In the preliminary data analysis, there was little difference between the actual usage and the default load schedules in the energy prediction software for bedrooms and family (daily use) areas. The main contributor to the observed mismatch was found to be the visitor zone. This occupies around 40 percent of the total area of a Saudi Arabian house. The survey data obtained in this study (Figure 5) showed that 51 percent of people used the visitor zones between one and ten times per year, 18 percent used them between 11 and 20 times per year, and 10 percent between 21 and 30 times a year. In addition, 65 percent of respondents had separate guestrooms for males and females. Furthermore, 42 percent mentioned that the dining room was only used when they had visitors.

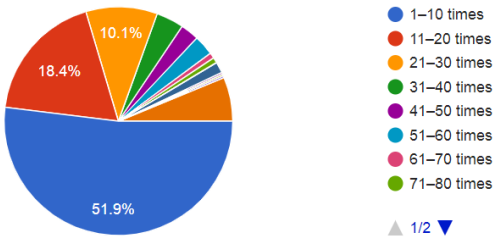

Fig.5. The use of the male guest room

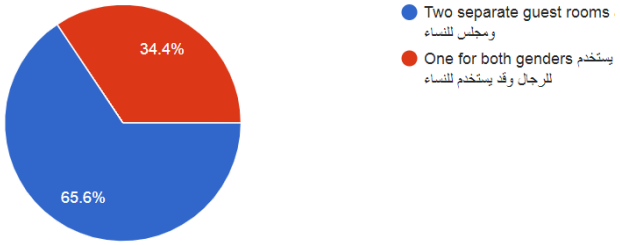

Fig.6. The number of guest rooms in the Saudi Arabian houses

Further analysis of data showed that there was a link between the age of the homeowner, the number of people living in the house, and the size of the guest zone (Figure 7). It was found that older people cared more about the size of the guest zone than younger people did. Younger people found a single guestroom and a shared dining room to be acceptable as they did not expect many visitors. Guest rooms could be used by either male or female visitors as long as their visits did not coincide. In other words, separation of male and female guest facilities was divided by time of occupation rather than physical space.

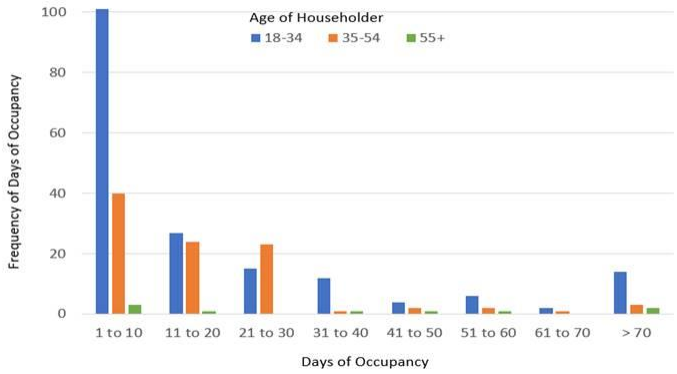

Fig.7. The relationship between the age of the house holder and the use of the guest rooms

Living hall

The space used for the 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 the living hall occasionally. The survey data showed that there was a relationship between the size of the house and how spaces were used for family gatherings (Figure 8): families living in houses smaller than 500 m^2 in size tended to use the living hall for family gatherings, while larger houses tended to have a separate family room.

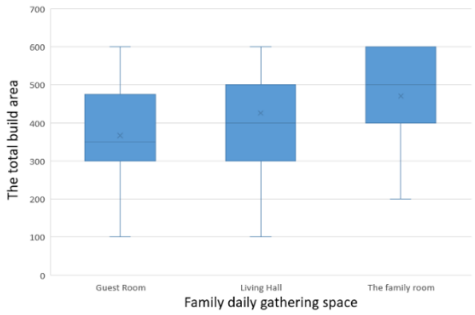

Fig. 8. Standard deviation for the total build area and the family gathering space.

Male guest zone

Figure 9 illustrates the number of times that the male guest room is occupied each year against the number of males per household. Most respondents reported that occupancy was less than ten times a year, but 20 percent reported usage in excess of 60 times a year. This indicates a major difference in the normal daily usage schedule

which is included in the energy-simulation software. Failure to recognise this variation leads to inaccurate energy estimates.

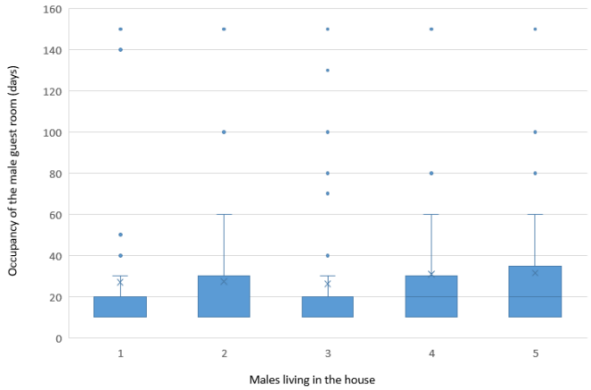

Fig. 9. Occupancy of the male guest room against the number of males per household

Female guest zone

Saudi Arabian culture impacts on building design. Traditional views requires that the female zone must not be visible from the male zone. The female zone has its own entrance, reception room, and bathroom. Nowadays, some houses only have one zone which is used by both males and females. Figure 10 shows that houses smaller than 300 m^2 are more likely to have just one zone for both genders while houses over 300 m^2 are more likely to have two zones. Figure 11 shows the relationship between the number of females living in the house and the use of the female guest room. The data shows that families who have one or two females in the house either have one zone for both genders or use the female guest room fewer than 20 times a year. Homes that have three to five female inhabitants use the female guest room, on average, 30 to 50 times a year.

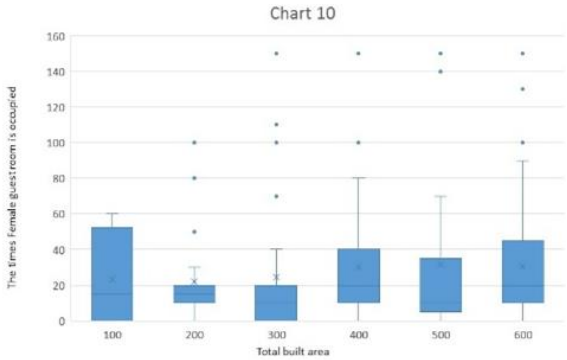

Fig. 10. Occupancy of the female guest room against the total build area.

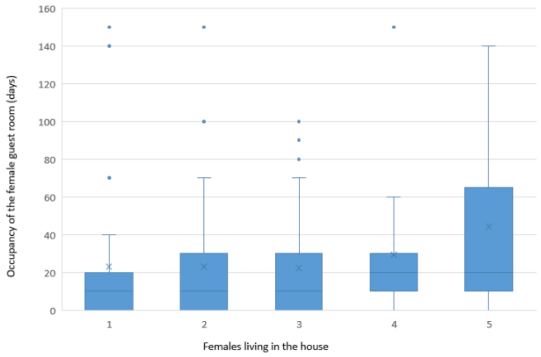

Fig. 11. Occupancy of the female guest room against the number of females per household

3.2 Create the model (Step 2)

Select a performance measure

Data for each space in the house were analysed in order to identify the mismatches between predicted and actual energy use. After testing the 12 algorithms identified in figure 12, it was found that the best results could be obtained by using linear regression method. It produced the lowest error among the algorithms tested. The Roof Mean Square Error (RMSE) was applied as a performance measure when conducting the analysis. The RMSE identifies how much error the system makes in predicting Y (**Y1** is the male zone, **Y2** is the female zone and **Y3** is the living hall).

$$y = \alpha + \beta x,$$

$$y_i = \alpha + \beta x_i + \varepsilon_i.$$

$$\hat{\varepsilon}_i = y_i - a - bx_i.$$

$$\sum_{i=1}^{n} \hat{\varepsilon}_i^{\,2} = \sum_{i=1}^{n} (y_i - a - bx_i)^2$$

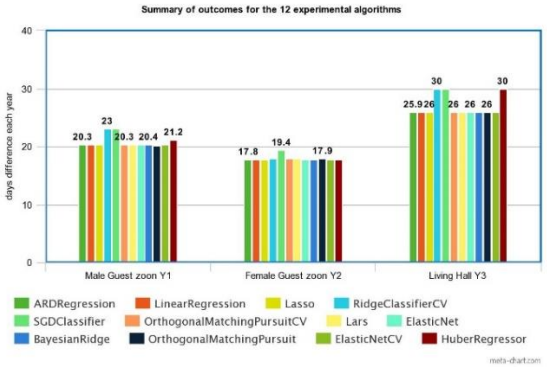

Fig. 12. Summary of outcomes for the 12 experimental algorithms summary

Obtain the data and create the workspaces

In this study, we used Python and Scikit-learn to develop a TUD model for estimating the occupancy schedules in Saudi Arabian houses. As mentioned in the previous section, the linear regression algorithm was used.
Figure 13 shows how the linear regression model delineated the relationship between the times when the male guest room was occupied (Y1) and the number of males living in the house(X1). By using the linear regression algorithm, it was possible to obtain an error estimation for Y1(male guest zone): with a mean squared error was 1659 with a square root of approximately 41. The estimation for the male guest room error, therefore, was 41 days, i.e. 20.5 days higher or lower than real energy use.

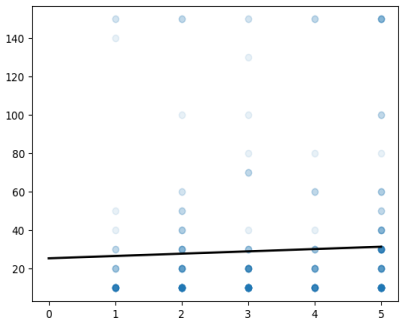

Fig. 13. Occupancy of the male guest room against the number of males living in the house

Occupation of the female guest room also had a strong relationship with the number of females living in the house (X2), the year the house was built (X3) and the number of times the female guest room was used (Y2). By using the linear regression algorithm, it was possible to obtain an error estimation for Y2: with a mean squared error was 1278.97 with a square root of approximately 35. The estimation for the female guest room error, therefore, was 35 days, i.e. 17.5 days higher or lower than real energy use. Figure 14 shows how the model determined the relationship between the times the female guest room was occupied and the number of females living in the house.

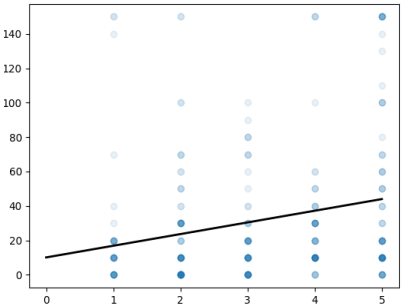

Fig. 14. Occupancy of the female guest room against the number of females living in the house.

The living room can be used for both visitors and family gatherings. Running the model, it could be established there was a strong relationship with the year the house was built and the size of the house: the year that the house was built as (X4) gave slightly lower error readings than the number of male living in the house (X5) and the size of the house (X6). Applying the male occupants as (X5), rather than female occupants, the rate of error was reduced slightly. By using the linear regression algorithm, it was possible to obtain an error estimation for Y3 (living room): with a mean squared error was 2692.79 with a square root of approximately 51. The estimation for the living room error, therefore, was 51days, i.e. 25.5 days higher or lower than real energy use.

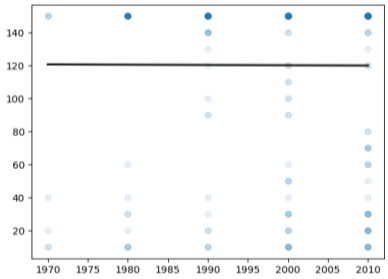

Fig. 15. Occupancy of the living room against the year that the house was built

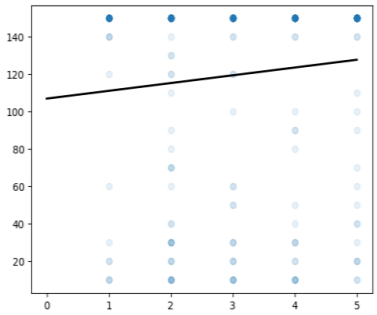

Fig.16. Occupancy of the living room against the number of males living in the house.

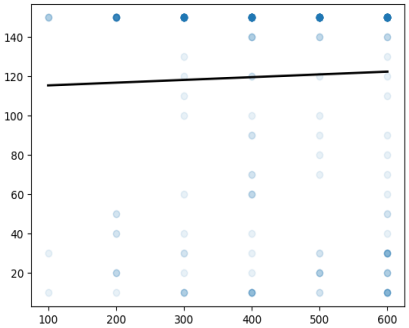

Fig.17. Occupancy of the living room against the size of the house

3.3 Testing the model (Step 3)

The adopted research method used 30 percent of the data to verify the linear regression model. Analysis showed that the model predicted a mismatch of 26 days between estimated and real energy usage.
The outcomes were applied the model presented by the author [16], which used a case study based on a villa in Riyadh. The villa analysed was built in 2005 and comprised a basement, ground floor, first floor and a second floor. It was occupied by six people (three males and three females). The villa's total floor area was 500 m^2. There was one guest room which was used for both male and female guests. In this case study, the guest room was used around 90 times a year. Applying the model predicted use of the guest room as 29.6 days for males and 31.7 days for females. This estimated a combined usage, therefore, of 61.3 days. This revealed a 28.7-day gap mismatch between estimated and real usage (i.e. 90 days – 61.3 days). For the living room, the model produced a more accurate prediction - 133 days of occupancy against a real measure of 140 days.

4 Conclusions and Future Work

This paper has shown how consideration of culture and behaviour in the load schedules for energy simulations can provide a more accurate prediction of energy use. Absence of this factor in energy simulations was identified as being particularly important in Saudi Arabia as the cultural use of living space differs from circumstances in other cultures. For instance, on average, people in Saudi Arabia use the visitor zone (which covers approximately 40 percent of the total house area) for maximum 52 days each year.

The study also showed that the number of males and females living in a house, the size of the house and the year of its construction had a significant impact on the use of the guest zones. The year the house was built is an important factor to be considered as Saudi Arabians tend to use the guest zone less frequently now. This is reflected in the design of more modern buildings.

The principal aim of collecting data was to develop a more sophisticated computational model that can produce a more accurate prediction of energy usage in the context of Saudi Arabian culture. The model presented here was successful in achieving this and the mismatch was reduced significantly. Using data collected from villas in Riyadh, the error in predicting energy usage was reduced by about 90 percent from 300 percent to 32 percent. (Table 1).

Table 1. Outcomes of the model comparing real energy usage using the default settings.

	Real energy uses for the male and female guest rooms (case study)	The model prediction	default schedule for energy plus
The day's space occupied each year	90 days	61.3 days	360 days
The gap between the prediction and the real use (%)	0%	32%	300%

For future study, it is intended that more focus is given to improving the online survey method by using open-ended questions. Other research will focus on methods for integrating load schedule predictions into energy simulation workflow. Together, these will provide greater insight into the use of space and home energy in Saudi Arabia and allow the simulation models to be improved further.

References

1. De Wilde, P.: The Gap between Predicted and Measured Energy Performance of Buildings: A Framework for Investigation, Automation in Construction no. 41: 40-49. https://doi.org/10.1016/j.autcon.2014.02.009 (2014a)
2. Martinaitis, V. and Edmundas, Z. and Violeta, M. and Tatjana V.: Importance of Occupancy Information When Simulating Energy Demand of Energy Efficient House: A Case Study, Energy and Buildings no. 101: 64-75. https://doi.org/10.1016/j.enbuild.2015.04.031 (2015)
3. Yu, Z. and Faribor, H. and Benjamin, C.M. and Edward, M. and Hiroshi Y.: A Methodology for Identifying and Improving Occupant Behavior in Residential Buildings, Energy no. 36 vol. 11: 6596-608 (2011)
4. Korjenic, A. and Thomas B.: Validation and Evaluation of Total Energy Use in Office Buildings: A Case Study, Automation in Construction no. 23: 64-70. https://doi.org/10.1016/j.autcon.2012.01.001 (2012)
5. Armstrong, M. and Mike, S. and Hajo, R. and Ian, B. and Jocelyn, M.: Synthetically Derived Profiles for Representing Occupant-Driven Electric Loads in Canadian Housing.' Journal of Building Performance Simulation no. 2, vol 1: 15-30. https://doi.org/10.1080/19401490802706653 (2009)
6. Yoshino, H. and Tianzhen, H. and Natasa N.: IEA EBC Annex 53: Total Energy Use in Buildings—Analysis and Evaluation Methods, Energy & Buildings no. 152: 124-36. https://doi.org/10.1016/j.enbuild.2017.07.038 (2017)
7. Sunikka-Blank, M. and Ray G.: Introducing the Prebound Effect: The Gap between Performance and Actual Energy Consumption, Building Research & Information no. 40 vol. 3: 260-73. https://doi.org/10.1080/09613218.2012.690952 (2012)
8. Al-Said, F.: 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 (2003)
9. Othman, Z. and Rosemary, A. and Laurie, B.: Privacy, Modesty, Hospitality, and the Design of Muslim Homes: A Literature Review. Frontiers of Architectural Research no. 4 vol. 1: 12-23. https://doi.org/10.1016/j.foar.2014.12.001 (2015)
10. Wilke, U. and Frédéric, H. and Jean-Louis, S. and Darren R.: A Bottom-up Stochastic Model to Predict Building Occupants, Time-Dependent Activities, Building and Environment. https://doi.org/10.1016/j.buildenv.2012.10.021 (2012)
11. Lamedica, A. Prudenzi, A.: A Bottom-up Approach to Residential Load Modeling.' IEEE Transactions on Power Systems no. 9 vol. 2: 957-64. https://doi.org/10.1109/59.317650 (1994)
12. Norford L.K, R.H. Socolow, E.S Hsieh, G. V. Spadaro. 1994. 'Two-to-One Discrepancy between Measured and Predicted Performance of a "low-Energy" Office Building: Insights from a Reconciliation Based on the DOE-2 Model.' Energy & Buildings no. 21 vol. 2:121-31. https://doi.org/10.1016/0378-7788(94)90005-1.
13. Paatero Jukka V, Peter D. Lund, 2006. 'A Model for Generating Household Electricity Load Profiles.' International Journal of Energy Research no. 30 vol. 5: 273-90. https://doi.org/10.1002/er.1136.

14. Tanimoto, Jun, Aya Hagishima, and Hiroki Sagara. 2008b. 'A Methodology for Peak Energy Requirement Considering Actual Variation of Occupants' Behavior Schedules.' Building and Environment no. 43 vol. 4: 610-19. https://doi.org/10.1016/j.buildenv.2006.06.034.
15. Widén, J. and Andreas, M. and Kajsa, E.: Models of Domestic Occupancy, Activities and Energy Use Based on Time-Use Data: Deterministic and Stochastic (2012)
16. Richardson, I. and Murray, T. David, I.: A High-Resolution Domestic Building Occupancy Model for Energy Demand Simulations, https://doi.org/10.1016/j.enbuild.2008.02.006 (2008).
17. Richardson, I. and Murray, T. and David, I. and Alice, D.: Domestic Lighting: A High-Resolution Energy Demand Model, Energy & Buildings no. 41 vol. 7: 781-89. https://doi.org/10.1016/j.enbuild.2009.02.010 (2009a)
18. Richardson, I. and Murray, T. and David, I. and Alice, D.: A Modelling Framework for the Study of Highly Distributed Power Systems and Demand Side Management, In Sustainable Power Generation and Supply, 2009. SUPERGEN '09. International Conference On, 1–6. IEEE Publishing 10.1109/SUPERGEN.2009.5348274 (2009b)
19. Tanimoto, J. and Aya, H. and Hiroki, S.: Validation of Probabilistic Methodology for Generating Actual Inhabitants Behavior Schedules for Accurate Prediction of Maximum Energy Requirements.' Energy & Buildings no. 40 vol. 3: 316-22. https://doi.org/10.1016/j.enbuild.2007.02.032 (2008a)
20. Stokes, Melody, Mark Rylatt, Kevin; Lomas. 2004. 'A Simple Model of Domestic Lighting Demand.' Energy & Buildings no. 36 vol. 2: 103-16. https://doi.org/10.1016/j.enbuild.2003.10.007.
21. Author et al.: Understanding the Occupant Behaviours in Islamic Homes to Close the Gap in Building Performance Simulation: A Case Study of Houses in Riyadh, IBPSA2019 (2019)
22. Al-Hussayen. M.: Building Problems in Saudi Arabia: The Need for Building Research, and the Development of Building Research Approaches, PhD Dissertation, The University of Michigan, Ann Arbor (1980)
23. Al-Hemaidi, W.: The Dilemma of Regulating Privacy: Planning Regulations, Privacy and House Form; The Case Study of Low-Density Single-Family Dwellings in Saudi Arabia, PhD Thesis, University of London (1996)