

## Development of a Modelling Framework for Refined Residential Occupancy Schedules

Diba Malekpour Koupaei<sup>1</sup>, Farzad Hashemi<sup>2</sup>, Vinciane Tabard-Fortecoëf<sup>3</sup>, Ulrike Passe<sup>1</sup>

<sup>1</sup>Iowa State University, Ames, USA

<sup>2</sup>Penn State University, State College, USA

<sup>3</sup>INSA Lyon, Lyon, France

### Abstract

Occupancy schedules are a leading source of uncertainty in building energy simulations. This uncertainty often results in an undesired gap between actual performance and predictions. Past efforts to develop more proper occupancy behavioural models have often been either oversimplified and hence underestimating or more precise but simply too complicated to become common practice. This study is an effort to find a balance between complexity and accuracy among the aforementioned models for the sole purpose of energy use studies in the context of residential neighbourhoods. We start with a probabilistic occupancy model based on the American Time Use Survey (ATUS) and use input from the population of study to develop representative occupancy schedules.

### Introduction

According to the 2018 Annual Energy Outlook (U.S. Energy Information Administration, 2018), the building sector (residential and commercial) accounted for more than 1/4th of the total U.S. delivered energy in 2017. Moreover, projections for the future are unfortunately just as disappointing and unsettling as the current state of consumption. Consequently, energy use in buildings has been a growing concern of both the public and professionals in the field for quite some time now (Yan et al., 2017). So far, the majority of the previous efforts have been concerned with developing state of the art physical, technological and economical advances to minimize the energy use in the buildings sector. However, energy use in buildings is not merely a technological problem but more importantly, it is a social phenomenon that demands more attention. After all, “buildings don't use energy: people do” (Janda, 2011).

While historically, a tendency existed to direct a disproportionate share of attention towards energy systems or technological efficiency improvements and the human dimension of energy use in buildings was often neglected and overlooked, recent years witnessed an increase in the body of literature concerned with energy-related occupant behaviour in buildings (Yan et al., 2017). It is now an established fact that whether one is concerned with studying the current state of energy use in buildings or is proposing energy saving measures for future improvements, an understanding of occupancy behaviour and its implications for the energy demand is vital

(Mahdavi, Lambeva, Mohammadi, Kabir, & Pröglhöf, 2007). Some researchers even go as far as suggesting that residential utility demand profiles are mostly shaped by their corresponding occupancy patterns. For instance, the Tyndall Centre's report on microgrids (2005) states that “electricity load profile depends mainly on the household size and occupancy pattern.” (Abu-Sharkh et al., 2005)

As a result, today almost all energy modelling and simulation software tools use some sort of data linked to occupant behaviour and treat it as a defining parameter in their calculations of yearly profiles for heating, cooling, lighting, ventilation, and even plug loads. The most common form of such occupancy data is known as “diversity profiles” which is a schematic occupant presence profile of a space or thermal zone over a given period of time. Such profiles intend to reproduce the real occupancy of the space in order to accurately estimate the impact of peoples' presence and activity levels on energy load demand calculations of buildings (Abushakra, Sreshthaputra, Haberl, & Claridge, 2001). These profiles usually consist of a combination of weekday and weekend schedules for the particular type of building (for instance residential or commercial) in discussion. Software users are often provided with the choice of using the predefined generic schedules in the simulation tool's default library or defining their own profiles instead. Although the second option is meant to give the user flexibility and higher precision, the reality is that high quality occupancy data in all its stochastic variety is scarce (Paatero & Lund, 2006) and often times, the user is left with no choice but to use the predefined generic schedules.

While the inclusion of occupancy inputs in energy simulations is an undeniably big step forward, the remaining problem is that the use of such generic schedules as an input to the model is a leading source of uncertainty that results in large undesired gaps between the predicted and actual energy use of buildings (Page, Robinson, Morel, & Scartezzini, 2008). This issue becomes even more crucial when the designated population for a project has unique characteristics and behavioural patterns. For instance, the presented study focuses on a compact low-income residential neighbourhood where, as will be discussed in the next sections of the manuscript, the behavioural patterns are far from what would be considered a typical lifestyle as defined by current standards and guidelines. This means that if we were to use generic schedules as the input to our

model, we would be dealing with an even higher level of uncertainty.

As a consequence of everything discussed up until this point and with the global aim of enhancing simulation approaches, multiple efforts have been made to generate high-resolution occupancy schedules and use them as a substitute for generic predefined schedules. A quick review of these efforts (please refer to section 4 of Yan et al., 2015 for more details), including but not limited to Agent-Based Models (ABM) or the probabilistic models, would reveal that the developed techniques tend to be relatively more precise and representative of reality (Yan et al., 2017). However, such models' accuracy and precision comes at the price of them being overwhelming, complicated and in uncompromising need of a high-resolution database of large magnitude. All of this means that such techniques are still far from being practical enough to become common practice.

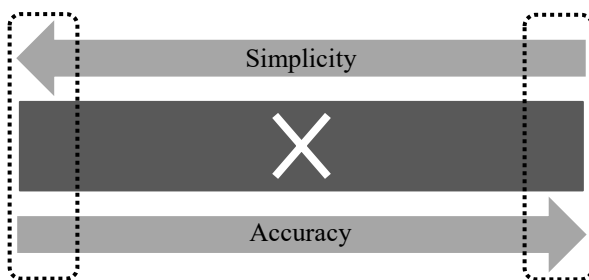


Figure 1: A schematic diagram of available occupancy models.

In other words, if we were to put all the proposed occupancy behavioural models on a spectrum of accuracy, previously discussed generic and simplified occupancy schedules would land on one end, while recently developed sophisticated models would probably land on the opposing end of the spectrum (Figure 1). None of these two sets of options represents an ideal scenario where a balance between complexity and accuracy exists. Moreover, when faced with this wide range of occupancy behavioural models, and considering the advantages and shortcomings of each type, the question of efficiency arises with every choice, (Mahdavi, 2011). An important aspect to note here is that this choice is not independent of the types of queries that the simulation model is expected to provide answers for (Mahdavi & Tahmasebi, 2016). This means that the task of finding the point of balance between practicality and accuracy of occupancy models is context-dependent. For instance, in this study, we were aiming to find this point of balance for an existing residential neighbourhood in an urban context. What we propose here might not be ideal for other simulation queries and future effort is needed to specifically address such needs.

In the following sections, we will first introduce our research query and goals. Then, we explain the proposed technique in high detail and end by testing the efficiency and performance of our proposed technique in a pilot case study with the help of Urban Modeling Interface (*umi*) (Reinhart, Dogan, Jakubiec, Rakha, & Sang, 2013).

## Method

### Research query and introduction to the case study

This study focuses on the use of occupancy presence schedules in the energy use simulation of a predominantly residential neighbourhood in Des Moines, Iowa. The Capitol East neighbourhood (Figure 2), which is the pilot study area for this urban energy simulation, is located just east of the State of Iowa Capitol complex, near downtown Des Moines. The Capitol East community is primarily of low income, and their settlement pattern in this neighbourhood is quite dense when compared to other parts of the city (Iowa State University Planning Team, 2014).

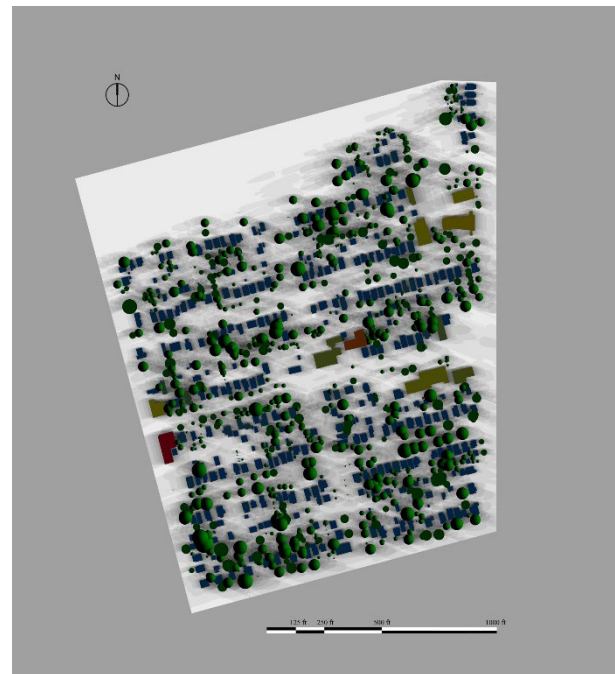


Figure 2: Top view of the Capitol East neighborhood as modeled in the *umi* environment.

Overall, our ultimate goal is to develop relevant and promising retrofit and climate change adaptation strategies that would guide our population towards a more sustainable future. Since developing such strategies is deeply tied to an understanding of the current energy-related behavioural patterns prevalent among the residents, our goal in this study is to identify these patterns and use the findings as a communication tool for community outreach and stakeholders' decision making.

### A description of the proposed technique

In the previous sections of this manuscript, we mentioned that more recent sophisticated occupancy models are successfully able to represent occupants' energy-related behaviours in high resolution. However, these models are often developed on the basis of high-resolution databases of large magnitude and require complex calculations. Building such databases as input and completing the required calculations for every project can be overwhelming and in many cases impossible. This is perhaps why most of these models only exist in scholarly publications and have not been able to replace relatively

less accurate but more simplified static models. However, we hypothesized that the use of a previously defined model based on a publicly available dataset could be beneficial if we were to refine its results according to project-specific inputs. This procedure would let us take advantage of the accuracy of such sophisticated models without jeopardizing the overall practicality of the process.

Here we used a probabilistic occupancy model based on the American Time Use Survey (ATUS) (U.S. Bureau of Labor Statistics, 2017) and refined its results according to input from our own study population based on a neighbourhood wide survey. Probabilistic models use statistical data to predict the probability that certain behaviour, in this case leaving and returning home, occurs. The stochastic process involved in the probabilistic models' calculations considers the occupancy status as a random variable and at each time step, the model determines the probability of presence according to the previous status. It is important to note that it is possible that non-deterministic factors would influence the presence status as well. Hence, even with the same initial condition, different directions with different possibilities may be achieved (Yang, Santamouris, & Lee, 2016). The Markov Chain model type, which is the specific probabilistic model type we chose to use in this study, is one of the most applied stochastic ABMs. Previous studies have reported an average of 73% accuracy on the occupancy number detection by Markov models (Dong et al., 2010).

The Markov Chain transition probability in the selected probabilistic model was applied to the 2017 ATUS database. ATUS is a publicly available survey of how people use their time and includes detailed 24-hour diaries, completed at pre-defined time intervals by many thousands of participants. Usually, a Time Use Survey (TUS) includes the location of the participants at each time step in the diary, and can thus be used to identify the number of present occupants in a residential unit (Richardson, Thomson, & Infield, 2008). Therefore, TUSs are completely in line with the goals of a probabilistic occupancy model. A complete description of the probabilistic model used and its preliminary findings can be found in Malekpour Koupaei et al., 2019.

Earlier, we mentioned that this is a two-step process: obtaining an original model (previously defined based on a publicly available database) and refining its output based on input from the project's sample. However, the need for refining the results of a highly accurate model could be questioned. The answer to this question lies in the large scope of the TUSs. TUSs are usually conducted at national level and their results can only represent a typical lifestyle in the same level. When a project considers a non-traditional population with unique characteristics and behavioural patterns, such results can be misleading, under- or overestimating and non-representative (Abraham, Maitland, & Bianchi, 2006). This is why it is important to go one step further and refine the results of this sophisticated model to find representative matches.

The main reference for the refinement process in this study is a survey conducted by the Sustainable Cities Research Group at Iowa State University (Iowa State University, n.d.) in the selected neighbourhood to understand how residents make energy-related decisions in their houses and make use of HVAC and lighting systems. The survey was sent to about 1,000 household addresses in three adjoining neighbourhoods (i.e., Capitol East, Capitol Park, and MLK Jr Park) in the Des Moines metropolitan area. Although the sample size seems reasonably large, the response rate for this survey, calculated as the number of completed forms divided by the eligible sample size, was only 6.3%. This is surprisingly low, given the fact that this survey was purposefully designed to be simple, straight forward and quick. This rather low response rate is further validation of our initial hypothesis that acquiring enough data for building a customized sophisticated model from scratch is time consuming, expensive and in short, impossible for smaller projects like this. What this number indicates is the need for developing an accurate enough technique that requires a smaller database as its input.

Since no other database for this type of neighbourhood is available, the survey response data was taken as the basis for this project regardless of the low response rate. In our survey, one question (and its corresponding answers) is particularly relevant to occupancy profiles and thus of interest for the research project at hand. The aforementioned question was:

Question 1: In an average week:

- a. What percent of your Monday-Friday daytime hours is spent at home? (WDD)
- b. What percent of your Monday-Friday evening hours is spent at home? (WDN)
- c. What percent of your weekend daytime hours is spent at home? (WED)
- d. What percent of your weekend evening hours is spent at home? (WEN)

Responses to this question (Question 1) were diverse, covering a range of all the possible values between 0% and 100%. Therefore, creating one typical aggregated schedule with the help of an arithmetic average of the reported percentile numbers would have sacrificed this diversity in behaviour among the residents. Accordingly, what we needed here was a number of reliable and representative common schedules generated by a clustering/classification method and not a single schedule generated by averaging all the answers. Our initial concept for this clustering step was to find the link between the respondents' answers to this question and some of their general characteristics as reflected in other parts of the survey. These characteristics, which included respondents' ages, genders, economic activities and education levels, were addressed with the following questions in our survey:

Question 2: What is your gender?

1 = Male

2 = Female

3 = Other, Non-binary

Question 3: What is your age category?

1 = 18-30

2 = 31-40

3 = 41-50

4 = 51-60

5 = 61-70

6 = 71-80

7 = 81 or older

Question 4: What is the highest degree or level of school you have completed?

1 = Did not complete high school

2 = High School or equivalent (GED)

3 = Some college, no degree

4 = Trade/Technical/Vocational training

5 = Associate degree (2-year)

6 = Bachelor's degree (4-year)

7 = Master's degree

8 = Professional or doctorate degree

Question 5: What is your current employment status?

1 = Employed for wages

2 = Self-employed

3 = Unemployed and looking for work

4 = Unemployed but not looking for work

5 = Homemaker

6 = Student

7 = Military

8 = Retired

9 = Unable to work

As can be seen here, these questions (Questions 2-5) were all multiple-choice and finding any type of link between these answers and that of the presence rate question (Question 1) would have facilitated this desired clustering. However, none of the general characteristics addressed by the questions above (Questions 2-5) seemed appropriate and relevant in terms of explaining the differences between presence rates on its own. In other words, we were not able to identify a simple direct correlation between any of the respondents' pre-clustered groups (as defined by Questions 2-5) and their answers to the presence rate question (Question 1). Instead, what proved to be relatively more successful was using a combination of these respondent characteristics as the defining factors for the presence rate. Here we used the "rpart" package in R (Therneau, Atkinson, & Ripley, 2010) to cluster our data into different groups with the help of a regression-based decision-tree classification

method. To avoid overfitting of the model, we set the complexity parameter to 0.1 and generated the best reasonably sized classification trees for each of the four presence rates separately. These classification trees are represented in Figure 3 (a-d).

*Table 1: Correlation between presence rates and respondent characteristics.*

Regression Method		Correlation			
		Q 2	Q 3	Q 4	Q 5
Least Squares Line	WDD	0.27	0.27	0.14	0.57
	WDN	0.19	0.10	0.06	0.42
	WED	0.21	0.29	0.09	0.46
	WEN	0.15	0.22	0.08	0.39
Best Classification Tree	WDD	0.62			
	WDN	0.63			
	WED	0.46			
	WEN	0.41			

One aspect to consider here is that association is not causation. For instance, an apparent association between education level and daytime presence rate must not be taken for granted. What it means in this context is that we were able to explain the differences in daytime presence rates in our dataset with the help of education level and this may not hold true in cases not included in our dataset. Moreover, one cannot argue that because a classification tree suggests that level 1 of education is linked to a presence rate of 50.89%, education level of 1 is what causes that presence rate to be 50.89%. Following the logic of classification criteria for the four trees below (Figure 3), we were able to identify three weekday and three weekend schedule types prevalent in the sample.

Then, each of these schedule profiles was developed based on the following 3-step procedure:

Step 1: First, for ease of use all the presence rates were rounded (up or down) to their closest multiple of five percentile (Table 2). Now, each one of the desired day type schedules is recognized by two variables: a mean daytime presence rate (MDPR) and a mean nighttime presence rate (MNPR). If these two were to be translated into hourly values (which is the format that most energy simulation tools accept as input for occupancy schedules), daytime presence rate would be of a dynamic nature, while nighttime presence rate can be considered static in most cases. This is due to the fact that daytime is usually a vibrant period of time for a household where changes in the hourly presence rates are expected, while the nighttime period is considered to be of a more stable nature (López-Rodríguez, Santiago, Trillo-Montero, Torriti, & Moreno-Munoz, 2013). Therefore, we decided to use the ATUS-generated schedules and refine them for the daytime period only. Step 3 explains the necessary modifications for all hourly rates that fall into the nighttime period.



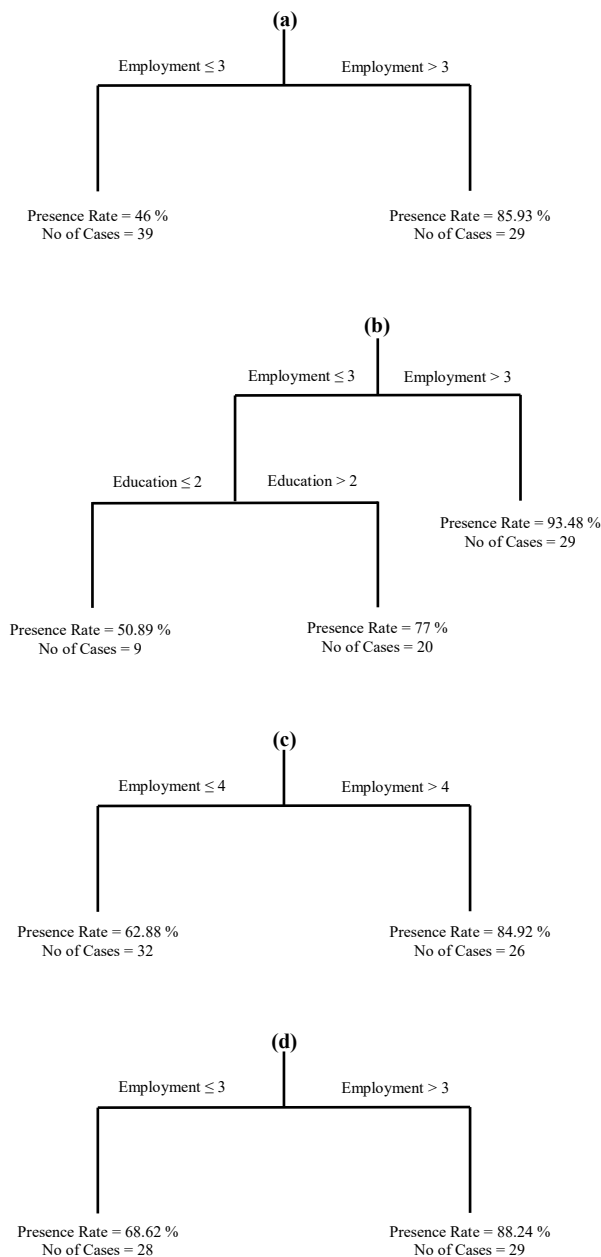


Figure 3: Best classification trees for (a) weekday daytime (WDD), (b) weekday nighttime (WDN), (c) weekend daytime (WED) and (d) weekend nighttime (WEN) mean presence rates.

Table 2: 24-hour schedules and their characteristics.

Day Type	24-hour Schedule ID	Presence Rate	
		MDPR	MNPR
Weekday (WD)	WD-a	45	50
	WD-b	45	75
	WD-c	85	95
Weekend (WE)	WE-a	65	70
	WE-b	65	90
	WE-c	85	90

Step 2: Typically, the probabilistic TUS-based models require two inputs in order to create a customized 24-hour schedule: number of people in the household and the type of day (weekday/weekend). Accordingly, we first generated 100 ATUS-based occupancy profiles for weekday schedules and another 100 for weekend schedules of three-person households (the number of people in each household was identified to be around 2.5, for more information please refer to Iowa State University Planning Team, 2014), to find matches for the desired MDPR values defined in the last step. Then, these generated schedules were arranged in terms of their daytime presence percentage and averaged using the following criteria:

“IF  $(MDPR - 5\%) \leq MDPR \leq (MDPR + 5\%)$   
 THEN average all”

The logic behind this extra step was to avoid using a rare occurrence of a specific schedule and receive a more common profile instead. This was necessary since each schedule types was intended to represent the common occupancy profile of its group and not of a specific individual. It is worth noting here that this method does not work for daytime presence rates smaller than 5% or larger than 95%. This should not come as a surprise, given the fact that such low/high presence rates hardly allow room for any changes in their corresponding hourly presence values. Therefore, when such presence rates are desired, the hourly rates should be set as constants equal to the determined MDPR or MNPR instead.

Step 3: Finally, we modified the nighttime presence rates to match what Table 2 indicates as the MNPR. All the hourly rates that fall into the nighttime period were set as constants equal to the MNPR value.

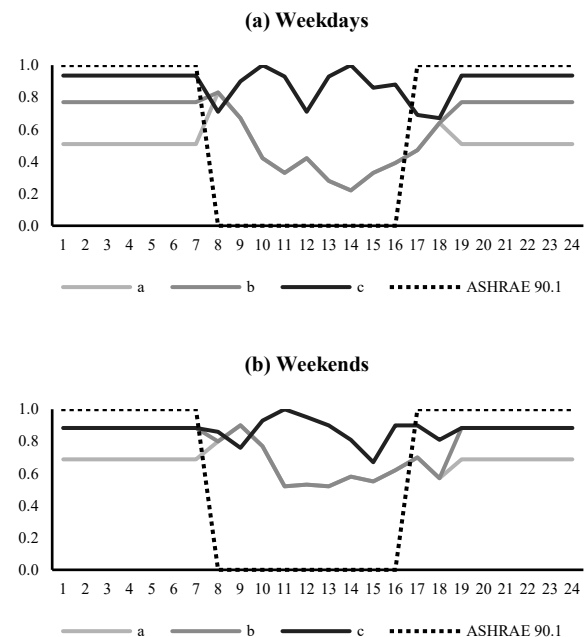


Figure 4: 24-hour schedules and their hourly presence rates for (a) weekdays and (b) weekends. The dotted line represents the standard ASHRAE 90.1 schedule for multifamily buildings.

Accordingly, six 24-hour schedule profiles as visualized in Figure 4 (a-b) were defined in the *umi* template library editor interface, where information about materials, constructions, schedules, thermal loads, spaces and buildings is stored (Cerezo, Dogan, & Reinhart, 2014). This template library editor is purposefully designed to be separate from the actual modeling interface of the *umi* software, to allow better cooperation between designers and those who are assigned the task of template generation and modification.

### Technique application and model development

Lighting, plug loads, and domestic hot water schedules were developed based on the Building America (BA) Data for the Des Moines area (Hendron & Engebrecht, 2010). A point to consider here is that the normalized hourly values in the BA profiles are set as such to be representative of that specific one-hour period's share of energy use as a fraction of the total daily use. However, *umi* uses another definition of hourly use, according to which each hourly value stands for that hour's energy use as a fraction of the maximum hourly use available. Therefore, BA profiles were accordingly scaled to match the needs of the *umi* software. After defining both daily and yearly periods of cooling, heating, and natural ventilation using Climate Consultant and Typical Meteorological Year (TMY) datasets for Des Moines (Milne, 2016), we were able to create the required heating, cooling, and natural ventilation schedules as well.

To generate weekly schedules out of these 24-hour schedules, one can argue that there are nine possible combinations available. However, only four of these nine schedules were logically possible (based on the logic behind the generation of each of the weekday/weekend schedules separately). Moreover, the occurrence probability of each of these week schedules was different and depended on the number of cases that represented this particular week schedule in our original dataset. Our model needed to reflect this variety in the probabilities among different schedules. Otherwise, we would have not been able to reach a realistic view of our neighborhood's unique characteristics. The following table describes these four types of week schedules and their characteristics.

Table 3: Week schedule and their characteristics.

Week ID	Probability (%)	Composition	
		Weekday ID	Weekend ID
W-1	15 %	a	a
W-2	35 %	b	a
W-3	5 %	c	b
W-4	45 %	c	c

The four resulting building templates (per construction template) were then assigned to the 272 residential buildings in our neighborhood with the help of a randomizing script in *Grasshopper* (this script takes advantage of the "Heteroptera" randomizing plugin (Bahrami, 2018)). According to this *Grasshopper* script,

our buildings were first clustered into three different groups based on their size and then, taking the probability of each schedule's occurrence into account (as represented in Table 3), a template was randomly assigned to that building within its own size-defined group. The criteria for clustering buildings in this step was defined by size. Buildings smaller than the 25% quantile of all residential buildings considered in this study were clustered into the "small buildings" group, while those that fell into the upper 25% quantile were considered "large buildings" and all others were labeled as "medium buildings". This extra caution for the random distribution was deemed necessary because building sizes were quite diverse and a homogeneous distribution of the schedules according to their probabilities would not have been possible otherwise.

### Results

In order to test the performance of the schedules defined and generated with the technique described above, five runs of the study neighbourhood were simulated in the *umi* environment. These five runs all shared the same geometry and material inputs but were assigned different occupancy schedules. This assignment of occupancy schedules was based on a single probability distribution function with the seed in the randomizing script in *Grasshopper* changing for each run.

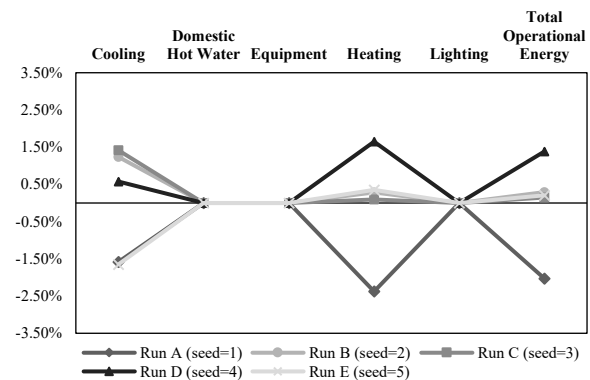


Figure 5: Comparison between the changes of different energy load components from their arithmetic mean among the five randomized simulation runs.

Overall, the results of these energy simulation runs were very close in terms of both their total yearly energy consumption values and the composition of their operational energies. As can be seen from Figure 5 above, total operational energy loads for each of these runs are almost equal in all cases and no more than 2% apart. With regard to the components of this total value, equipment, lighting, and domestic hot water loads are not directly impacted by the occupancy schedule changes. Moreover, of all the other load components calculated by *umi* that did change, none of them witnessed a change from the mean (of all five sets of results) that is equal to or higher than 2.4%. Therefore, energy load composition also remained the same and hardly changed between different runs. This suggests that the randomization process was successful in creating a homogeneous distribution of the generated schedules.

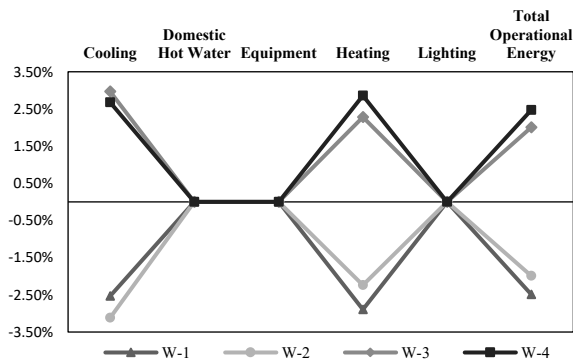


Figure 6: Comparison between the changes of different energy load components from their arithmetic mean among the four non-diverse simulation runs.

If however, we were to use any single one of our introduced schedules for all the residential buildings modeled (instead of using a combination of all four of our defined schedules), the disparity between the results would have been much higher and thus non-negligible (Figure 6). This suggests that it was crucial for our model to reflect the diversity in behavior and a combination of all the schedules did in fact provide us with results that are more realistic representations of the sample.

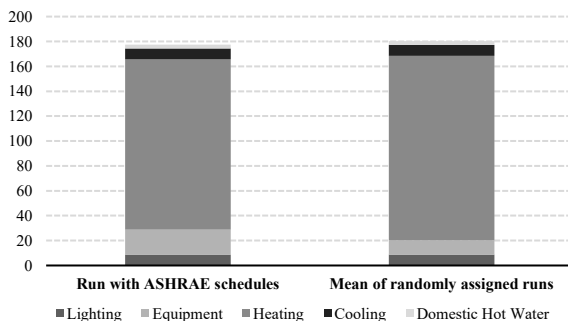


Figure 7: Total energy consumption comparison between the mean of randomized simulation runs and a standard run (normalized kWh per m<sup>2</sup>).

Finally, if we compare the arithmetic mean of the results of the five randomized runs with a simulation run that uses ASHRAE 90.1 standard occupancy schedule (ASHRAE, 1989) as its input, total energy consumption decreases by nearly 2%. This 2% difference translates into a 3.5 kWh/m<sup>2</sup> gap between the two modeling approaches (Figure 7). Therefore, the small but crucial adjustment of developing representative schedules was a big step towards bridging the gap between simulation results and actual energy use of these residential buildings.

Finally, it has been suggested before that there are three major dimensions of model resolution: (1) temporal, (2) spatial, and (3) occupancy (Melfi, Rosenblum, Nordman, & Christensen, 2011). Temporal resolution refers to the precision with which the timing of events is modeled. Spatial resolution refers to the precision of the physical scale. Finally, occupancy resolution refers to how the model specifies people. On this basis, Yan et al. (2015) developed evaluation criteria for occupancy models. When comparing our presented technique with their

evaluation criteria, it can be seen that we were able to maintain a relatively high level of temporal, spatial and state resolution for generating our occupancy schedules without jeopardizing the overall simplicity and practicality of the process (Figure 8).

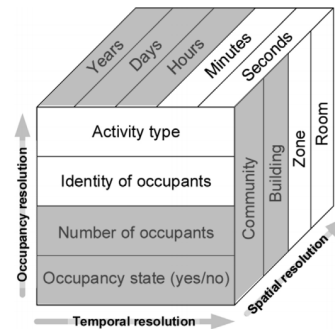


Figure 8: An illustration of our technique's temporal, spatial, and state resolutions. Based on and adapted from Yan et al. (Yan et al., 2015).

## Conclusion

This paper describes a process to develop occupancy schedules for energy use studies of existing residential neighborhoods based on residents' specific behavioral characteristics. The goal was to balance between accuracy and complexity of an occupancy model by developing a technique that takes advantage of a sophisticated probabilistic model based on the ATUS and then refines its results according to locally collected data. Application of this technique in a pilot case study showed a difference of nearly 2% in yearly energy consumption when compared to the use of a standard generic schedule. Since our developed schedules are more representative of the actual energy use patterns of the selected population, their use in the model gave us a more realistic view of the current state of energy consumption in the neighborhood. This process can be automated and adapted to other similar simulation queries, provided that project-specific input is gathered from the population. Current limitations are due to the fact that our findings are yet to be validated with actual metered energy consumption data. Our future work will use aggregated energy use data by zip code provided by the utility companies involved in the region to address this shortcoming.

Finally, the developed methodology and resulting preliminary data can serve as communication tools for community outreach. For instance, neighborhood-specific retrofit strategies can be developed, which would relate more appropriately to the actual characteristics of residents' behaviors and would thus be more realistic and thus successful.

## Acknowledgement

This work was funded by the 2016 Iowa State University Presidential Interdisciplinary Research Initiative (PIRI) on Data-Driven Science.

## References

- Abraham, K. G., Maitland, A., & Bianchi, S. M. (2006). Nonresponse in the American time use Survey: Who

- is missing from the data and how much does it matter? *International Journal of Public Opinion Quarterly*, 70(5), 676–703.
- Abu-Sharkh, S., Li, R., Markvart, T., Ross, N., Wilson, P., Yao, R., ... Arnold, R. (2005). *Microgrids: distributed on-site generation*. Tyndall Centre for Climate Change Research.
- Abushakra, B., Sreshthaputra, A., Haberl, J. S., & Claridge, D. E. (2001). Compilation of Diversity Factors and Schedules for Energy and Cooling Load Calculations, ASHRAE Research Project 1093-RP, Final Report.
- ASHRAE. (1989). *ASHRAE Standard 90.1-1989 Energy Efficient Design of New Buildings Except Low-Rise Residential Buildings*.
- Bahrami, A. (2018). Heteroptera. Retrieved from <https://www.food4rhino.com/app/heteroptera>
- Cerezo, C., Dogan, T., & Reinhart, C. (2014). Towards standardized building properties template files for early design energy model generation. In *Proceedings of the ASHRAE/IBPSA-USA building simulation conference* (pp. 25–32).
- Dong, B., Andrews, B., Lam, K. P., Höynck, M., Zhang, R., Chiou, Y.-S., & Benitez, D. (2010). An information technology enabled sustainability test-bed (ITEST) for occupancy detection through an environmental sensing network. *Energy and Buildings*, 42(7), 1038–1046.
- Hendron, R., & Engebrecht, C. (2010). *Building America house simulation protocols*. National Renewable Energy Laboratory Golden, CO.
- Iowa State University. (n.d.). ISU Sustainable cities. Retrieved from <https://sustainablecities.cber.iastate.edu/>
- Iowa State University Planning Team. (2014). *Capitol East Neighborhood Plan 2014*. Retrieved from [https://lib.dr.iastate.edu/resilientneighborhoods\\_plans/1/](https://lib.dr.iastate.edu/resilientneighborhoods_plans/1/)
- Janda, K. B. (2011). Buildings don't use energy: people do. *Architectural Science Review*, 54(1), 15–22.
- López-Rodríguez, M. A., Santiago, I., Trillo-Montero, D., Torriti, J., & Moreno-Munoz, A. (2013). Analysis and modeling of active occupancy of the residential sector in Spain: An indicator of residential electricity consumption. *Energy Policy*, 62, 742–751.
- Mahdavi, A. (2011). The human dimension of building performance simulation. In *Proceedings of Building Simulation* (pp. K16–K33).
- Mahdavi, A., Lambeva, L., Mohammadi, A., Kabir, E., & Pröglhöf, C. (2007). Two case studies on user interactions with buildings' environmental systems. *Bauphysik*, 29(1), 72–75.
- Mahdavi, A., & Tahmasebi, F. (2016). The deployment-dependence of occupancy-related models in building performance simulation. *Energy and Buildings*, 117, 313–320.
- Malekpour Koupaei, D., Hashemi, F., Tabard-Fortecoëf, V., & Passe, U. (2019). A Technique for Developing High-Resolution Residential Occupancy Schedules for Urban Energy Models. In *The Symposium on Simulation for Architecture and Urban Design*. Atlanta.
- Melfi, R., Rosenblum, B., Nordman, B., & Christensen, K. (2011). Measuring building occupancy using existing network infrastructure. In *Green Computing Conference and Workshops (IGCC)*, 2011 International (pp. 1–8). IEEE.
- Milne, M. (2016). Climate Consultant v6. 0.
- Paatero, J. V., & Lund, P. D. (2006). A model for generating household electricity load profiles. *International Journal of Energy Research*, 30(5), 273–290.
- Page, J., Robinson, D., Morel, N., & Scartezzini, J.-L. (2008). A generalised stochastic model for the simulation of occupant presence. *Energy and Buildings*, 40(2), 83–98.
- Reinhart, C. F., Dogan, T., Jakubiec, J. A., Rakha, T., & Sang, A. (2013). Umi-an urban simulation environment for building energy use, daylighting and walkability. In *13th Conference of International Building Performance Simulation Association, Chambery, France*.
- Richardson, I., Thomson, M., & Infield, D. (2008). A high-resolution domestic building occupancy model for energy demand simulations. *Energy and Buildings*, 40(8), 1560–1566.
- Therneau, T. M., Atkinson, B., & Ripley, M. B. (2010). The rpart package. Chicago.
- U.S. Bureau of Labor Statistics. (2017). *American Time Use Survey*. Retrieved from <https://www.bls.gov/tus/>
- U.S. Energy Information Administration. (2018). *Annual Energy Outlook 2018: with projections to 2050*.
- Yan, D., Hong, T., Dong, B., Mahdavi, A., D'Oca, S., Gaetani, I., & Feng, X. (2017). IEA EBC Annex 66: Definition and simulation of occupant behavior in buildings. *Energy and Buildings*, 156, 258–270.
- Yan, D., O'Brien, W., Hong, T., Feng, X., Gunay, H. B., Tahmasebi, F., & Mahdavi, A. (2015). Occupant behavior modeling for building performance simulation: Current state and future challenges. *Energy and Buildings*, 107, 264–278.
- Yang, J., Santamouris, M., & Lee, S. E. (2016). Review of occupancy sensing systems and occupancy modeling methodologies for the application in institutional buildings. *Energy and Buildings*, 121, 344–349.