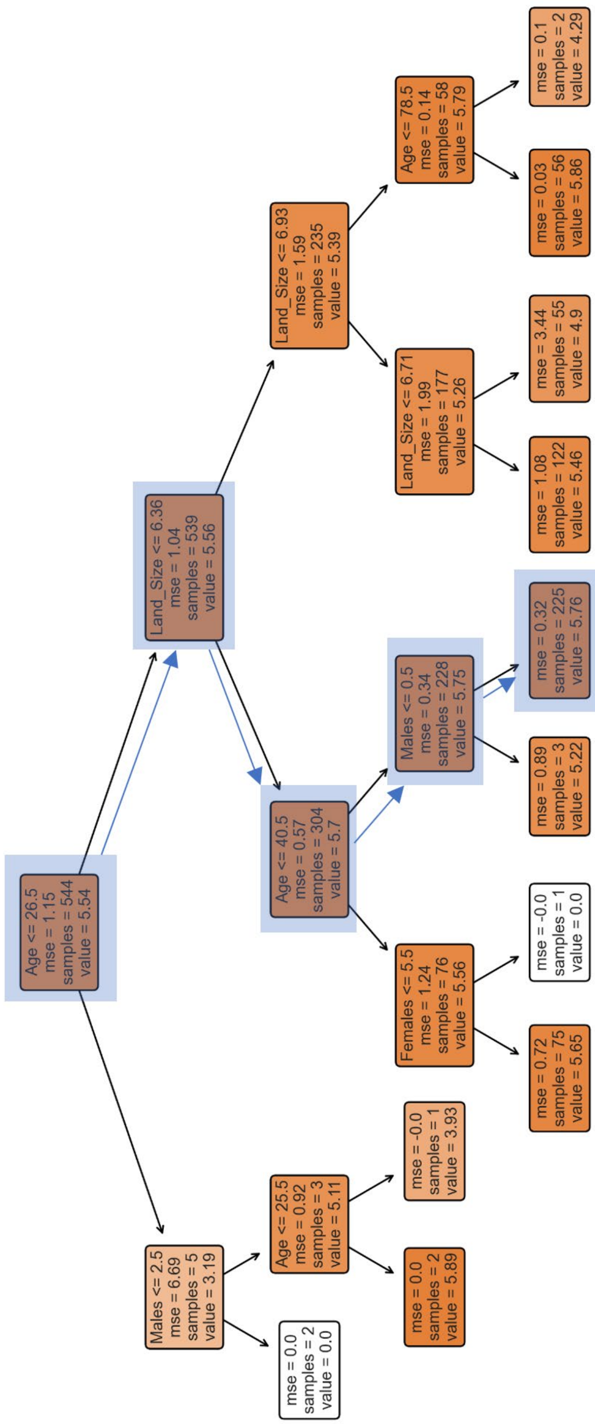


Fig. 6.7. Decision tree from a Random Forest model for predicting the time-use for the living hall.

6.2. Time-use frequency prediction model for energy performance simulation

6.2.1. The energy performance simulation workflow

The five main energy modelling inputs are weather, building geometry, HVAC systems, internal loads, and operation schedules (Fig. 6.8). The weather inputs are default data extracted from external repositories based on the building location. The user models geometry input, with the help of the simulation software. For the HVAC system, the user could design the whole system with its size and calculations. EnergyPlus allows the user to import this information directly from IFC files (Corrado & Fabrizio 2019).

Internal loads include occupancy patterns, activities, lighting, and equipment. They are also used as inputs for HVAC systems to calculate the occupation time and measure the energy needed to cool or heat zones. The HVAC load is based on internal and external temperature differences to a setpoint predefined in the RIUSKA database. The temperature setpoint can be adjusted manually to account for any unique behaviour in rooms.

One of the biggest challenges in any energy simulation is the time-use schedules of internal loads because occupant behaviour is constantly changing. Until today, there has been no tool that can represent actual time-use (Ding et al. 2020). Two additional considerations for measuring the life cycles of systems are the commission and operation. The more details a user can provide, the higher performance validation the software gives (Maile, Fischer and Bazjanac 2007). The researcher found occupant behaviour was the main contributor that influenced the building energy-performance gap (Choi 2017). With the currently available tools and technology, the best solution to solve the energy-performance gap problem is the grey-box approach. The advantage of the grey-box approach is the ability to integrate occupant-behaviour data into the physical model input with the ML model to improve occupant input accuracy (Laaroussi et al. 2020). The next section reviews the process of integrating the RF model, which was created early on in this research, with the energy simulation modelling inputs.

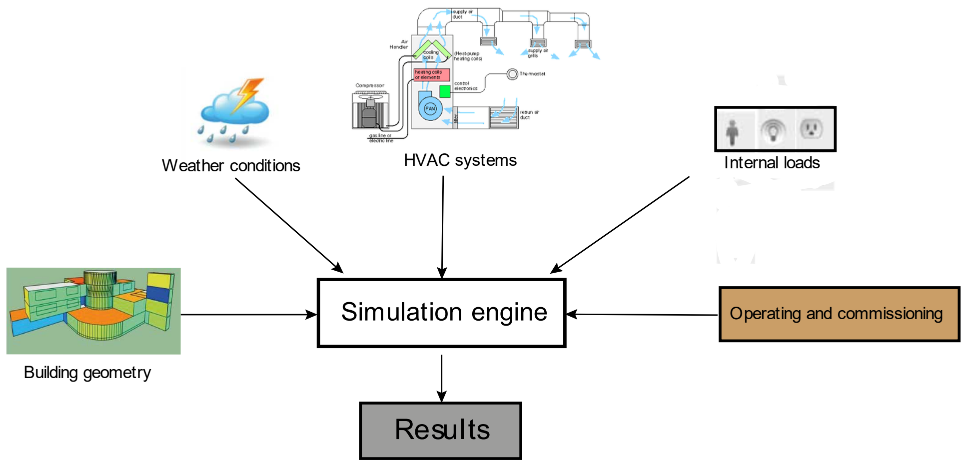


Fig. 6.8. General data flow of simulation engines.

6.2.2. Elaboration of the time-use frequency (TUF) model application

6.2.2.1. Time-use frequency prediction model

The Random Forest model was used in this study to predict time-use schedules for male guest zones, female guest zones, and living halls for single-family houses in Riyadh, Saudi Arabia. In doing so, the researcher provided the model with five main input parameters: 1) location of the house, 2) age of the householder, 3) number of male members, 4) number of female members, and 5) land size. Study survey data and the RF model (as described earlier in this chapter) were used to predict the daily time-use for each zone. The hourly time-use of targeted zones were fixed schedules based on the primary survey data. Fig. 6.10 shows how OpenStudio represents the hourly and day/weekly schedules.

There are multiple steps involved with the integration of the engineering simulation with the ML model. The first step requires users to develop the energy model using any energy simulation software that supports EnergyPlus. This research used OpenStudio software for energy modelling. The second step involves highlighting the rooms under each category; this study included rooms in the male guest zone, female guest zone, and living hall. Using the RF model, the next step is to enter the five parameters required for the RF model to predict the time-use for each zone. After entering the parameters, the next step is to run the RF model and generate the daily time-use for each zone in the year. The final step involves replacing the old schedules with new predicted schedules for the targeted spaces (Fig. 6.13). The tool is then integrated with the energy model after the projected model, using any energy simulation software that supports EnergyPlus. Fig. 6.9 shows the integration process between the energy software and the tool. Step 1 uses information from existing data sources to create individual building energy simulations. Step 2 employs machine learning to predict the time-use for the targeted zones and replace the old time-use schedules with the new predicted schedules.

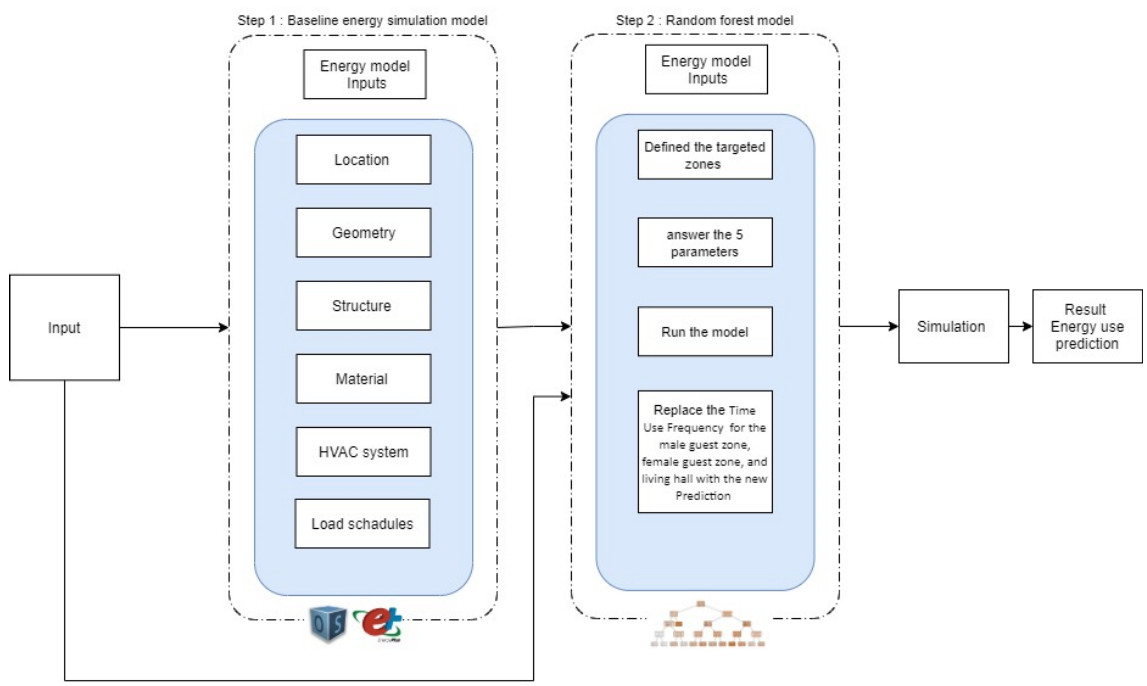


Fig. 6.9. The framework of the integration between energy simulation and machine learning.

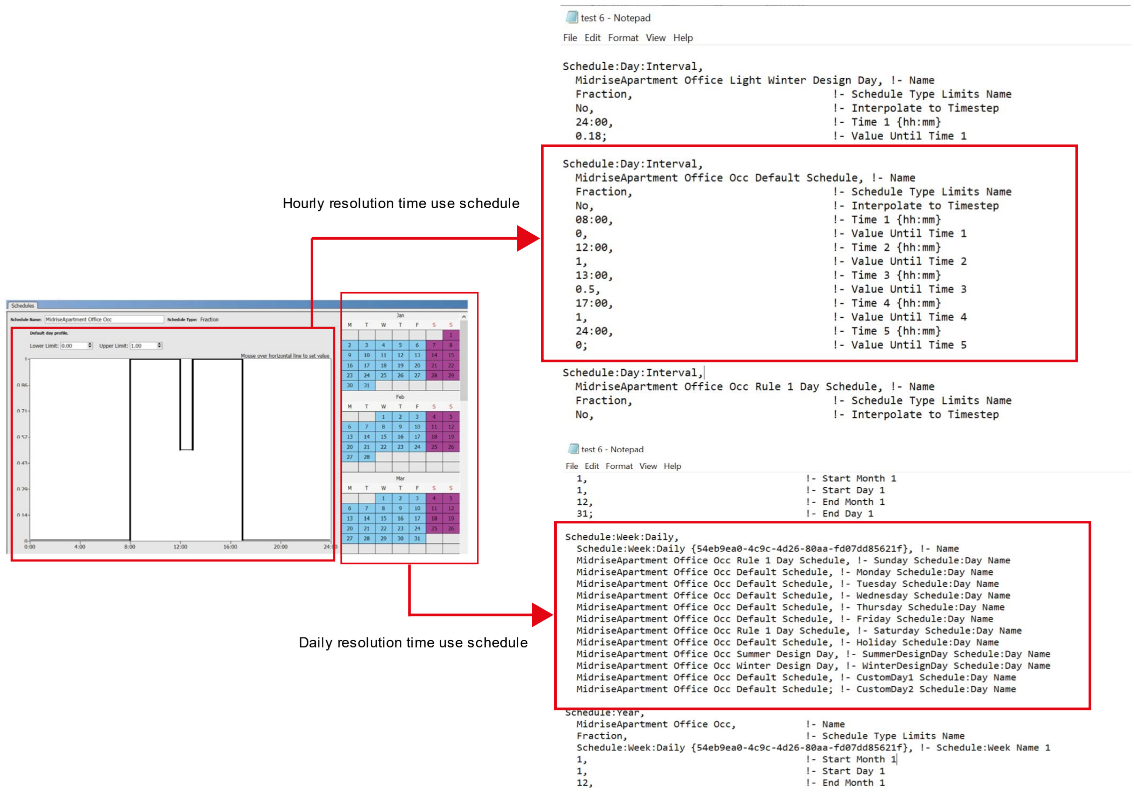


Fig. 6.10. EnergyPlus energy model as a text file.

The EnergyPlus data inputs include the load schedules for occupants, lighting, equipment, heating and cooling, hourly resolution, weekly resolution, and day schedules for the whole year. In this study, the researcher focused on predicting daytime-use resolution for the whole year. The prediction focused on the daytime-use resolution because of the lack of information related to the daily use of male guest zones, female guest zones, and living halls. For the hourly resolution, the tool used fixed hourly-resolution generated from the TUD primary survey for the male and female guest zones (Fig. 6.11) and the living hall (Fig. 6.12). The fixed schedules were applied to the number of days the RF model was predicted to be occupied.

The RF model predicted the number of days (resolution) rooms were occupied. The first survey showed higher time-use demand for guest zones during the weekends compared to weekdays. The daytime-use distribution of the prediction filled the weekends first, then the weekdays. For example, if the model predicted 12 days of use for the male guest zone in a year, then the model would distribute the days to fill the first day of the weekend for each month.

OpenStudio software uses Building Component Library (BCL) for load schedules. In the library, the living room space-type schedule for the residential building type is operational all days of the year (360 days). In addition, there are no male or female guest room space types in the library. Based on the survey analysis, people in Riyadh use male guest zones for, on average, 40 days per year in comparison to 70 days per year for female guest zones. The living hall is occupied for an average of 227 days per year (more information on the time-use survey data is given in Data and Analysis, Chapter 5). The hourly resolution (Fig. 6.11) shows a comparison between the default time-use fraction with the improved schedules from the primary survey outcomes. This demonstrates that the gap does not differ greatly between the people's answers regarding how many hours they use the room and the default schedule. The average time-use in both default and predicted models was 6 to 8 hours, so the researcher focused on improving the daily-use prediction and used fixed hourly schedules generated from primary survey answers.

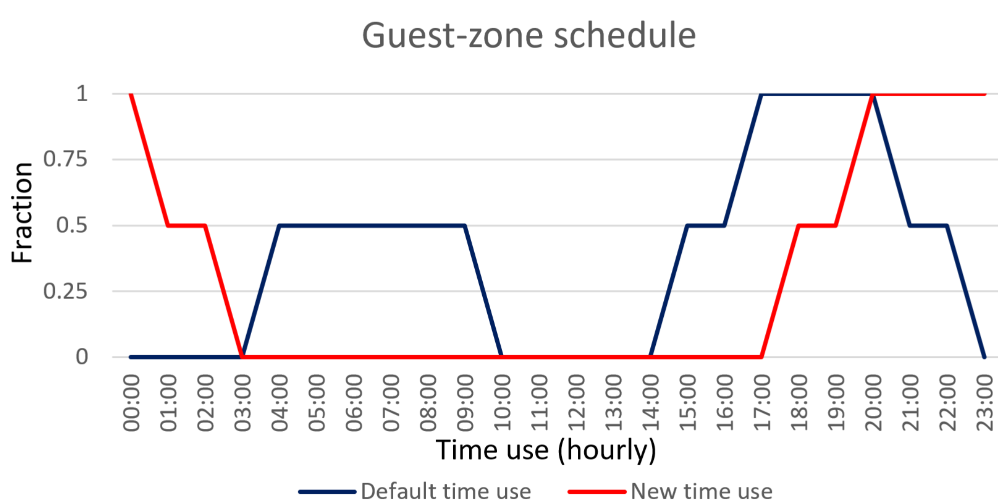


Fig. 6.11. The hourly time-use fraction from the TUD survey and the default hourly use.

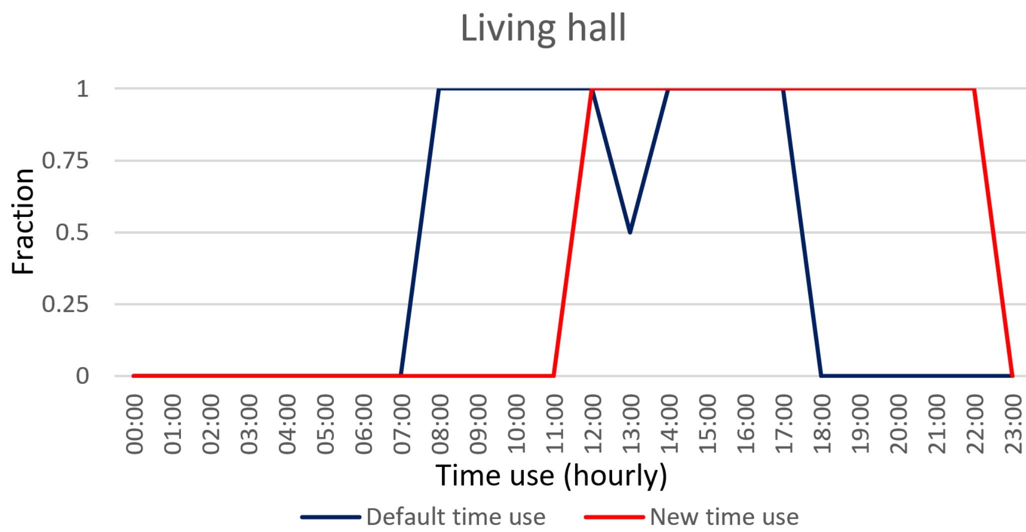


Fig. 6.12. The hourly time-use fraction from the TUD survey and the default hourly use.

6.2.3. Application tool UI and workflow diagram

As mentioned previously, this chapter covers the steps taken to develop a load-schedule prediction model for Saudi house occupancy. The goal of the occupancy model is to improve load schedules for the male and female guest zones and the living hall. A secondary goal is to make the tool user-friendly and easily accessible for any user with an average knowledge of energy-simulation software. This section reviews the tool's processes and features. To be able to use the tool, the user first needs to model their project using an energy modelling software like OpenStudio, which supports whole-building energy modelling using EnergyPlus. The process of the tool from "in" to "out" is covered in six main steps. Fig. 6.13 illustrates those steps: 1) import the IDF file, 2) define the spaces that need improvement (i.e., male guest zone, female guest zone, and living hall), 3) answer the five parameter questions, 4) generate the time-use prediction using the RF model, 5) adjust the time-use prediction, if needed, and 6) replace the defined spaces with the new schedules in each category and generate the IDF file with the newly updated schedules. A more detailed overview of each step is discussed next.

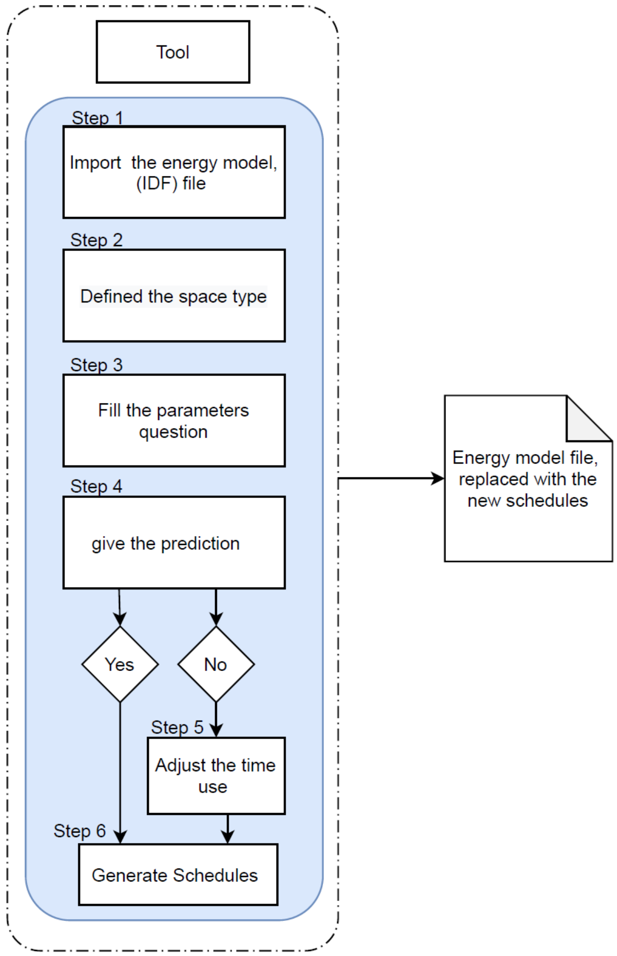


Fig. 6.13. How the tool processes the energy file from 'in' to 'out.'

The first step involves importing the energy-model file for a project that should be already fully modelled using energy simulation software. The file should be in IDF format. The IDF file format sorts all energy model input in a text file that can be read by any energy software. It is usually used to export an energy model from one software application to another (see Fig. 6.14).

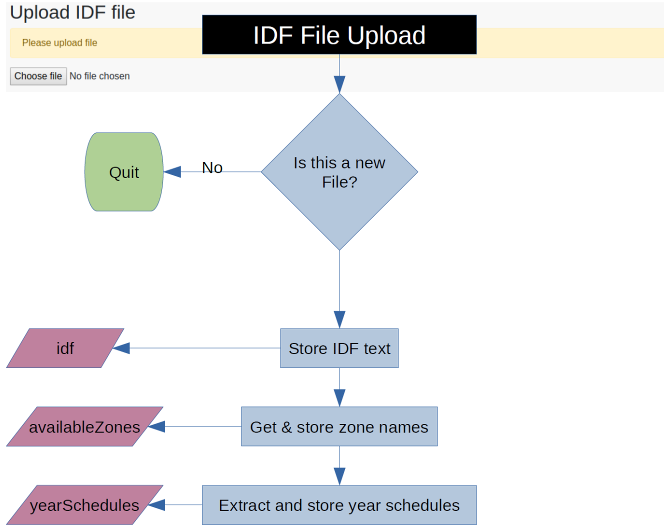


Fig. 6.14. First step diagram (import the IDF file).

The second step involves defining the zone. The tool only includes load-schedule inputs in the IDF data. The tool also shows the user all the thermal zones named in the project. The user needs to remember the names in order to categorise the zones. Choose the names of the spaces that come under the male guest zone, the female guest zone, and the living hall (see Fig. 6.15).

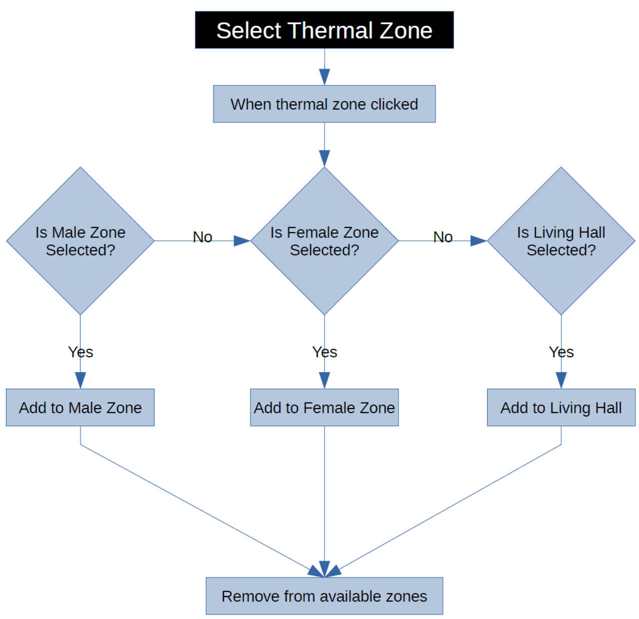


Fig. 6.15. The space-type section.

The third step involves generating the prediction schedules, which includes the machine learning model. Fig. 6.16 illustrates the processes in this step. The time-use schedule tool is run behind the tool for the user, using WebStorm (WebStorm: The Smartest JavaScript IDE by JetBrains n.d.). It is linked to the server PyCharm development environment tool (JetBrains, n.d.) for the data pipeline and the RF algorithm to generate the prediction schedules (see Fig. 6.16). To predict the new load schedules, the user needs to complete the five parameter questions: 1) location, 2) age of the householder, 3) number of male members, 4) number of female members, and 5) land size. The tool will then generate the time-use prediction for each zone and develop the new schedule template.

For certain rooms that have unique occupant behaviour, the tool allows for adjustments. A feature had been added to the tool to adjust the time-use prediction of these rooms. As mentioned previously, one of the advantages of the RF algorithm is that the model gives the confidence level of the prediction based on the data. A kernel density function was also added to the model, giving the user the ability to see the confidence level for any point over the course of the 360 days. This is presented in a bar from 0 to 360 days (see Fig. 6.17), with the confidence level indicated by colour and number. The tool automatically chooses the highest confidence level for each zone, but the user can adjust this if necessary. The last step is to generate the energy model with new load schedules for each zone. The tool will automatically make updates with the new schedules and download the new IDF file to the user's computer.

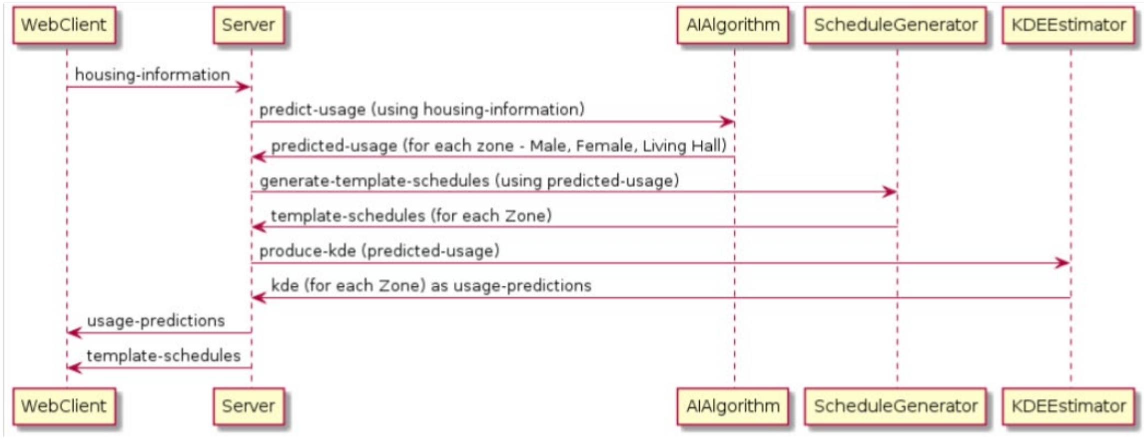


Fig. 6.16. The tool process diagram.

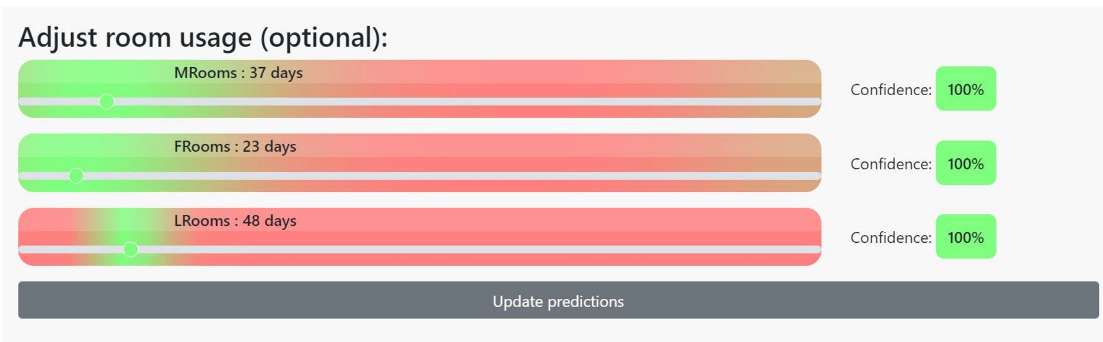


Fig. 6.17. Daytime-use prediction and adjustment.

The data structure for time-load schedules (people, lighting, equipment, cooling and heating) in the energy model file (IDF) follows this order: hourly, daily, then weekly for the whole year (52 weeks). For the energy model to run, the prediction schedules must follow the IDF-schedule template stipulated for EnergyPlus. For each space type, the file must have time-use schedules for hourly-use fraction units, then weekly for the year. This template must be repeated for each load related to any particular space type in the house. As a result, several schedules are produced for each case study. For example, if the model predicts 12 days of time-use for the male guest zone, then each load has two fixed-day schedule hourly fraction values. The first schedule includes the time-use data generated from the primary survey. The second is an empty time-use schedule (all hours are zero fraction values). Then each day of the weekly schedule refers to one of the two schedules based on the daily prediction model. In this example, one day of each month is occupied in the first weekend (Fig. 6.18). This schedule will be repeated in the first week of each month and the rest of the week schedules will use empty daytime-use schedules.

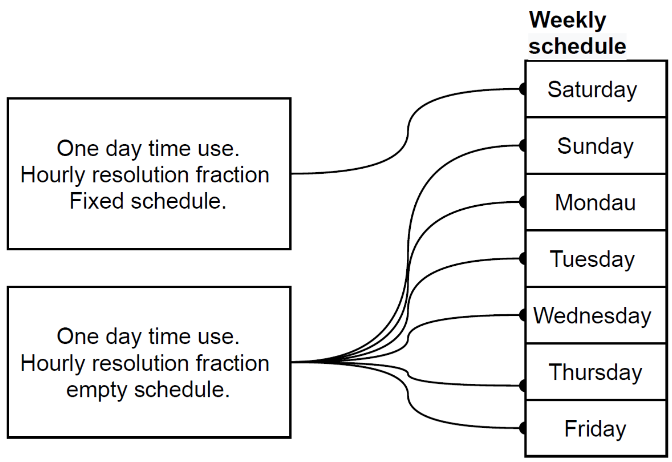


Fig. 6.18. One-week time-use diagram.

6.3. Case studies

This section reports the individual test findings of the five case studies. Two experiments were conducted and compared for each of the five case studies in order to validate the performance of the tool. The process for the first experiment started by creating an energy model for each case study. The simulation was run using input data provided by the householder (except for the time-use schedules and the weather data; these two inputs were provided by the BCL of the software) and the energy use prediction outcome compared with the householder's real electricity bill (without using the occupant model tool) (Fig. 6.19).

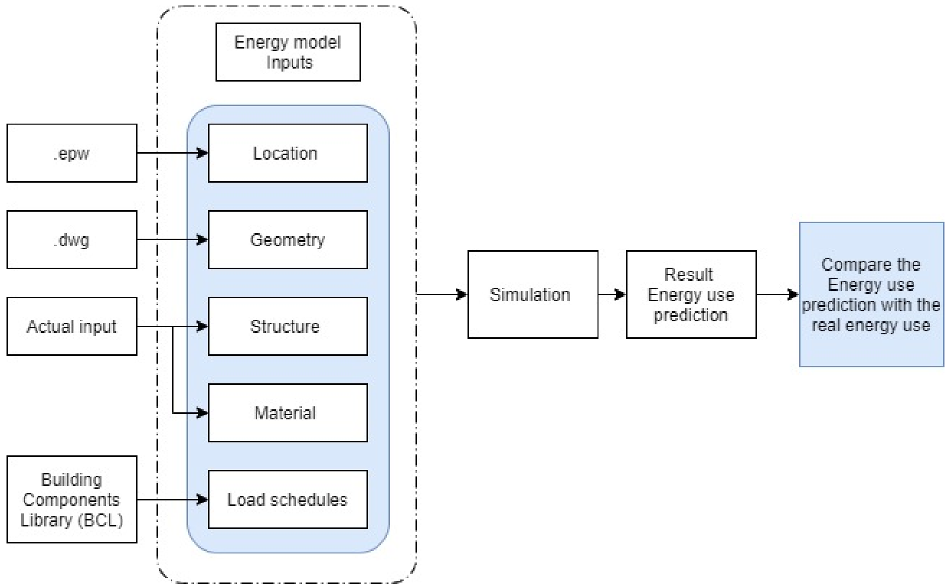


Fig. 6.19. Experiment 1 process.

The second experiment was conducted to determine how much the RF model could improve the performance of the prediction and close the gap between energy prediction and actual energy use. This was achieved by importing the energy models made in the first experiment into the time-use schedule tool. Then, the five parameters for the case study were inputted into the RF model. Householders provided all case study details, including house plans, building details, electricity bills, and the five parameter questions that were used as input for the model to predict time-use schedules (location, householder age, number of male members, number of female members, and land size) in order to generate the new occupant schedules.

A new time-use prediction for the targeted zone was then generated. Then the time-use schedules for the male and female guest zones and the living hall were replaced with the new predicted schedules. The last step involved running the simulation and comparing the outcome for experiment 2 with the real energy-use and the first experiment to define how much the model closed the gap (Fig. 6.20). These two outcomes were compared with the actual electricity bill, and the process was repeated for all five case studies.

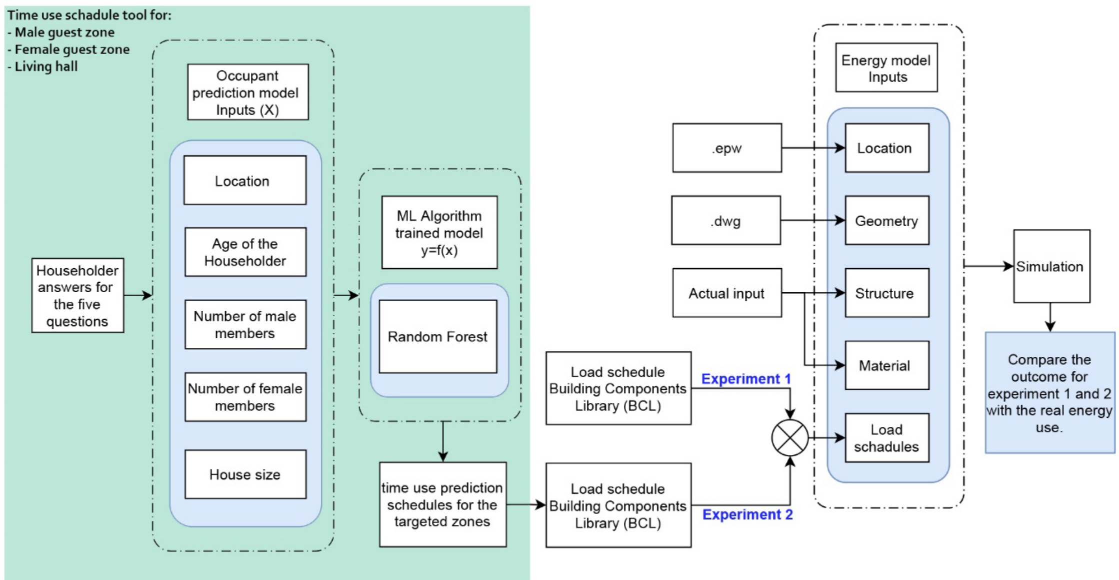


Fig. 6.20. The method used for processing the experiment.

Case studies were purposefully selected to represent each of the main areas of Riyadh city (north, south, east, west and central). The researcher also selected case studies in which all parameters were in the first standard deviation range. The researcher chose this range to be close to the mean of the dataset (higher data density). The more data samples the model has from similar inputs, the higher the model accuracy, which means a higher confidence level in the prediction outcome. As noted earlier, the householders provided all case study buildings and occupant details. OpenStudio software and EnergyPlus were used to run the experiments. OpenStudio software was used to create the energy model for the first experiment. For the second experiment, the energy model improved using the time-use schedule tool, and EnergyPlus was used to run the simulation. Fig. 6.20 shows the different processes for data input preparation between the two experiments.

6.3.1. *First case study*

6.3.1.1. *House information*

The first case study was for a young family; the householder's age was 43. The house was located in the northern area of Riyadh city. The house size was 250 m^2, which is the smallest area allowed by law for a single-family house in Saudi Arabia. Five people lived in the house: three males and two females, including the parents. Based on survey data results (as reported in Chapter 5), five occupants is the average size of household for the size and location of the house. Because of the house size, there is just one guest zone, used by both males and females. Usually, in cases like this where there is no female guest zone or family room in the house, the occupants use the living hall for the daily family gathering, and also for female guests if the male guest zone is occupied. In this case study, 32% of the total built area of the house was for visitors (the male guest zone and living hall). The answers for the five parameters were as follows:

Parameters	Location	Age of the householder	Male members	Female members	Land size
First case study	North	43 years old	3	2	250 m^2

The space types for this case study were as follows:

Y1	Male guest zone: Reception room, dining room, and bathroom.
Y2	Female guest zone: NONE.
Y3	Living hall: Ground-floor living hall and first-floor living hall.
Daily use	1 storage room, driver’s room, 2 ground WC, kitchen, 1 master bedroom, 3 bedrooms, 3 first floor WC, housekeeper’s room, laundry room, housekeeper’s WC

6.3.1.2. Model prediction for the first case study

The RF model (developed and described earlier in this chapter) was used to generate time-use predictions for the three zones (male and female guest zones and living hall). The five parameters: 1) location, 2) householder age, 3) number of male members, 4) number of female members, and 5) land size were inputted into the model. Fig. 6.21 shows the graphic interface for the prediction outcome, the parameters section, and the case study's prediction outcome. It also illustrates the confidence level by colours and per cent. The user can change the number of days within the range from 0 to 360 days, and the model will provide a confidence level for the new adjustment.

The first model prediction was for the male guest zone; in this case, the model predicted that the male guest zone would be used for 12 days each year. The model then predicted that the annual use for the female guest zone would be four days. When the model predicts less than 12 days (as in this case with the female guest zone), the model prediction output indicates that the house does not have a female guest zone. The model reports output in this manner because of the female guest zone question on the survey. The question was, "How many days do you use the female guest zone per year? Put zero if you do not have a female guest zone in your house." The model predicted four days for the female guest zone each year, which is close to zero (and this is true in this case). For the living hall, the model predicted that the annual daily use was 360 days, which is likewise true in this case because the house is small and does not have a separate daily family room. Fig. 6.16 shows the two different behaviour predictions for this type of input: "daily use 360 days" in green on the right end of the bar and "special occasion" use, also labelled in green, on the left of the bar. The highest confidence level for the "special occasion" use was 78% in 24 days. In this case, the model will choose the highest confidence level (360 days, 100%

confidence), which is the number at the right end of the bar. The user can adjust the time prediction if needed.

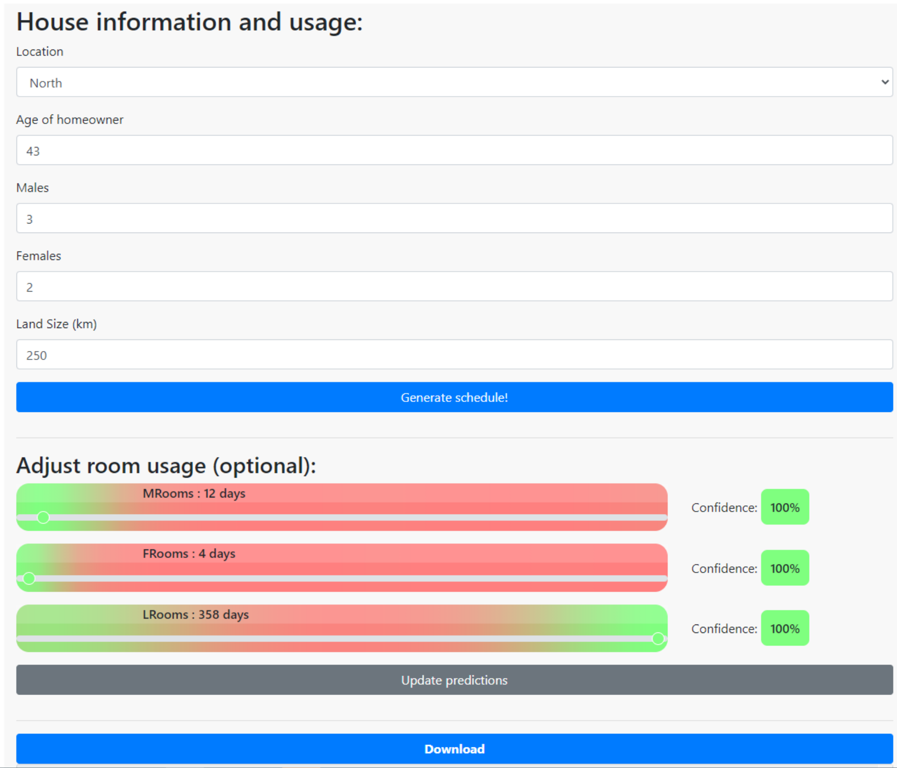


Fig. 6.21. A screenshot of the tool's graphic interface, showing the parameters and prediction section for the first case study.

6.3.1.3. Incorporated the model into an energy simulation experiment

The goal of this section is to validate the new predicted time-use schedules within the case study energy simulation prediction outcomes. Two experiments were conducted to validate the new predicted time-use schedules. In the first experiment, the model was run using all data inputs provided by the householder, except for the weather data and load-time schedules (which were provided by default schedules using the Building Components Library [BCL]) from EnergyPlus. The BCL library had no specific schedules for the study target zones (male guest zone, female guest zone, and living hall); the closest time-use schedule for the targeted zones was the default load schedule for the family gathering room. The operation of the default load schedule was from

3:00 to 10:00, and from 14:00 to 23:00 every day, 360 days a year. Without using the prediction schedules, the electricity-prediction simulation outcome was 140 GJ. According to the electricity bill provided by the householder, the real energy use was 108 GJ. The gap between the electricity prediction and real energy use was 29.6%.

For the second experiment, the same energy model was used as in the first experiment, but the time-use schedules for the male and female guest zones and living hall were replaced with the prediction schedules generated in the previous section. The male guest zone operation prediction was 12 days a year, using fixed hourly time-use from 17:00 to 03:00. The female guest zone operation prediction was zero because the case study did not have such a zone. Three hundred and sixty days were entered for the living hall daily use from 12:00 to 00:00. In the final step, the outcome of the two experiments was compared with the actual electricity bill. The outcome of the electricity prediction in this experiment was 127 GJ, compared to the actual energy use of 108 GJ; the gap was reduced to 17.5%.

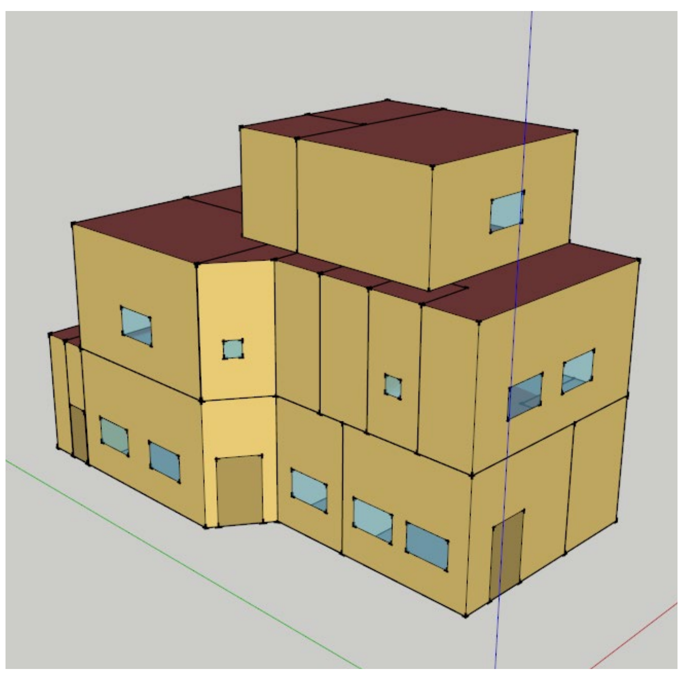


Fig. 6.22. Energy model geometry for the first case study.

6.3.2. *Second case study*

6.3.2.1. *House information*

The second case study was a house located in the east of Riyadh. The size of the house was 754 m^2. Compared to the average house sizes in other areas in Riyadh, this house could be considered a big house. However, in Riyadh, based on study survey data and previous studies discussed in Chapter 2, this house size is considered average, as the mean house size in the area is 750 m^2. The age of the householder was 55, which was also close to the mean age of householders in the study; the average age of householders was 53. There were nine people in the household: five males and four females, which was also close to the mean values represented in the study; the mean for the number of male members living in the house was 5.2, and 3.29 for female members. This house had more rooms compared to case study 1, with one zone for male guests and a separate zone for female guests. However, the total proportion of guest space was 34% of the total built area, which was still within the average range (30-40%) for guest zones in Saudi houses. The householder responses to the five parameter questions were as follows:

Parameters	Location	Age of the householder	Male members	Female members	Land size
Second case study	East	55 years old	5	4	754 m^2

The space types for this case study were as follows:

Y1	Male guest zone: Reception room, dining room, and bathroom.
Y2	Female guest zone: Reception room.
Y3	Living Hall: Ground-floor living hall and first-floor living hall.
Daily use	3 storage rooms, 3 ground WC, kitchen, 1 master bedroom, 4 bedrooms, 4 first floor WC, housekeeper’s room, laundry room, and housekeeper’s WC.

6.3.2.2. Model prediction for the case study

For the second case study, the model predicted an annual use of 37 days for the male guest zone, 23 days for the female guest zone, and 48 days for the living hall. This house had a separate family gathering room; therefore, the living hall was only used occasionally.

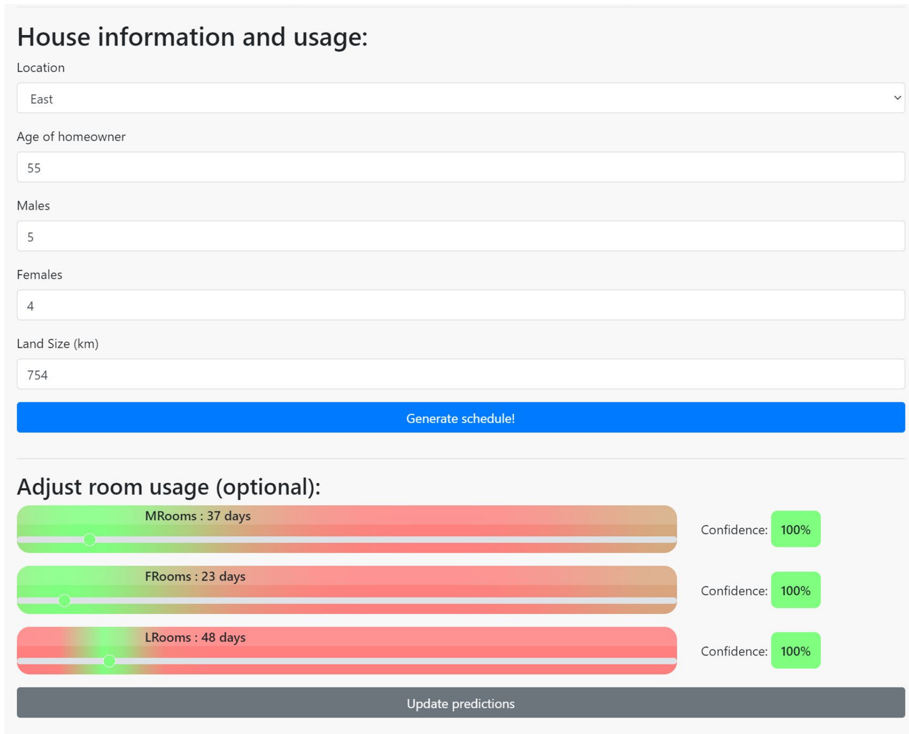


Fig. 6.23. Screenshots of the tool's graphic interface, showing the parameters and prediction section for the second case study.

6.3.2.3. Incorporated the model into an energy simulation experiment

The same method used in the first case study was used to run the first and second experiments in this case study. The first experiment was to define the gap between the energy prediction and the real energy use without using the model. The result showed that the prediction outcome was 238 GJ, and when compared to the real energy-use of 156 GJ, the gap was 52.5%. The second experiment, which used the tool to improve the time-use for the targeted zones, showed the energy-use prediction outcome was 149 GJ compared to the real energy-use, 156 GJ, reducing the gap to 4.5% between energy prediction and the actual energy-use.

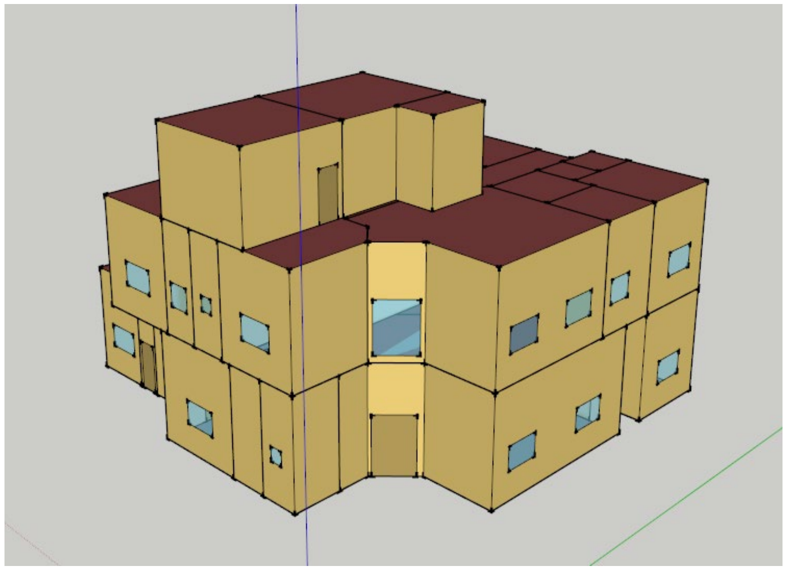


Fig. 6.24. Energy-model geometry for the second case study.

6.3.3. *Third case study*

6.3.3.1. *House information*

The third case study was a house located in the southern area of Riyadh, which was considered the oldest area in the city. Usually, people who live in this area have a lower income than those in the other areas. This house was built in the late 1970s and had more rooms, compared to the open plan of modern houses (for details on changes in house layouts over time in Saudi Arabia, see Chapter 4). Houses built during this time had greater privacy levels than houses built in more recent years. The male guest zone had two male guest reception rooms, a dining room, and a bathroom. Similarly, the female guest zone had two female reception rooms and a bathroom. In addition to the living hall on the ground floor and first floor, the house had a separate daily family gathering room. Study survey data showed that the typical number of people living in homes in the south was higher than in other areas (see Chapter 5). In this case, eight people lived in the house, including the grandfather (92 years old) who built the house. The grandfather's eldest son (56 years old) and his family were living in the house; and the eldest son was now responsible for the house. In this case study, the eldest son was interviewed and asked if there was any unique behaviour. The interview showed that no unique factors impacted occupant behaviour, and this case could be used as a representative case for the southern area of Riyadh. Householder responses to the five parameters were as follows:

Parameters	Location	Age of the householder	Male members	Female members	Land size
Third case study	South	56 years old	4	4	528 m^2

The space types for this case study were as follows:

Y1	Male guest zone: reception room 1, reception room 2, dining room, and bathroom.
Y2	Female guest zone: reception room 1, reception room 2, and bathroom.
Y3	Living Hall: ground-floor living hall and first-floor living hall.
Daily use	1 family room, kitchen, 6 bedroom, 5 bedrooms, housekeeper’s room, laundry room, housekeeper’s WC

6.3.3.2. Model prediction for the case study

For the third case study, the model predicted annual use of 14 days for the male guest zone. For the female guest zone, the model predicted two alternative patterns of behaviour: 36 days a year with a 100% confidence level and daily use with a 60% confidence level (see Fig. 6.26). The model chose 36 days because it had the highest confidence level. For the living hall, the model predicted an annual use of 45 days a year, which was considered only occasional use. The living hall findings were expected because the house had a family room for daily gatherings.

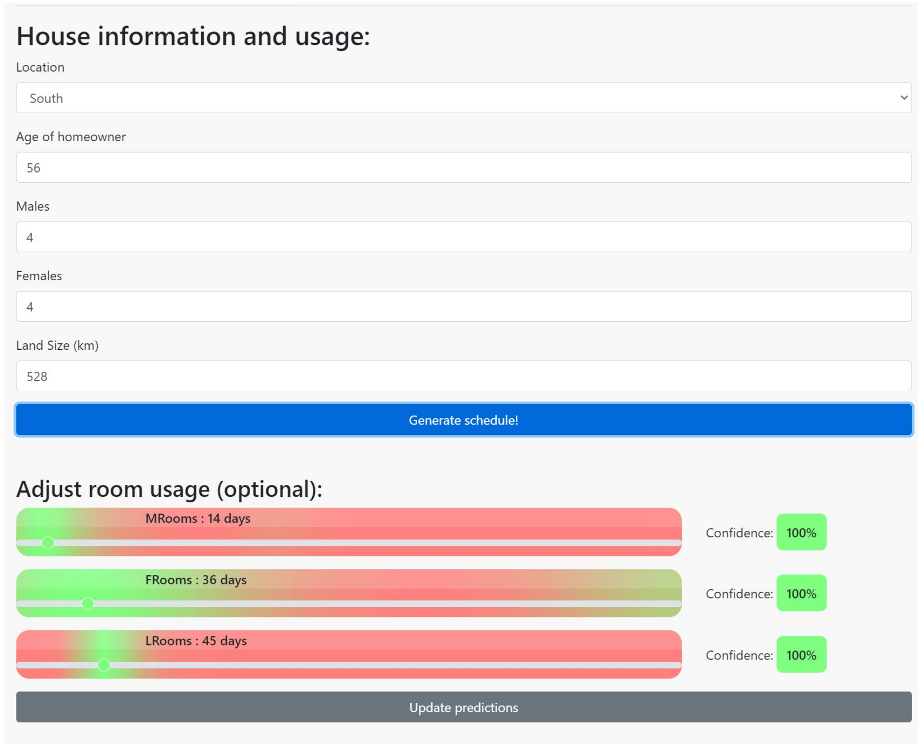


Fig. 6.25. The tool's graphic interface, showing the parameters and prediction section for the third case study.

6.3.3.3. Incorporating the model into an energy simulation experiment

The same method used for the previous case studies was repeated. After creating the energy model and running the simulation, the prediction outcome was 186 GJ, compared to the actual energy use of 131.7 GJ, resulting in a 41.9% gap. Findings suggest old houses use less electricity than newer houses because rooms in old houses are smaller compared to new, open-plan houses. For example, in this case study, the male guest zone had two reception rooms; the size of the first one was 30 m^2, and the second one was 18 m^2. When they had guests, they just needed to use one room and they very rarely used both. However, in new houses, males have one big room for guests and cooling a big room (48 m^2) requires more energy. The outcome using the prediction tool model was 126.7 GJ, as compared to the actual energy use of 131.7 GJ, reducing the gap to 3.8%.

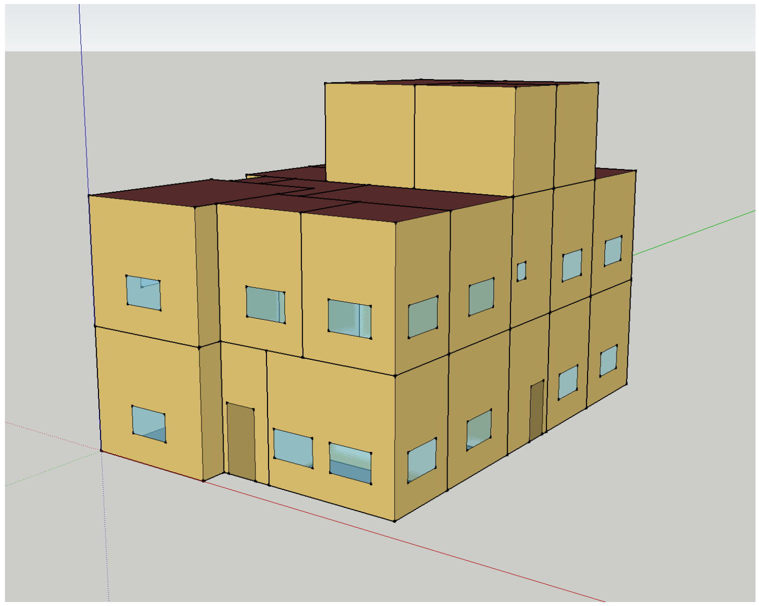


Fig. 6.26. Energy-model geometry for the third case study.

6.3.4. *Fourth case study*

6.3.4.1. *House information*

The fourth case study house was in the western area of Riyadh, built in 2004, and had a modern design. The owner was 32 years old and lived with his mother and two younger brothers. The house had a floor area of 360 m^2, with around 35% of the total built area for guests. The house had one guest zone, which was used for both males and females. The guest zone had a reception room with a bathroom, a dining room on one side, and a living hall on the other side. The front part of the house was for male guests and the back of the house was for the family. In this case, the dining room was located between the two sections, with two doors: one facing the male guest zone and the other facing the family zone. The reason for having two doors is for circulation and visibility purposes, so the family can prepare dinner while the guests are in the reception room without violating their privacy. Householder responses to the five parameter questions were as follows:

Parameters	Location	Age of the householder	Male members	Female members	Land size
Fourth case study	West	32 years old	3	1	360 m^2

The space types for this case study are as follows:

Y1	Male guest zone: reception room, dining room, and bathroom.
Y2	Female guest zone: (NONE)
Y3	Living Hall: ground-floor living hall and first-floor living hall.
Daily use	2 ground WC, kitchen, 1 master bedroom, 4 bedrooms, 2 first floor WC, housekeeper's room, laundry room, and housekeeper's WC

6.3.4.2. Model prediction for the case study

After providing the model with the five parameter inputs, the model predicted an annual use of 8 days for the male guest zone (reception room, dining room, and bathroom). No female guest zone was present in this case study. For the living hall, the model predicted a daily room use of 360 days. As noticed in the analyses in Chapter 5, small houses usually do not have separate family gathering rooms, so they use the living hall for daily gatherings.

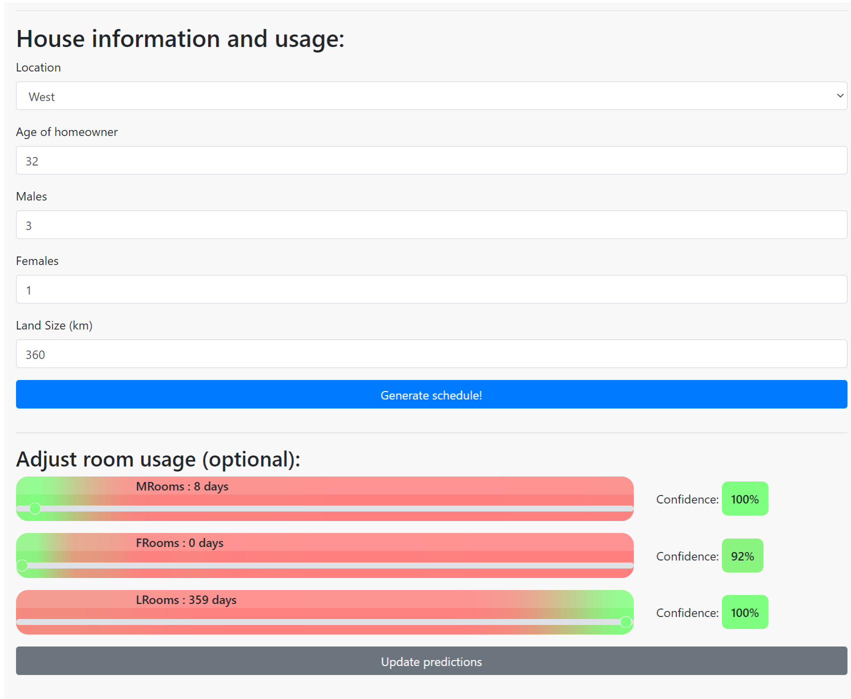


Fig. 6.27. The tool's graphic interface, showing the parameters and prediction section for the fourth case study.

6.3.4.3. Incorporating the model into an energy simulation experiment

The outcome of the first experiment, using the default schedules, predicted the electricity use as 244 GJ, compared to the actual electricity bill of 138 GJ—a gap of 76% between energy prediction and the actual energy consumed. For the second experiment, the model included the survey data for the five parameters, processed by the ML algorithm developed in the previous section. In the next step, the old schedules from experiment 1 were replaced with the new prediction schedules, generated using the model for the time-use of the male and female guest zones and the living hall.

The energy simulation outcome was 218 GJ, after replacing the default schedules with the new predicted schedules. The gap between the prediction and the actual energy use dropped to 57%. Thus, the occupant model reduced the gap by approximately 20%. The gap was still high so the researcher interviewed the householder to understand their behaviour and identify possible explanations for the gap. From the discussion with the house owner, the researcher discovered some unique behaviour related to the operation of the house. The father, who had a wife and four children, built the house. However, he passed away, leaving the eldest son (32 years old) responsible for the family. The house was designed based on the father's requirements, but the eldest son no longer used the house or rooms as his father did before him. For example, some daily rooms were not occupied, one bedroom was not used at all, and the master bedroom was only used when his mother was in the house.

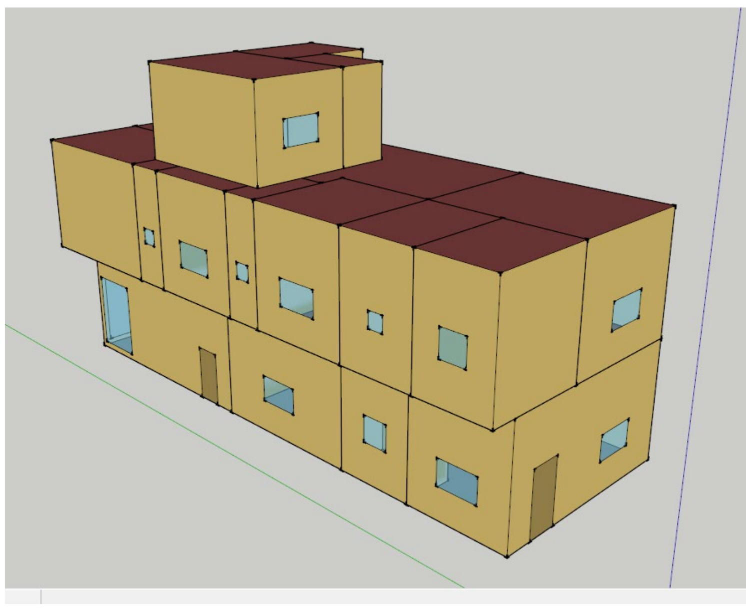


Fig. 6.28. The energy-model geometry for the pilot study.

6.3.5. Fifth case study

6.3.5.1. House information

The fifth case study was a house in the middle of the city. The house was a modern open-plan design built in 2007 with three male and three female occupants. The total land size was 500 m^2, and the householder was 61 years old. The house had a male guest zone and a large living hall, including the female guest zone area. In the open-plan style, houses still separated the male guest zone from other sections in the house for privacy and to prevent circulation overlap. Householder responses to the parameters were as follows:

Parameters	Location	Age of the householder	Male members	Female members	Land size
Fifth case study	Middle	61 years old	3	3	500 m^2

The space types for this case study are as follows:

Y1	Male guest zone: reception room, dining room, and bathroom.
Y2	Female guest zone: (NONE)
Y3	Living Hall: (ground-floor living hall and first-floor living hall).
Daily use	Kitchen, 1 master bedroom, 3 bedrooms, family room, 7 WC, housekeeper's room, laundry room, and housekeeper's WC

6.3.5.2. Model prediction for the case study

The five parameters provided by the householder were used to run the model and predict time-use for the male guest room, the female guest room, and the living hall. The model predicted an annual use of 20 days for the male guest zone per year. The model also predicted a usage pattern for the female zone; however, this case study did not have a female guest zone. For the living hall, the model predicted an annual use of 52 days.

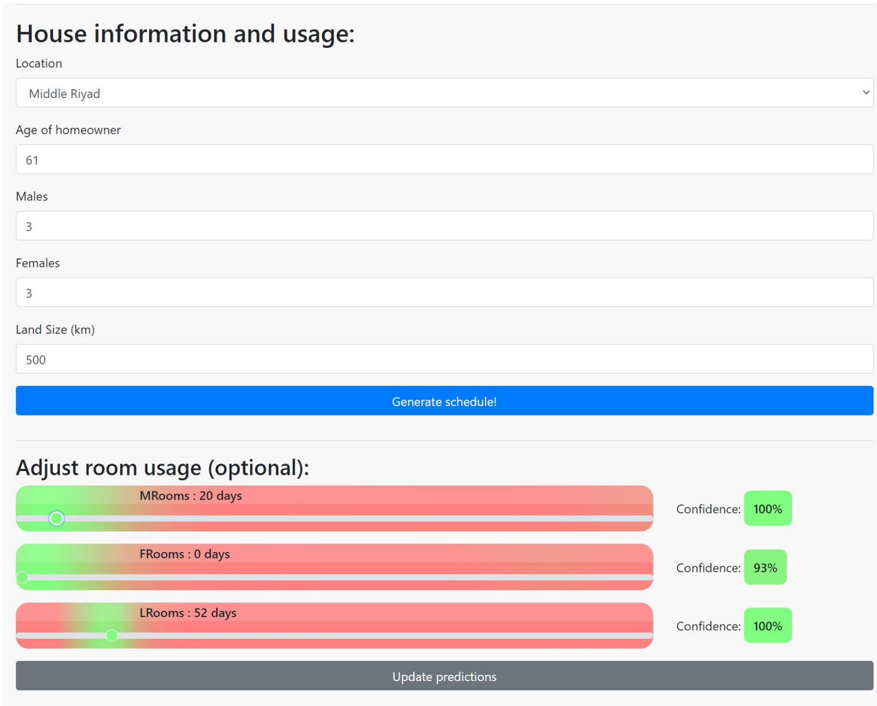


Fig. 6.29. The tool's graphic interface, showing the parameters and prediction section for the fifth case study.

6.3.5.3. Incorporating the model into an energy simulation experiment

This section evaluates the prediction model effect on the total energy use of the case study. The outcome of the first experiment was 348 GJ. When compared to the real electricity use of 263 GJ, the gap was 32.6%. In the second experiment the outcome was 278.7 GJ, using the new predicted time-use schedules,. When compared to real energy use, the gap dropped to 6%. This case study had a basement, which increased the energy-use outcome. Moreover, the open-plan design required a lot of energy to cool the open-plan spaces, particularly because of the hot desert climate that exists in Riyadh.

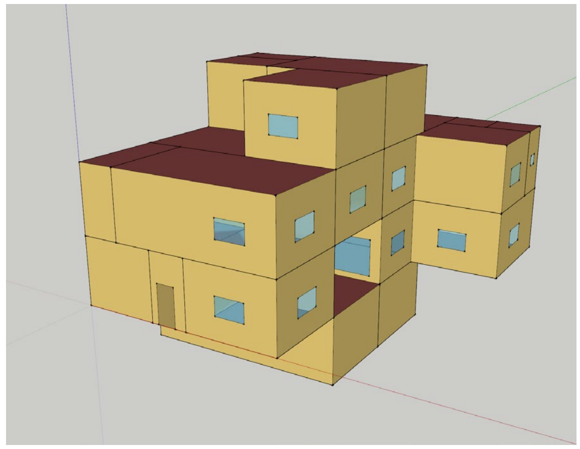


Fig. 6.30. The energy-model geometry for the pilot study.

6.4. Discussion

The five case studies represented the five main locations in Riyadh city as shown in Fig. 6.32. The Random Forest (RF) models were applied in all cases using five specific inputs: 1) location, 2) householder age, 3) number of male members, 4) number of female members, and 5) land size for each case to predict time-use for the male and female zones, and the living hall. All case study parameters were in the first standard deviation range, which covered 68% of the data (the range was between the two-dotted blue lines in Fig. 6.33).

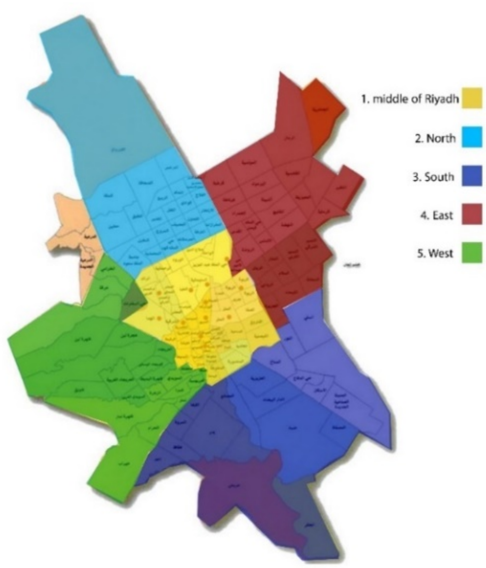


Fig. 6.31. Map of Riyadh city highlighting the five main locations.

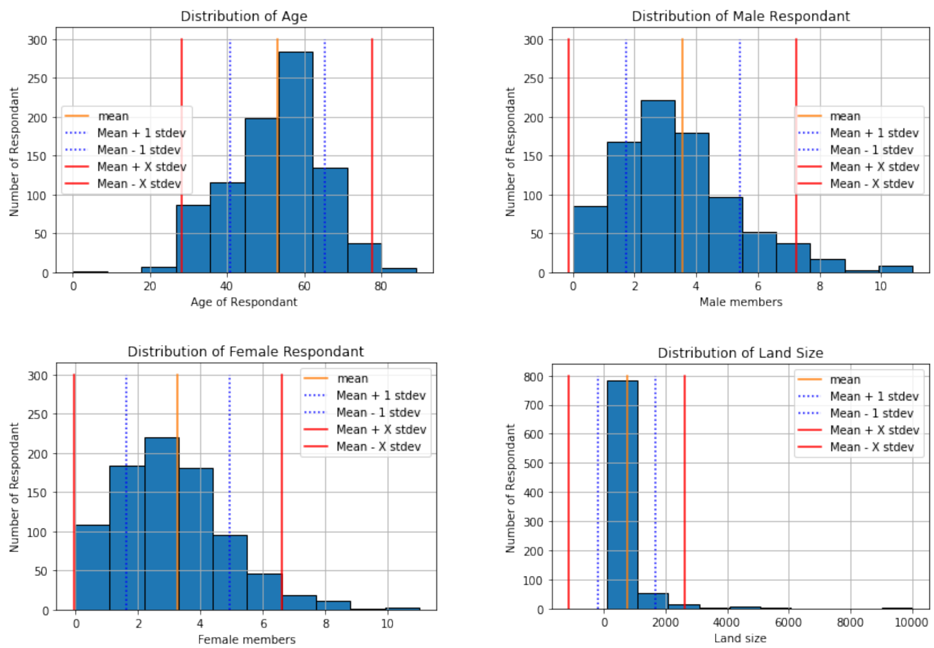


Fig. 6.32. Data distribution for the four parameters showing the means and the first and second standard deviations.

The outcomes of the machine learning model were integrated with the energy simulation to evaluate the influence of the new daily schedules on the predictions. The outcome showed that the gap between energy simulation prediction and actual energy consumption could be reduced to between 17.5% and 3.8% (Table 6.9).

Table 6.9
Case-study outcomes.

Case Study	The five parameters input: the first column is the location (X1), the second is the age of the householder (X2), the third is the number of male members (X3), the fourth is the number of female members (X4), and the fifth is the land size (X5) (m^2 unit)					The time-use prediction for the male guest zone (Y1), female guest zone (Y2), and living hall (Y3)(days unit).			Experiment 1, using default time-use. The first column is the energy-use prediction (Gigajoules unit), and the second is the % difference from the actual energy use.		Experiment 2, using the time-use schedules ML. The first column is the energy-use prediction (Gigajoules unit), and the second is the % difference from the actual energy use.		Actual energy use. The first column is the energy-use prediction (Gigajoules unit), and the second is the % difference from the actual energy use.	
Unit	X1	X2	X3	X4	X5	Y1	Y2	Y3	Gj	%	Gj	%	Gj	%
1	North	43	3	2	250	12	0	360	140	+29.6%	127 GJ	+17.5%	108	0
2	East	55	5	4	754	37	23	48	238	+52.5%	149GJ	-4.5%	156	0
3	South	56	4	4	528	14	36	45	186	+41.9%	126.7	-3.8%	131	0
4	West	32	3	1	360	8	0	360	244	+76%	218	+57%	138	0
5	Centre	61	3	3	500	20	0	52	348	+32.6%	278.7	+6%	263	0

6.4.1. *Discussion of the findings*

A lack of information about Saudi culture and occupant behaviour caused low accuracy inputs for energy estimation, leading to inaccurate energy-use predictions. This study focused on bridging this gap by predicting the usage of three zones: the male guest zone, the female guest zone, and the living hall. Because of the high level of privacy in Saudi culture, these three zones must be separated to avoid circulation overlap. Thus, all Saudi houses must have a male guest zone, even if the owner knows that they will only use it for a few days a year. Householders keep guest zones closed and do not use them unless there are guests in the house (Fig. 6.33). Having designated guest zones is a sign of prestige and generosity, important values in Saudi culture.

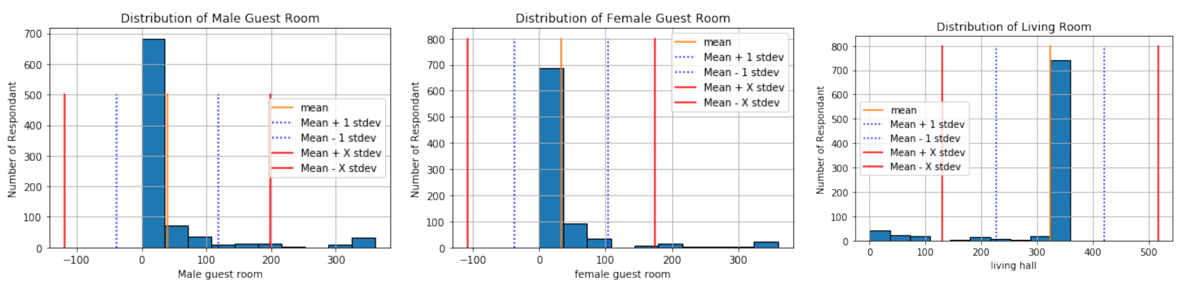


Fig. 6.33. The time-use data distribution of the targeted zones.

Saudi house design history has evolved, as discussed in Chapter 4. This study reviewed the three different periods representing changes in house designs over time. Traditional houses were built using mud, with courtyards in the centre. Modern concrete houses used the setback technique with front yards. Open-plan houses also followed the setback technique and had a front yard. In this study, the third case study in the south was a house built in the 1970s and followed the second-period design. Walls were used to segregate rooms to maintain a high level of privacy. The male guest zone, for example, had two separate guest rooms: a reception for visitors and a separate dining room. The same applied to the female zone. As a result, the energy consumption was low, not because of the techniques used to reduce the heat but because they were not using all of the rooms in the house. These findings suggest house occupants likely spent most of the day in a small room and did not take advantage of the other available rooms in the house. On the other hand, case study five had an open-plan house, which was considered a new modern building style built in 2007. This house showed high-energy consumption because of the energy needed to cool the big open space of the living hall (Table 6.10), which also opened into the female zone as well. The designer attempted to remedy the increased energy consumption by using location and self-shading techniques. But the energy use was higher compared to closed-plan houses.

Table 6.10
Comparison of the land size with real electricity use.

Unit	Land size (m^2)	Electricity-use annually (Gj)
Case study #1	250	108
Case study #2	754	156
Case study #3	528	131
Case study #4	360	138
Case study #5	500	263

In the fourth case study, the person who built the house passed away and the occupant behaviour of the remaining family members did not reflect the original build-operation design. This showed that the energy-use consumption changes based on the behaviour and choices of current occupants. If the house were designed with different considerations than its operation predicted during the design stage, then the gap between the energy consumption prediction and real energy use would be large. Findings from the RF model implementation and case study experiments offer promising results to help reduce the gap in prediction and actual energy use.

7. Conclusion

This chapter presents the research findings of the dissertation. In particular, this study provides important information about Saudi cultural behaviour related to energy use in houses, as well as a methodology that can be applied to other cultures. A review of study limitations and opportunities for future research and practice is also presented. The goal of this research is to better understand culture-specific behaviour on energy use, leading to improved energy predictions for domestic buildings. The author hypothesized that cultural considerations related to guest zone use in Saudi Arabia have a significant impact on energy use. Through a mixed-method approach, the study explored critical factors from Saudi culture on occupancy behaviour and its energy-using implications. The researcher collected TUD survey data to fill the gap in energy use linked with unique cultural behaviour. The data collected was further used to develop a prediction model for occupancy in Saudi houses.

The researcher chose the Random Forest out of 18 prevailing regression models because it has one of the lowest MAE, and because of its strength in learning and predicting from social behaviour parameters. The model evaluation was carried out using five case studies. The occupancy prediction model was implemented and deployed as a stand-alone tool to support energy simulation using EnergyPlus. This occupancy prediction tool generates the time-use pattern for guest zones (male guest zone, female guest zone, and living hall) in Saudi houses. As a result, energy simulations using culture-driven energy use patterns were able to reduce the gap between predicted and real energy use.

The Conclusion Chapter (Chapter 7) covers five main sections:

- Research findings and discussion.
- Summary of the theoretical contribution of the study.

- Limitations of this research.
- Avenues for future research.

7.1. Research findings and discussions

7.1.1. Layout

Research findings in the pilot study (Chapter 4) demonstrate that a courtyard layout in Saudi culture is better for circulation, privacy, and visibility than the setback technique used in modern houses. The visibility analysis, using space syntax, shows that traditional houses have the highest visibility compared to the four other modern houses. In our findings, the first standard deviation in traditional houses is much higher than the others. This suggests that traditional houses have more open space and guest rooms have low visibility leading to a large discrepancy in space visibility. This finding supports the privacy requirement in Saudi Arabia as traditional houses usually have more open spaces for the family zone, which are visually protected from visitors and neighbours. The experiment also shows clear visual and circulation segregation between the three main zones. It was noted that each zone has one or two reception rooms, and some houses had up to four reception rooms in total. This shows that homes had varying layouts and explored unique behaviour in the use of these zones. Several studies explain the primary reasons for this separation between zones related to privacy and gender, which is reflected in the Saudi culture (AlKhateeb, Humphries-Smith & Eves 2014).

7.1.2. Importance of guest zone (size and use in Saudi houses)

Every culture has its own values and beliefs. For instance, generosity and privacy are very important aspects of Saudi culture. Saudi people often want to invite guests to their homes without violating their family's privacy. This is reflected in their homes, as every house must possess at least one guest zone. The guest zone features either one or two reception rooms, a dining room, and a bathroom for guests. Some households possess two separate guest zones, one for male guests and another for female guests. The average per cent of guest zone areas equals 30 to 40% of the house's total built area, demonstrating the generous space guests are offered when they visit families. Another critical factor related to privacy is limiting the overlap between family members and visitors during circulation. To achieve this, male and female guest zones have separate entrances and are visually separated from the family zone. The study revealed that the average time-use for the male guest zone in Riyadh houses equals 40 days per year, and the average use for the female guest zone equals 33 days each year.

7.1.3. Impact of considering cultural behaviour on energy performance evaluations

Study findings demonstrate that by adding the cultural behaviour considerations (time-use patterns) to the energy model, the gap between the actual and the predicted energy use can be reduced by an average of 11% (with a range of 17.5%–4.5%). In addition, the data show there is different behaviour in guest zone usage related to family size, and to single-family houses compared to multiple-family houses. The data show increasing trends in the time-use of the male guest zone compared to the number of male members living in the house. However, the trend declines after six male members live in the house. The trends are similar for females, with an increase in time-use related to the number of female family members and the female guest zone. However, the declining trend changes after eight female members live in the house.

7.1.4. Using occupancy-driven schedules can improve energy-use predictions

Understanding Saudi house guest zones and associated time-use behaviour has not been well considered until now in energy performance evaluations. In typical energy performance simulations, the assumption considers that all rooms in homes are used daily. However, this is not the case for Saudi houses. As noted earlier, 30–40% of the total built area (guest zones) in Saudi homes are only used occasionally. Therefore, energy usage in these homes differs. Using the default load schedules provided by the energy simulation software, such as EnergyPlus, in five case studies without consideration of unique time-use for the targeted zones, leads to a significant gap between energy predictions and the real energy use. These gaps range between 29% and 52%. The gap was reduced to 3.8% by evaluating the occupancy prediction tool, replacing the default load schedules with the new predicted schedules.

7.1.5. Applicable design strategies and suggestions

The study also found that the best solution to creating an energy-efficient house, while also respecting cultural values, is to keep the male guest zone near the main entrance. This design layout makes access to the male guest zone direct from the street, so guests and family members do not overlap with their circulation (which considers the privacy of family members). Also, the house should have at least two entrances: one entrance in the front of the house that leads to the male guest zone and a second entrance, which opens directly into the family area and is usually located at the back of the house to maintain privacy. The courtyard technique is the best layout to achieve the goal of privacy and energy efficiency, with guest zone windows facing public areas and family zone windows facing private areas. While guest zones are only used occasionally, Saudi culture dictates that every Saudi house must have a guest zone.

7.1.6. Hourly operational schedules for the houses of Riyadh

The hourly time-use data (weekdays and weekends) show that the people use the guest zones often at night. In Saudi Arabia, it is common to have gatherings at night due to the hot, dry climate during the day. People prefer to meet at night when the weather is cooler. People typically meet after the last prayer of the day, which is Isha prayer time in Riyadh, at around 8 pm. To accommodate visitors in their home, homeowners start to prepare the guest zone by turning on the

HVAC to cool the zone, one to two hours before guests arrive. It is also more common to invite guests over at the weekend (and rare to use guest zones on weekdays). If a house is small and there is just one guest zone for both genders, homeowners invite a male guest on one day and female guests on a separate day. The data show there is slightly different hourly time-use behaviour between the male guest zone and the female guest zone. For the male guest zone, the data show that guests leave at around midnight, whereas the female guests in the female guest zone stay two hours longer on average. These data demonstrate how the Islamic culture influences operational schedules in houses in Riyadh, Saudi Arabia. It also demonstrates the uniqueness of occupant behaviour and how it varies across and within houses.

7.2. Summary of the theoretical contributions of the research

Occupant behaviour is one of the main factors that impact energy use in buildings (Yoshino, Hong & Nord 2017, p. 53). Occupant behaviour is driven by several elements and culture is one of them. There is a dearth of research that examines the unique nature and lifestyle of Saudi Arabian culture and its impact on energy use. Aljammaz et al. (2019) stated that the lack of information regarding Saudi occupant behaviour could lead to an average gap of 40% between energy prediction and real energy use. Reliable energy use predictions can help inform the effectiveness of house designs, and considering behaviour is essential to improve energy predictions. The accuracy of energy predictions is dependent on representative load schedules that can be derived from existing TUD survey data, as demonstrated in countries such as the UK, Sweden and France. However, the influence of the unique Saudi culture on energy-use behaviour is not well understood. Thus, this research aims to close this knowledge gap and strengthen the need for acquiring representative energy use data linked to occupant behaviour as a key consideration for energy predictions. As demonstrated in this study, the developed occupancy model that incorporates culture-specific time-use schedules of Saudi houses could reduce the gap between energy predictions and real energy use to as low as 3.8%.

7.3. Limitations

The current research focused on single-family houses in Riyadh, Saudi Arabia. The results from the study cannot be generalised to other countries and other household types. The research highlights the importance of considering cultural occupant behaviour that affects energy use to enable accurate energy predictions.

With the focus on single-family houses in Riyadh, Saudi Arabia, the present study included a convenience sample of participants from King Saudi University. Faculty members were recruited during a conference on sustainability solutions in neighbourhoods in the Architecture School in 2018. As such, the primary survey sample is made up predominately of highly educated participants and may not represent the wider community. In future research, the author will investigate if the educational background of householders could have a significant impact on the observed behaviour, particularly in relation to the time-use of specific zones identified in this study.

Furthermore, this study only focused on cultural behaviour related to occupants' use of guest zones in their homes. The research did not consider other behaviour that may be driven by Saudi culture and also influence energy use. For instance, the study did not examine the use of windows, courtyards or backyards, which may also impact energy consumption. Instead, the study investigated the impact of Saudi cultural behaviour on the time-use of targeted zones. Cultural behaviour has an impact on the use of other features, such as the time-use, direction, and size of windows (Aljammaz 2016). In addition, the geometry and the orientation of the house (courtyard or using the backyard) also contribute to occupant behaviour and energy use (AlKhateeb, Humphries-Smith & Eves 2014).

To validate the methodology, five case studies were chosen, each representing one of the five main areas in Riyadh city (north, south, east, west and centre). The case studies' parameters are within the range of the first standard of each parameter, which represents 68% of the data density. For the householder's age, the first standard deviation ranges between 41 and 65. The first standard deviation for the males living in the house ranges between 1.76 and 5.36 people. For the females who live in the house, the range is between 1.6 and 5 people. Regarding the date the house was built, all the case studies were in the first standard deviation of the data, between 1992 and 2014. The exception would be the house in the southern area, which was built before that (as the South is an old area, and most of the houses in it are old as well). The final parameter concerns land size. All case studies also reside in the range of the first standard deviation, between 120m^2 and 2,628m^2. The data above and below the first standard deviation were not validated due to the low number of the dataset on the other ranges.

7.4. Future research

Time-use data are a key metric in supporting energy predictions. However, current fixed TUD do not account for underrepresented user groups (such as Saudi houses), subsequently leading to inaccurate energy predictions. This research collected time-use data for guest zones in Saudi houses and developed an applicable occupancy prediction tool for end-use energy estimations. In the future, more time-use data is envisaged to continuously improve occupancy predictions. For instance, hourly fixed schedules can be improved with flexible hourly inputs linked with dynamic occupant behaviour. Also, this study only focused on the energy use pattern of guest zones (inclusive of reception room, living room, bathroom, and so forth). Future examination of the time-use of each room is needed to further improve energy performance evaluations.

This research covers cultural behaviour and associated energy use of houses in Riyadh, Saudi Arabia. By including more households with different house types (single- and multiple-family homes), the researcher will continue investigating culture considerations and their significant impact on energy use in different regions of Saudi Arabia. Future research areas include investigating the extent to which this phenomenon is specific to Riyadh and whether culture is as influential on energy use predictions in other parts of Saudi Arabia.

Furthermore, the researcher will also widen data collection on houses out of the first standard deviation range for the parameters. For example, data analysis (reported in Chapter 5) discovered that the use of the guest zone increased from zero to six male members, but the trend decreased with more than six male occupants. A similar pattern is found for females, but with eight occupants and more. The researcher will validate the new data for the big number of family and big houses using representative case studies for the new data, then using the same methodology to evaluate the impact of the new data on closing the gap between the energy prediction and the real energy use for each area.

In addition, this research shows that there are multiple rooms in guest zones in Saudi houses only used a few times a year. This finding does not diminish the necessity of these rooms in houses. Every Saudi house must have at least one zone for guests, as these zones are part of cultural considerations related to privacy and gender. Reducing the number of guest rooms with adaptive use patterns could be an alternative design strategy to improve energy efficiency while abiding by the high levels of privacy demanded by Saudi culture.

The occupancy prediction tool can be developed to be an easy-to-use interface tool, integrated with the energy simulation process to predict more accurate time-use schedules for the targeted zones in the Saudi houses. The tool can be used by any user who has average knowledge of how to use energy simulation software, and can answer the five main parameter question (location, age of the householder, number of males in the house, number of females in the house, and the size of the house).

7.5. Summary

In summary, the current research aimed to assess the impact of Saudi culture on energy use in houses in Riyadh, Saudi Arabia. Through a series of surveys and experimental investigations, the researcher found that Saudi culture was reflected in house designs, circulation of family members and guests, and gender-specific guest zones. Families typically separate male guest, female guest, and family living zones. To understand this behaviour, time-use data were gathered and used as inputs into an occupancy prediction model. The outcome showed that using the improved time-use schedules in energy modelling could reduce the gap between predicted and actual energy use by 3.8%. These results reinforce the significance of culture and its effects on energy use occupant behaviour as well as its applications in building performance evaluations.

Findings from this study can benefit many professionals. For example, architects could use these findings to examine representative end-use energy consumption to inform energy performance evaluations for prospective buildings in Saudi Arabia. Furthermore, understanding Saudi occupant behaviour will help inform architects on the best house layouts to improve energy efficiency, while also respecting the privacy and culture of residents. Incorporating sustainable techniques and layouts, such as courtyards to keep privacy high and allow sunlight to enter homes, can be one of many applicable design recommendations. This study helps illuminate the impact of Saudi culture and home designs on energy use evaluations.

References

Ahmad, M. W., Mourshed, M. and Rezgui, Y. (2017). Trees vs neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption. *Energy and Buildings* [online]. **147**, 77–89. Available from: doi.org/10.1016/j.enbuild.2017.04.038

Akbar, J. A. (1980). *Support for court-yard houses: Riyad, Saudi Arabia*. Thesis, Massachusetts Institute of Technology.

Almeida, N. *et al.* (2010) 'A framework for combining risk-management and performance-based building approaches', *Building Research & Information*, 38(2), pp. 157–174. doi: 10.1080/09613210903516719.

Al-Said, F. (2003) 'The pattern of structural transformation of the Saudi Contemporary Neighbourhood: The case of Al-Malaz, Riyadh, Saudi Arabia', in *39th International Society of City and Regional Planners Congress*. Available at: http://www.isocarp.net/data/case_studies/347.pdf.

Al Surf, M., Susilawati, C. and Trigunarsyah, B. (2012). Analyzing the literature for the link between the conservative Islamic culture of Saudi Arabia and the design of sustainable housing. In: A. Ali, R. Ibrahim, J. Goulding, and F. Pour Rahimian. eds., *Proceedings of the 2nd International Conference on Socio-Political and Technological Dimensions of Climate Change,* pp. 3–16. Kuala Lampur, Malaysia: University Putra Malaysia Press. Available from: http://www.frsb.upm.edu.my/stdcc/index.php/STDCC/STDCC

Al-Hathloul, S. (2017). Riyadh development plans in the past fifty years (1967-2016). *Current Urban Studies* [online]. **5**(1), 97–120. Available from: doi.org/10.4236/cus.2017.51007

Al-Hemaidi, W. (1996) *The dilemma of regulating privacy: Planning regulations, privacy and house form: The case study of low-density single-family dwellings in Saudi Arabia.* Ph.D. Thesis, University of London.

Alahmadi, M., Atkinson, P. and Martin, D. (2013). Estimating the spatial distribution of the population of Riyadh, Saudi Arabia using remotely sensed built land cover and height data. *Computers, Environment and Urban System*s [online]. **41**, 167–176. Available from: doi.org/10.1016/j.compenvurbsys.2013.06.002

AlGhamdi, A. (2019). Data insight: Electricity tariff changes in Saudi Arabia. Riyadh: KAPSARC.

Ali, Q. et al. (2020). The performance gap in energy-efficient office buildings: How the occupants can help?" *Energies* [online]. **13**(6), 1480. Available from: doi.org/10.3390/en13061480.

Aljammaz, M. (2016). *Daylighting in Riyadh and Los Angeles: Comparison of Cultural Factors in potential market penetration*. M.B.S. Thesis. University of Southern California.

Aljammaz, M., Wang. T. H. and Peng, C. (2020). Understanding occupant behaviour in Islamic homes to close the gap in building performance simulation: A case study of houses in Riyadh. *Proceedings of the 16th IBPSA International Conference and Exhibition.* Rome, Italy, Sept. 2-4.

Aljammaz, M., Wang, T. H. and Peng, C. (2019). The influence of Saudi Arabian culture on energy use: Improving the time-use schedules in energy simulation for houses in Riyadh. In Ji-Hyun Lee, ed. *Hello, Culture! Proceedings of the 18th International Conference, CAAD Futures 2019.* Daejeon, Korea, pp. 273-289. CUMINCAD.

AlKhateeb, M., Humphries-Smith, T. and Eves, B. (2014). Space design and privacy in a Saudi House." In: Conference: Time, Space, and the Body 2014, 7–9 Sept. Oxford

Alkhedheiri, A. A. (1998) *The role of secondary cities in the national development process of Saudi Arabia.* Ph.D. University College London (University of London). Available at: http://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.287873 (Accessed: 17 June 2017).

Almahmood, M. (2017). Mapping the gendered city: Investigating the socio-cultural influence on the practice of walking and the meaning of walkscapes among young Saudi adults in Riyadh. *Journal of Urban Design* [online]. **22**(2), 229–48. Available from: doi.org/10.1080/13574809.2016.1273742

Alshahrani, J, and Boait, P. (2019). Reducing high energy demand associated with air-conditioning needs in Saudi Arabia. *Energies* [online]. **12** (1), 87. https://doi.org/10.3390/en12010087.

Alwetaishi, M. (2018). Toward sustainable building design in hot and arid climate with reference to Riyadh City, Saudi Arabia. *International Journal of Civil and Environmental Engineering* **12**(5), 572–75.

Amaratunga, D. and Baldry, D. (2001). Case study methodology as a means of theory building: Performance measurement in facilities management organisations. *Work Study* [online]. **50**(3), 95–105. Available from: doi.org/10.1108/00438020110389227

Amlôt, M. (2021). Saudi Crown Prince reveals plan for Riyadh to be top 10 largest city economy in the world. *Alarabiya News*, 28 January; Updated 30 January.

Amran, Y. H. A. et al. (2020). Renewable and sustainable energy production in Saudi Arabia: The direction of Saudi vision 2030. *Journal of the Knowledge Economy* [online]. **8**(2), 536–64. Available from: doi.org/10.1007/s13132-017-0479-8

Ao, Y. et al. (2019). The linear Random Forest algorithm and its advantages in machine learning assisted logging regression modeling. *Journal of Petroleum Science and Engineering* [online]. **174**, 776–89. Available from: doi.org/10.1016/j.petrol.2018.11.067

Asfour, O. S. (2020). A comparison between the daylighting and energy performance of courtyard and atrium buildings considering the hot climate of Saudi Arabia. *Journal of Building Engineering* [online]. **30**, 101299. Available from: doi.org/10.1016/j.jobe.2020.101299

Babangida, H. and Katsina, H. S. (2018). Integrating Islamic design principles for achieving family privacy in residential architecture [online]. *Journal of Islamic Architecture* [online]. **5**(1) 9–19. Available from: doi.org/10.18860/jia.v5i1.4407

Barnston, A. G. (1992). Correspondence among the correlation, RMSE, and Heidke forecast verification measures; Refinement of the Heidke score. *Weather and Forecasting* [Online]. **7**(4), 699–709. Available from: doi.org/10.1175/1520-0434(1992)007<0699:CATCRA>2.0.CO;2

Barthelmes, V. M. et al. (2018). Profiling occupant behaviour in Danish dwellings using time use survey data. *Energy and Buildings* [online]. **177**, 329–40. Available from: doi.org/10.1016/j.enbuild.2018.07.044

Breadsell, J. K., Byrne, J. J. and Morrison, G. M. (2019). Pre- and post-occupancy evaluation for resident motivations for and experiences of establishing a home in a low-carbon development. *Sustainability* [online]. **11**(14), 3970. Available from: doi.org/10.3390/su11143970

Briller, D. L. (2013). Adapting to a new reality—Strategies for building energy design in a changing climate. *Strategic Planning for Energy and the Environment* [online]. **33**(1), 7–65. Available from: doi.org/10.1080/10485236.2013.10750245

Capasso, A., Lamedica, R. and Prudenzi, A. (1994). A Bottom-up approach to residential load modeling. *IEEE Transactions on Power Systems* [online]. **9**(2), 957–64. Available from: doi.org/10.1109/59.317650

Carlucci, S. et al. (2020). Modeling occupant behavior in buildings. *Building and Environment* [online]. **174**, 106768. Available from: doi.org/10.1016/j.buildenv.2020.106768

Choi, J. H. (2017). Investigation of the correlation of building energy use intensity estimated by six building performance simulation tools. *Energy and Buildings* [online]. **147**, 14–26. Available from: doi.org/10.1016/j.enbuild.2017.04.078

Ciulla, G. and D'Amico, A. (2019). Building energy performance forecasting: A multiple linear regression approach. *Applied Energy* [online]. **253,** 113500. Available from: doi.org/10.1016/j.apenergy.2019.113500

Clarke, J. A. (2001). *Energy Simulation in Building Design*. 2nd ed. Oxford, UK: Butterworth-Heinemann.

Clevenger, C. M. and Haymaker, J. (2006). The impact of the building occupant on energy modeling simulations. In: *Joint International Conference on Computing and Decision Making in Civil and Building Engineering*, Montreal, Canada, pp. 1–10.

Corrado, V. and Fabrizio, E. (2019). Steady-state and dynamic codes, critical review, advantages and disadvantages, accuracy, and reliability. In: F. Asdrubali and U. Desideri , eds. *Handbook of Energy Efficiency in Buildings*. Butterworth-Heinemann. pp. 263–94.

Crawley, D. et al. (2008). Contrasting the capabilities of building energy performance simulation programs. *Building and Environment* [online]. **43**(4), 661–73. Available from: doi.org/10.1016/j.buildenv.2006.10.027

Cruz, P. J. S. (2019). Structures and architecture: Bridging the gap and crossing borders. In: *Proceedings of the Fourth International Conference on Structures and Architecture (ICSA 2019)*, July 24-26, Lisbon, Portugal.

D'Oca, S., Hong, T. and Langevin, J. (2018). The human dimensions of energy use in buildings: A review. *Renewable and Sustainable Energy Reviews* [online]. **81**(1), 731–42. Available from: doi.org/10.1016/j.rser.2017.08.019

Demirbas, A., Hashem, A. A. and Bakhsh, A. A. (2017). The cost analysis of electric power generation in Saudi Arabia. Saudi *Energy Sources. Part B, Economics, Planning and Policy* [online]. **12**(6), 591–96. Available from: https://doi.org/10.1080/15567249.2016.1248874

Denney, A. S. and Tewksbury, R. (2013). How to write a literature review. *Journal of Criminal Justice Education* [online]. **24**(2), 218–234. Available from: doi.org/10.1080/10511253.2012.730617

de Wilde, P. (2014). The gap between predicted and measured energy performance of buildings: A framework for investigation. *Automation in Construction* [online]. **41**,40–49. Available at: doi.org/10.1016/j.autcon.2014.02.009

Diao, L. *et al.* (2017) 'Modeling energy consumption in residential buildings: A bottom-up analysis based on occupant behavior pattern clustering and stochastic simulation', *Energy and Buildings*, 147, pp. 47–66. doi: 10.1016/j.enbuild.2017.04.072.

Ding, Y. et al. (2020). A prediction model coupling occupant lighting and shading behaviors in private offices. *Energy and Buildings* [online]. **216**, 109939. Available from: doi.org/10.1016/j.enbuild.2020.109939.

Ding, Y. et al. (2015). Uncertainty sources and calculation approaches for building energy simulation models. (2015). *Energy Procedia* [online]. **78**, 2566–2571. Available from: doi.10.1016/j.egypro.2015.11.283

DOE2.Com Home Page. (2021). Available from: doe2.com/

Dong, B. et al. (2018). Modeling occupancy and behavior for better building design and operation—A critical review. *Building Simulation* [online]. **11**(5), 899–921. Available from: doi.org/10.1007/s12273-018-0452-x

Dronkelaar, C. et al. (2016). A review of the energy performance gap and its underlying causes in non-domestic buildings. *Frontiers in Mechanical Engineering* [online]. **1**. Available from: doi.org/10.3389/fmech.2015.00017.

Du, J., Pan, W. and Yu, C. (2020). In-situ monitoring of occupant behavior in residential buildings—A timely review. *Energy and Buildings* [online]. **212**, 109811. Available from: doi.org/10.1016/j.enbuild.2020.109811

Du, J., Yu, C. and Pan, W. (2020). Multiple influencing factors analysis of household energy consumption in high-rise residential buildings: Evidence from Hong Kong. *Building Simulation* [online]. **13**(4), 753–69. Available from: doi.org/10.1007/s12273-020-0630-5

Dursun, P. (2007). Space Syntax in Architectural Design. Available from: .https://www.researchgate.net/publication/228646158_Space_Syntax_in_Architectural_Design

Dwyer, T. (2013). Knowledge is power: Benchmarking and prediction of building energy consumption. *Building Services Engineering Research and Technology* [online]. **34**(1), 5–7. Available from: doi.org/10.1177/0143624412471130

Edwards, B. et al., eds. (2004). *Courtyard housing: Past, present and future*. London, UK: Taylor and Francis.

Energy Independence and Security Act of 2007 – Policies [WWW Document]. (2017) . IEA. URL https://www.iea.org/policies/910-energy-independence-and-security-act-of-2007 (accessed 6.20.21).

EnergyPlus (2021) *EnergyPlus*. Retrieved October 3, 2021, from https://energyplus.net/

European Commission. (2016). Commission launches new database on the energy performance of buildings [WWW Document]. Energy - European Commission. URL https://ec.europa.eu/energy/en/news/commission-launches-new-database-energy-performance-buildings (accessed 2.16.19).

Fabi, V., Andersen, R. K. and Corgnati, S. (2015). Verification of stochastic behavioural models of occupants' interactions with windows in residential buildings. *Building and Environment* [online]. **94**(1), 371–83. Available at: doi.org/10.1016/j.buildenv.2015.08.016

Fadan, Y. M. (Yousef M. (1983) *The development of contemporary housing in Saudi Arabia (1950-1983) : a study in cross-cultural influence under conditions of rapid change*. Thesis. Massachusetts Institute of Technology. Available at: https://dspace.mit.edu/handle/1721.1/73256 (Accessed: 20 June 2021).

Fardous, I. S. (2019). *Adaptation of an architectonic tradition for a sustainable future in the Middle East: A case study of three building typologies in Riyadh City*. Ph.D. Dissertation. Robert Gordon University.

Felimban, A. et al. (2019). Assessment of current energy consumption in residential buildings buildings in Riyadh. *Space and Culture* [Online]. Available from: doi.org/10.1177/1206331220961264

Flick, U. (2018). *The SAGE handbook of qualitativedata collection.* London, UK: Sage.

Gaetani, I., Hoes, P. J. and Hensen, J. L. M. (2016). Occupant behavior in building energy simulation: Towards a fit-for-purpose modeling strategy. *Energy and Buildings* [online]. **121**, 188–204. Available from: doi.org/10.1016/j.enbuild.2016.03.038

Garba, S. B. (2004) 'Managing urban growth and development in the Riyadh metropolitan area, Saudi Arabia', *Habitat International*, 28(4), pp. 593–608. doi: 10.1016/j.habitatint.2003.10.008.

Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. Sebastopol, CA: O'Reilly Media.

Giddings, B., Almehrej, M. and Cresciani, M. (2020). The dilemma of Saudi Arabian homes in Ridyadh. *Space and Culture* [online]. Available from: doi.org/10.1177/1206331220961264

GNU Lesser General Public License v3.0. (2007). Available from: https://www.gnu.org/licenses/lgpl-3.0.en.html

Grube, E. J. and Michell, G. (1978) *Architecture of the Islamic world: its history and social meaning, with a complete survey of key monuments*. Morrow.

Hajj-Hassan, M. et al. (2020). A behavioral-based machine learning approach for predicting energy consumption. *Construction Research Congress 2020*. 1029–37. Available from: doi.org/10.1061/9780784482865.109

Han, M. et al. (2021). The reinforcement learning method for occupant behavior in building control: A review. *Energy and Built Environment* [online]. **2**(2), 137–48. Available from: doi.org/10.1016/j.enbenv.2020.08.005

Harbi, F. A. and Csala, D. (2019). Saudi Arabia's electricity: Energy supply and demand future challenges. In: *2019 1st Global Power, Energy and Communication Conference (GPECOM)*, 467–72. Available from: doi.org/10.1109/GPECOM.2019.8778554

Hong, T. et al. (2017). Ten questions concerning occupant behavior in buildings: The big picture. *Building and Environment* [online]. **114,** 518–530. Available from: https://doi.org/10.1016/j.buildenv.2016.12.006

Hong, T. et al. (2020). State-of-the-art on research and applications of machine learning in the building life cycle. *Energy and Buildings* [online]. **212**, 109831. Available from: doi.org/10.1016/j.enbuild.2020.109831

Horowitz, J. L. (2001). The Bootstrap. In: *Handbook of Econometrics,* J. J. Heckman and E. Leamer eds. Amsterdam, Netherlands: Elsevier.

Hu, M. and Xiao, F. (2020). Quantifying uncertainty in the aggregate energy flexibility of high rise residential building clusters considering stochastic occupancy and occupant behavior. *Energy* [Online]. **194**, 116838. Available from: doi.org/10.1016/j.energy.2019.116838

Hu, S. et al. (2020). A systematic review of occupant behavior in building energy policy. *Building and Environment* [online]. **175**, 106807. Available from: doi.org/10.1016/j.buildenv.2020.106807

Huchuk, B., Sanner, S. and O'Brien, W. (2019). Comparison of machine learning models for occupancy prediction in residential buildings using connected thermostat data. *Building and Environment* [online]. **160**, 106177. Available from: doi.org/10.1016/j.buildenv.2019.106177

Hyndman, R. J. and Koehler, A. B. (2006) 'Another look at measures of forecast accuracy', *International journal of forecasting*, 22(4), pp. 679–688. doi: 10.1016/j.ijforecast.2006.03.001.

International Energy Agency. (2014). Energy Efficiency Indicators [Electronic Resource]. *Fundamentals on Statistics*. Paris: International Energy Agency.

Ionescu, C. et al. (2015). The historical evolution of the energy efficient buildings. *Renewable and Sustainable Energy Reviews* [online]. **49**, 243–253. Available from: doi.org/10.1016/j.rser.2015.04.062.

JetBrains. (2021) *PyCharm: The Python IDE for professional developers* Retrieved October 3, 2021, from: https://www.jetbrains.com/pycharm/

Jia, M. and Srinivasan, R. (2020). Building performance evaluation using coupled simulationof EnergyPlus™ and an occupant behavior model. *Sustainability* [online]. **12**(10), 4086. Available from: doi.org/10.3390/su12104086

Jia, M. et al. (2021). Investigating the impact of actual and modeled occupant behavior information input to building performance simulation. *Buildings* [online]. **11**(32), Available from: doi.org/ 10.3390/buildings11010032

Jia, M., Srinivasan, R. S. and Raheem, A. A. (2017). From occupancy to occupant behavior: An analytical survey of data acquisition technologies, modeling methodologies and simulation coupling mechanisms for building energy efficiency. *Renewable and Sustainable Energy Reviews* [online]. **68**, 525–40. Available from: doi.org/10.1016/j.rser.2016.10.011

Jing, R. et al. (2017). A study on energy performance of 30 commercial office buildings in Hong Kong. *Energy and Buildings* [online]. **144**, 117–28. Available from: doi.org/10.1016/j.enbuild.2017.03.042

Kane, M. et al. (2014). Comparison of ARIMA and Random Forest time series models for prediction of Avian Influenza H5N1 outbreaks. *BMC Bioinformatics* [online]. **15**(1), 276. Available from: doi.org/10.1186/1471-2105-15-276

Korjenic, A. and Bednar, T. (2012). Validation and evaluation of total energy use in office buildings: A case study. *Automation in Construction* [online]. **23**, 64–70. Available from: doi.org/10.1016/j.autcon.2012.01.001

Krarti, M. (2019). Evaluation of Energy Efficiency Potential for the Building Sector in the Arab Region. Energies 12, 4279. https://doi.org/10.3390/en12224279

Krarti, M., Aldubyan, M., Williams, E. (2020). Residential building stock model for evaluating energy retrofit programs in Saudi Arabia. Energy 195, 116980. https://doi.org/10.1016/j.energy.2020.116980

Laaroussi, Y. et al. (2020). Occupant presence and behavior: A major issue for building energy performance simulation and assessment. *Sustainable Cities and Society* [online]. **63**, 102420. Available from: doi.org/10.1016/j.scs.2020.102420

Langevin, J., Wen, J. and Gurian, P. L. (2016). Quantifying the human–building interaction: Considering the active, adaptive occupant in building performance simulation. *Energy and Buildings* [online] **117**, 372–86. Available from: doi.org/10.1016/j.enbuild.2015.09.026

Li, H., Ng, S. T. and Skitmore, M. (2018). Stakeholder impact analysis during post- occupancy evaluation of green buildings—A Chinese context. *Building and Environment* [online]. **128**, 89–95. Available from: doi.org/10.1016/j.buildenv.2017.11.014

Li, N. et al. (2015). Probability of occupant operation of windows during transition seasons in office buildings. Renewable Energy [online]. **73**, 84–91. Available from: doi.org/10.1016/j.renene.2014.05.065

Li, X. and Yao, R. (2020). A machine-learning-based approach to predict residential annual space heating and cooling loads considering occupant behaviour. *Energy* [online]. **212**, 118676. Available from: doi.org/10.1016/j.energy.2020.118676

Li, Z., Han, Y. and Xu, P. (2014). Methods for benchmarking building energy consumption against its past or intended performance: An overview. *Applied Energy* [online]. **124**, 325–34. Available from: doi.org/10.1016/j.apenergy.2014.03.020

Liisberg, J. et al. (2016). Hidden Markov models for indirect classification of occupant behaviour. *Sustainable Cities and Society* [online]. **27**, 83–98. Available from: doi.org/10.1016/j.scs.2016.07.001

Maile, T., Fischer, M. and Bazjanac, V. (2007). Building energy performance simulation tools: A life-cycle and interoperable perspective. *Center for Integrated Facility Engineering (CIFE) Working Paper*. **107**, 1–49.

Martinaitis, V. et al. (2015). Importance of occupancy information when simulating energy demand of energy efficient house: A case study. *Energy and Buildings* [online]. **101**, 64–75.Available from: doi.org/10.1016/j.enbuild.2015.04.031

Mendes, G. *et al.* (2014) 'Regional analysis of building distributed energy costs and CO2 abatement: A U.S.–China comparison', *Energy and Buildings*, 77, pp. 112–129. doi: 10.1016/j.enbuild.2014.03.047.

Menezes, A. C. et al. (2012). Predicted vs. actual energy performance of non-domestic buildings using post-occupancy evaluation data to reduce the performance gap. *Applied Energy* [online]. **97**, 355–64. Available from: doi.org/10.1016/j.apenergy.2011.11.075

Mubarak, F. A. (2004). Urban growth boundary policy and residential suburbanization: Riyadh, Saudi Arabia. *Habit International* [online]. **28**(4), 456–591. Available from: doi.org/10.1016/j.habitatint.2003.10.010

Mubarak, F. and King, P. D. (2007). Cultural adaptation to housing needs: A case study, Riyadh, Saudi Arabia. *CiteSeer* [Online]. Available from: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.578.3267

Mun, S. H., Kwak, Y. and Huh, J. H. (2021). Influence of complex occupant behavior models on cooling energy usage analysis. *Sustainability* [online]. **1**(3), 1243. Available from: doi.org/10.3390/su13031243

Muroni, A. et al. (2019). Occupant behavior in identical residential buildings: A case study for occupancy profiles extraction and application to building performance simulation. *Building Simulation* [Online]. **12**(6), 1047–61. Available from: doi.org/10.1007/s12273-019-0573-x

Niu, S., Pan, W. and Zhao, Y. (2016). A virtual reality integrated design approach to improving occupancy information integrity for closing the building energy performance gap. *Sustainable Cities and Society* [online]. **27**, 275–86. Available from: doi.org/10.1016/j.scs.2016.03.010

Nurunnabi, M. (2017). Transformation from an oil-based economy to a knowledge-based economy in Saudi Arabia: The direction of Saudi vision 2030. *Journal of the Knowledge Economy* [online]. **8**(2), 536–64. Available from: doi.org/10.1007/s13132-017-0479-8

Oberkampf, W. L. and Roy, C. J. (2010). *Verification and validation in scientific computing.* Cambridge, UK: Cambridge University Press.

OpenStudio (2021) *OpenStudio*. Retrieved October 3, 2021, from https://openstudio.net/.

Paatero, J. V. and Lund, P. D. (2006). A model for generating household electricity load profiles. *International Journal of Energy Research* [Online]. **30**(5), 273–90. Available from: doi.org/10.1002/er.1136

Panchabikesan, K., Haghighat, F. and El Mankibi, E.(2021). Data driven occupancy information for energy simulation and energy use assessment in residential buildings. *Energy* [online]. **218**, 119539. Available from: doi.org/10.1016/j.energy.2020.119539

Paone, A. and Bacher, J. P. (2018). The impact of building occupant behavior on energy efficiency and methods to influence it: A review of the state of the art. *Energies* [online]. **11**(4), 953. Available from: doi.org/10.3390/en11040953.

Project Jupyter (2021) *Jupyter*. Retrieved October 3, 2021, from https://www.jupyter.org .

Renewable energy market analysis: The GCC Region. (2016). *International Renewable Energy Agency* [online]. Available from: https://irena.org/publications/2016/Jan/Renewable-Energy-Market-Analysis-The-GCC-Region

Rodriguez, J. D., Perez, A. and Lozano, J. A. (2010) ‘Sensitivity Analysis of k-Fold Cross Validation in Prediction Error Estimation’, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(3), pp. 569–575. doi: 10.1109/TPAMI.2009.187.

Rosenbaum, M. (2003). Understanding the energy modeling process: Simulation literacy 101. *The Pittsburgh Papers: Best of Green-Build 2003*. Pittsburgh, PA: BuildingGreen, Inc., 99–108

Saleh, M. A. (1997). Privacy and communal socialization: The role of space in the Security of traditional and contemporary neighborhoods in Saudi Arabia. *Habitat International* [online]. **21**(2), 167–84. Available from: doi.org/10.1016/S0197-3975(96)00055-0

Saleh, M. A. E. (2001). The changing image of Arriyadh City: The role of socio-cultural and religious traditions in image transformation. *Cities* [online]. **18**(5), 315–30. Availabble from: doi.org/10.1016/S0264-2751(01)00024-5

Sandels, C., Widén, J. and Nordström, L. (2014) ‘Forecasting household consumer electricity load profiles with a combined physical and behavioral approach’, *Applied Energy*, 131, pp. 267–278. doi: 10.1016/j.apenergy.2014.06.048.

Santamouris, M. (2014). On the energy impact of urban heat island and global warming on buildings. *Energy and Buildings* [online]. **82**, 100–113. Available from: doi.org/10.1016/j.enbuild.2014.07.022

Scikit-Learn: Machine Learning in Python (2021) Available from: https://scikit-learn.org/stable/

Space Syntax | Thriving Life in Buildings & Urban Places. (2021). *Space Syndax Symposium*. Istanbul, Turkey.

Sparks, R. S. (2021). A review of some sampling and aggregation strategies for basic statistical process monitoring: Discussion by Ross Sparks. *Journal of Quality Technology* [online]. **53**(1), 23–25. Available from: doi.org/10.1080/00224065.2019.1623588

SPSS Statistics. (2021) *Overview*. Retrieved October 3, 2021, from https://www.ibm.com/products/spss-statistics.

Sun, Y. et al. (2020). Data-driven occupant-behavior analytics for residential buildings. *Energy* [online]. **206**, 118100. Available from: doi.org/10.1016/j.energy.2020.118100

Sunikka-Blank, M. and Galvin, R. (2012). Introducing the prebound effect: The gap between performance and actual energy consumption. *Building Research and Information* [online]. **40**(3), 260–73. Available from: doi.org/10.1080/09613218.2012.690952

Tam, V. W., Almeida, L. and Le, K. (2018) ‘Energy-related occupant behaviour and its implications in energy use: A chronological review’, *Sustainability (Basel, Switzerland)*, 10(8), pp. 2635-. doi: 10.3390/su10082635.

Turner, C. and Frankel, M. (2008) ‘Energy Performance of LEED ® for New Construction Buildings’, *New Buildings Institute*, pp. 1–46.

UAE Government News. (2016). Dubai Supreme Council of Energy addresses latest updates in Dubai Clean Energy Strategy 2050.

Uddin, M. N. et al. (2021). Influence of occupant behavior for building energy conservation: A systematic review study of diverse modeling and simulation approach. *Buildings* [online]. **11**(2), 41. Available from: doi.org/10.3390/buildings11020041

Ullah, A. et al. (2020). Optimizing the efficiency of machine learning techniques. In: Y. Tian, T. Ma and M. K. Khan, eds. *Big Data and Security*. pp. 553–67. Singapore: Springer Singapore.

Vision Progress | Saudi Vision 2030. (2018). Available from: https://www.vision2030.gov.sa/

Wang, Z. and Srinivasan, R. S. (2017). A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models. *Renewable and Sustainable Energy Reviews* [online]. **75**, 796–808. Available from: doi.org/10.1016/j.rser.2016.10.079.

Wang, Z. et al. (2018). Random Forest based hourly building energy prediction. *Energy and Buildings* [online]. **171**, 11–25. Available from: doi.org/10.1016/j.enbuild.2018.04.008

WebStorm: The Smartest JavaScript IDE by JetBrains. (2021) JetBrains. Available from: https://www.jetbrains.com/webstorm/

Welcome to ESP-r. (2021). Available from: http://www.esru.strath.ac.uk/Courseware/ESP-/tour/

Westphal, F. S. and Lamberts, R. (2004). The use of simplified weather data to estimate thermal loads of non-residential buildings. *Energy and Buildings* [online]. **36**(8), 847–54. Available from: doi.org/10.1016/j.enbuild.2004.01.007

White, J. A., and Reichmuth, R. (1996). Simplified method for predicting building energy consumption using average monthly temperatures. In: *Proceedings of the 31st Itersociety Energy Conversion Engineering Conference*, 11–18 August.

Widén, J. and Karlsson, K. (2009) 'The influence of Swedish households' everyday activities and electricity-use patterns on the utilization of small-scale photovoltaic systems', in. *eceee 2009 Summer Study, La Colle-sur-Loup, France, June 1-6, 2009*. Available at: http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-122287 (Accessed: 25 March 2019).

Wilke, U. *et al.* (2012) 'A bottom-up stochastic model to predict building occupants' time-dependent activities', *Building and Environment*. doi: 10.1016/j.buildenv.2012.10.021.

Wogan, D., Carey, E. and Cooke, D. (2019). Policy pathways to meet Saudi Arabia's contributions to the Paris agreement. *Discussion Papers*. Available from doi.org/10.30573/KS--2019-DP49

Xie, Y. et al. (2018). Evaluation of machine learning methods for formation lithology Identification: A comparison of tuning processes and model performances. *Journal of Petroleum Science and Engineering* [online]. **160**, 182–93. Available from: doi.org/10.1016/j.petrol.2017.10.028

Yaman, R. et al. (2021). Post occupancy evaluation of space energy intensity on green building index energy efficiency (EE) criteria. *Engineering Journal* [online]. **25**(1), 233–43. Available from: doi.org/10.4186/ej.2021.25.1.233

Yan, C. et al. (2015). A multi-level energy performance diagnosis method for energy information poor buildings. *Energy* [online]. **83**, 189–203. Available from: doi.org/10.1016/j.energy.2015.02.014

Yan, D. et al. (2017) IEA EBC Annex 66: Definition and Simulation of Occupant Behavior. *Energy and Buildings* [online]. **156**, 258–70. Available from: doi.org/10.1016/j.enbuild.2017.09.084

Yoshino, H., Hong, T. and Nord, N. (2017). Total energy use in buildings—Analysis and evaluation methods. *Energy and Buildings* [online], **152**, 124–36. Available from: doi.org/10.1016/j.enbuild.2017.07.038

Yu, C., Du, J. and Pan, W. (2019). Improving accuracy in building energy simulation via evaluating occupant behaviors: A case study in Hong Kong. *Energy and Buildings* [online]. **202**, 109373. Available from: doi.org/10.1016/j.enbuild.2019.109373

Yu, Z. et al.(2011). A methodology for identifying and improving occupant behavior in residential buildings. *Energy* [online]. **36**(11), 6596–6608. Available from: doi.org/10.1016/j.energy.2011.09.002

Zou, P. X. W. *et al.* (2018) 'A mixed methods design for building occupants' energy behavior research', *Energy and Buildings*, 166, pp. 239–249. doi: 10.1016/j.enbuild.2018.01.068.

Zou, P. X. W., Wagle, D. and Alam, M. (2019). Strategies for minimizing building energy performance gaps between the design intend and the reality." *Energy and Buildings* [online]. **191**, 31–41. Available from: doi.org/10.1016/j.enbuild.2019.03.013

Appendix

Survey Approval letter:

Downloaded: 07/10/2021
Approved: 09/05/2018

Mohammed Aljammaz
Registration number: 160260683
School of Architecture
Programme: Architecture

Dear Mohammed

PROJECT TITLE: The Effect of the Islamic culture in the building energy simulation: Incorporating Saudi residents' domestic behaviour to reduce the prediction gap for the houses in Riyadh, Saudi Arabia
APPLICATION: Reference Number 018753

On behalf of the University ethics reviewers who reviewed your project, I am pleased to inform you that on 09/05/2018 the above-named project was **approved** on ethics grounds, on the basis that you will adhere to the following documentation that you submitted for ethics review:

- University research ethics application form 018753 (form submission date: 11/04/2018); (expected project end date: 30/12/2018).
- Participant information sheet 1042407 version 2 (11/04/2018).
- Participant consent form 1042406 version 2 (11/04/2018).

The following optional amendments were suggested:

Make sure the breadth and depth of data collection will be sufficiently representative of the average population. To what extent the conference event mentioned in the application will contribute to representation of the average population?

If during the course of the project you need to deviate significantly from the above-approved documentation please inform me since written approval will be required.

Your responsibilities in delivering this research project are set out at the end of this letter.

Yours sincerely

Chengzhi Peng
Ethics Administrator
School of Architecture

Please note the following responsibilities of the researcher in delivering the research project:

- The project must abide by the University's Research Ethics Policy: https://www.sheffield.ac.uk/rs/ethicsandintegrity/ethicspolicy/approval-procedure
- The project must abide by the University's Good Research & Innovation Practices Policy: https://www.sheffield.ac.uk/polopoly_fs/1.671066!/file/GRIPPolicy.pdf
- The researcher must inform their supervisor (in the case of a student) or Ethics Administrator (in the case of a member of staff) of any significant changes to the project or the approved documentation.
- The researcher must comply with the requirements of the law and relevant guidelines relating to security and confidentiality of personal data.
- The researcher is responsible for effectively managing the data collected both during and after the end of the project in line with best practice, and any relevant legislative, regulatory or contractual requirements.