EEB HUB
Energy Efficient Buildings Hub

REPORT

RUTGERS
THE STATE UNIVERSITY OF NEW JERSEY

Preliminary Report: Incorporating Information on Occupant Behavior into Building Energy Models

January 2014

Citation:
Figueroa, M., Putra, H.C., and Andrews, C.J. 2014. *Preliminary Report: Incorporating Information on Occupant Behavior into Building Energy Models*. Prepared by the Center for Green Building at Rutgers University for the Energy Efficient Buildings Hub, Philadelphia, PA.

Acknowledgment:
This material is based upon work supported by the Energy Efficient Buildings Hub (EEB Hub), an energy innovation hub sponsored by the U.S. Department of Energy under Award Number DE-EE0004261.

Table of Contents

Executive Summary

Accepted practice absolves building energy modelers of responsibility for capturing many of the effects of occupant behavior by assuming fixed comfort targets and ignoring "unregulated" loads. This paper asks what we can learn by incorporating more detailed information about occupant behavior into models. It compares results of three approaches: conventional practice, an augmentation incorporating detailed occupancy patterns, and an augmentation incorporating detailed behavioral responses of occupants to evolving comfort conditions. We apply these models to a highly-instrumented commercial building in Philadelphia, PA, USA, using EnergyPlus and extensions based in Markov chain modeling and agent-based modeling. We share preliminary findings only because the project schedule was disrupted. Key preliminary findings are that (1) better occupancy data greatly improves energy model accuracy, (2) standard assumptions about occupant schedules are often wrong so that a more sophisticated representation is warranted, (3) better data about occupants' adaptive responses only marginally improves energy model accuracy, (4) yet such data are quite valuable for predicting occupant satisfaction, and (5) incorporating occupancy data EnergyPlus needs additional hooks for incorporating occupant behavior.

1. Introduction

There is growing recognition that occupant behavior influences energy usage in buildings, but methods for incorporating its effects into energy modeling are not standardized. The purpose of this paper is to advance building energy modeling practice by explicitly comparing three models of the same building that incorporate occupant behavior in different ways.

Previous research has demonstrated that some occupant behaviors are more influential than others, so that temperature set point changes typically outweigh changes to internal loads [Blight and Coley 2013], especially in small buildings [Azar and Manassa 2012]. Some building and system types are more strongly affected by human behavior than others, so that low-mass buildings with large amounts of glazing are more influenced by external climate than by occupant behavior [Hoes et al 2009], buildings in cold climates served by district heating are not very sensitive [Kyrö et al 2011], whereas lighting and plug loads vary with occupancy schedule [Yun, Kim and Kim 2012].

Some occupant behavior is reasoned but much is habitual [Abreu, Pereira and Ferrão 2012; Masoso and Grobler 2010]. Occupants generally care more about thermal comfort than other indoor environmental quality factors [Frontczak and Wargocki 2011]. Occupants do not have homogeneous comfort preferences, and studies in a variety of countries have documented effects of personal attributes including age and gender, as well as social factors such as sense of control [Karjalainen 2013; Choi, Loftness and Aziz 2012; Indraganti and Rao 2010 (2); Santin 2011].

Occupants also rely on a variety of physiological, psychological, and behavioral adaptive responses to uncomfortable conditions [Yun et al 2012; Liu, Yao and McCloy 2012; Cao et al 2010; Indraganti 2010]. This variability in turn influences the efficacy of innovative building systems [Kalmár and Kalmár 2013; Saelens, Parys and Baetens 2011; Pfafferott and Herkel

2007]. In spite of the multi-faceted importance of occupants, a recent review concludes that few, if any, usable and validated models of the building-occupant system are available [Ryan and Sandquist 2012].

Previous research has employed a variety of methodologies for representing occupant behavior in building energy models. The standard approach adopted in deterministic building physics models such as EnergyPlus is to reduce the occupant to a set of assumptions about temperature set points, daily schedules, and internal loads [Fumo, Mago and Luck 2010], perhaps incorporating diversity factors [Davis and Nutter 2010]. Probabilistic modeling has proved useful for characterizing how frequently occupants make specific adaptive choices in response to changing environmental conditions [Haldi and Robinson 2008; Herkel, Knapp and Pfafferott 2008] and how their daily schedules vary [Duarte, van den Wymelenberg and Rieger 2013; Stoppel and Leite 2014; Tabak and de Vries 2010]. Monte Carlo analysis has identified which building parameters are sensitive to occupant behavior [Hopfe and Hensen 2011]. Neural networks have been able to predict user-sensitive appliance, lighting, and domestic hot water energy consumption [Swan, Ugursal and Beausolieil-Morrison 2011]. Markov process modeling has been effective for portraying more realistic occupancy schedules [Virote and Neves-Silva 2012; Page et al 2008; Richardson, Thomson and Infield 2008] lighting energy use [Widén, Nilsson and Wäckelgård 2009], and window operation [Yun, Tuohy and Steemers 2009], and provides a basis for quantifying remaining uncertainties [Meidani and Ghanem 2013]. Agent-based modeling has been used to simulate the diffusion of energy saving ideas through social networks within a building [Chen, Taylor and Wei 2012], predict user-controlled plug loads [Zhang, Siebers and Aickelin 2011] and it has allowed realistic representations of occupants' adaptive responses to changing comfort conditions [Andrews et al 2011].

The general trend is toward more differentiated analysis of occupants, activities, and schedules [Tanimoto, Hagishima and Sagara 2008]. Calibration and validation of models like those described in the previous paragraph requires substantial amounts of detailed data [Raftery, Keane and O'Donnell 2011]. Some studies rely on secondary data sets such as time-use surveys [Chiou, Carley, Davidson and Johnson 2011], but many collect primary data, as is done in the current paper.

The rest of this paper summarizes the data and methods used, shares and discusses results, and offers conclusions about the value added by each approach to modeling occupant behavior.

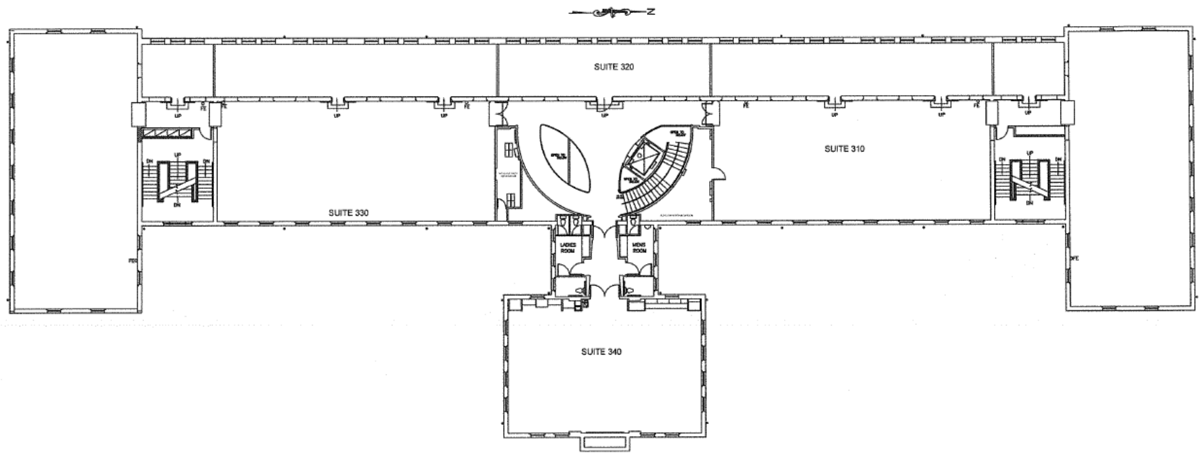
2. Calculation: Data and Methods

Data

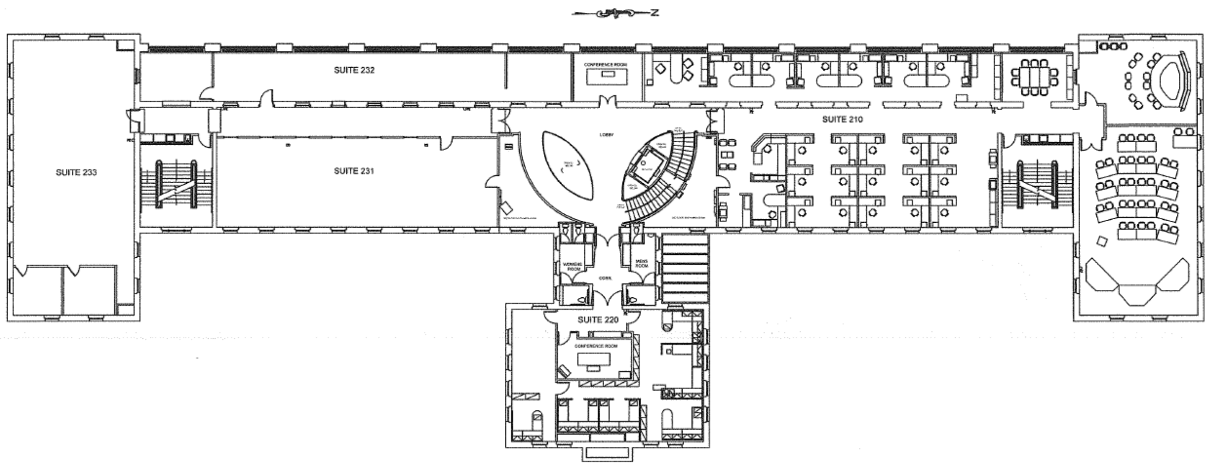
Building 101: Study site description

The study site is a 100+ year old brick structure that has been renovated over the years for changes in tenants. Currently owned by the Philadelphia Industrial Development Corporation (PIDC), it contains 75,156 ft^2 of gross building floor area with 61,000 ft^2 of multi-tenanted conditioned space (https://gpicspoint.ecs.psu.edu/gpic/Shared Documents/Hub Wide Meetings -- OC MEETINGS/BP3/2013.09.09-10 DOE External Review/04.4 - Building 101 Testbed R4 -

Rich.pptx *last accessed 1/31/14)*. The building is positioned on the north/south axis with each of 3 floors and basement level having two main wings (See Figure 1). As of August 2013 the building was reported to be 70% occupied.



3RD FLOOR



2ND FLOOR

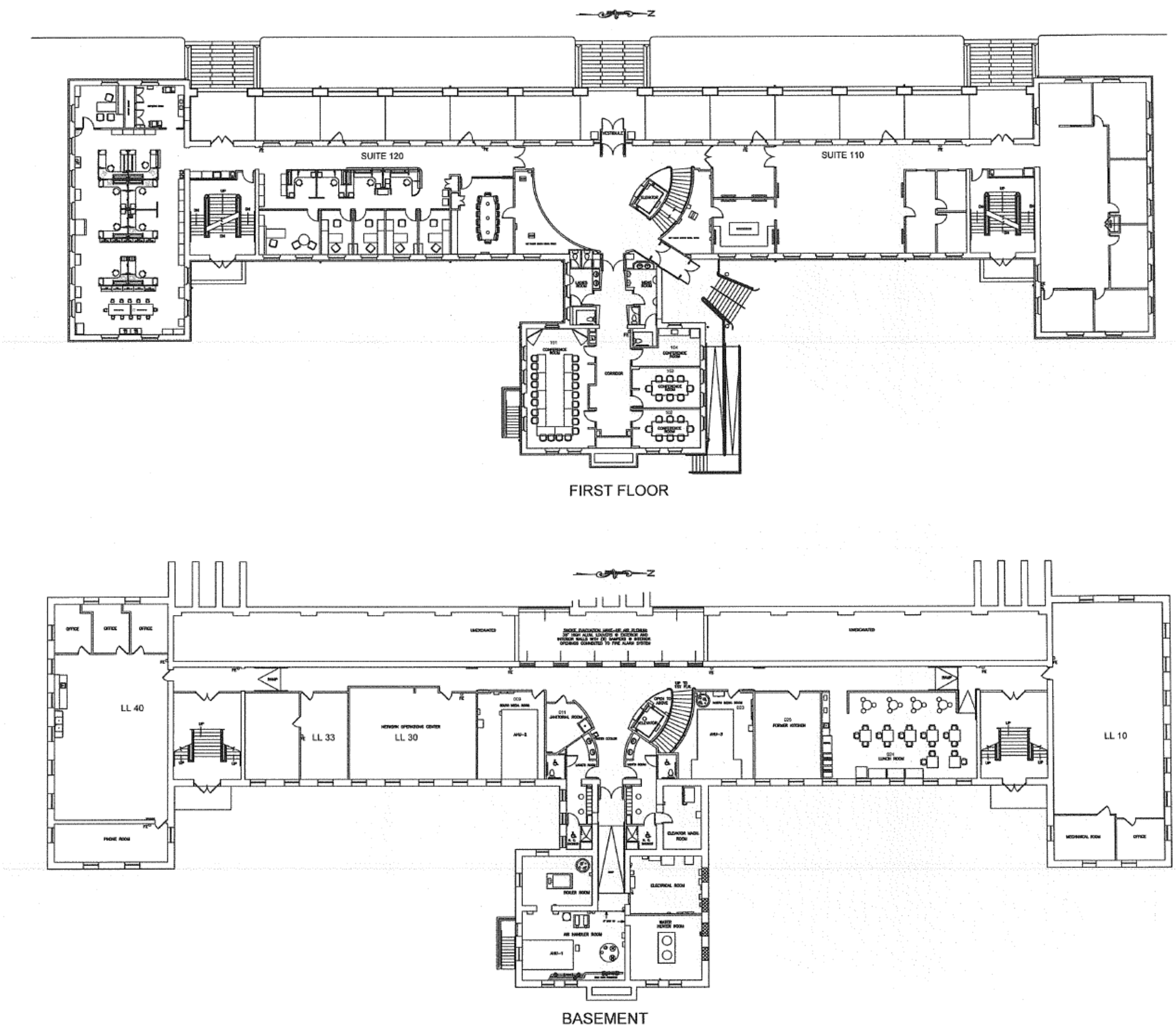


Figure 1: Building 101 Floor Plan

Building 101 has been the subject of much detailed engineering fieldwork by Hub researchers, hence that is not addressed further here. Refer to Xu (2012) for a full energy analysis of the building, including the development and calibration of the EnergyPlus model used as a baseline for this study.

Building 101 Information Portal Description

The CDH Energy website generates data for over 1500 physical measurements on Building 101 as CVS files (ASCII formatted text-files) which can be downloaded to local computers after system authentication. The measurements collect data for most of the systems operating in the building.

In particular, the building occupancy data, is displayed as integers representing the count of the number of people passing in each direction through 6 sensors for every 5-minute interval of the day, 7 days a week. The daily cumulative count for each sensor, which is reset every 24 hours at 11:00 PM, is also reported. Since for each sensor there are four measurements reported (entries, exits, cumulative entries, and cumulative exits) there is a total of 24 data streams available. The corresponding dates and times are available for each of these 24 data entries.

Four out of these six sensors are placed in such a way as to capture the entries and exists to the entire building (except for a small secondary door).

Data was extracted for the dates from February 1st 2012 up to June 11th 2013. The cumulative data showed to be incomplete, so it was decided to work with the raw counting data. Recalculating the cumulative data, and subtracting the exits from the entries (with a delay of one time step, namely 5 minutes) we can have a good approximation of the number of people present in the building for every 5-minute period in the data sample. A quick inspection reveals weekly cycles such as the one in Figure 2.

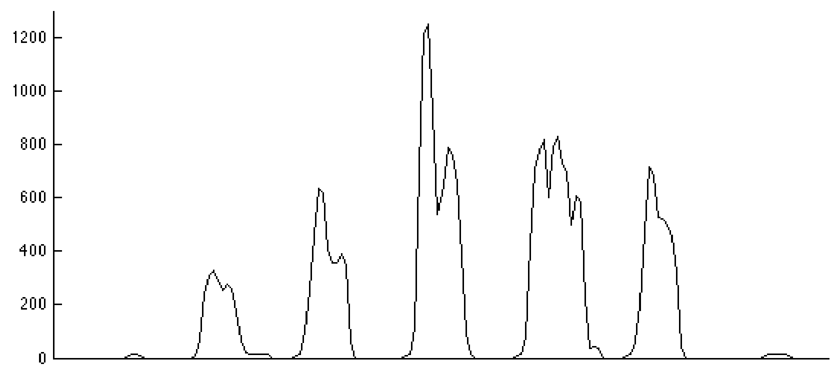


Figure 2: Sample Occupancy Data for a Week

In Figure 2 it is shown the weekly evolution of people in the building from Sunday to Saturday, peaking on Wednesday. This type of weekly evolution is typical throughout the data.

Building 101 occupant behavior fieldwork

Fieldwork included an assessment of occupant responses to energy efficient retrofits in tenanted spaces in Building 101. Our work thus far has illustrated that tenants of commercial office space can have an effect on building performance through the actions that they take to change conditions in their workspace that detract from thermal or lighting comfort or support for productivity. The “living lab” of Building 101 and its staff facilitated this investigation in an energy efficient, highly instrumented context that must compete with other commercial office building owners for tenant leases.

The study objectives were to better understand how occupants use their leased office spaces in ways that affect the projected energy performance of the building. The hypothesis is that integrated studies of leased office spaces will reveal readily available as well as longer term opportunities for greater energy savings with little to no additional investment that also improve tenant satisfaction.

Methods

IRB approval was secured for a multi-method approach to record independent observations and self-report by tenant representatives and occupants of leased office space at Building 101. An initial walk-through guided by facilities personnel incorporated initial observations and contacts with tenants as perspective participants in the study. Upon return, consenting tenants were engaged in in-depth interviews that developed information including the organization's mission, number of employees, their schedules in the office, and overall lighting and temperature fit with office needs. Short intercepts were conducted with tenant employees and included questions asking individuals to compare the typical lighting and temperature they have compared to their preferred levels. Photo documentation and spot measurements of temperature and light levels were taken in conjunction with interview comments. An online building-wide anonymous survey focusing on occupancy patterns and uses of space and equipment was also distributed through facilities management. Finally, targeted plug load metering was implemented in tenanted and some common spaces where their measurements could be used to compare appliance use and energy consumption to inform future strategies toward energy efficiency. At this point the integration of interview and observational data analysis is limited because of a change in funding support and so preliminary findings are reported herein.

Preliminary Findings

Eight of 11 tenants were engaged for this occupant behavior evaluation of their occupied offices. Eight tenant interviews and 20 intercept interviews were conducted. Interviews were analyzed for themes associated with temperature and lighting adaptations performed by occupants to improve their lighting and HVAC experience, and on the availability of information about saving energy.

Interviews and observations. Data was collected onsite between November 2013 and January 2014. Suites were serviced by forced hot air and for some also hydronic heat systems. Lighting fixtures across suites were primarily fluorescent pendants and sometimes included mini pendants. Tenants' were generally pleased with the building, location, views, and response by management personnel. Their experiences with temperature and lighting comfort, however, varied somewhat with the location and size of their suite and the source of heating, factors that contributed to the adaptive behaviors occupants took to improve their comfort in the workspace. Tables 1 and 2 summarize the observational and interview data that helped characterize occupant behaviors in response to work area temperature and lighting parameters.

Adaptive behavior to improve thermal comfort was influenced by location of workspace in the suite, size of office space, and type of heating among other factors. While many relied on using clothing layers, several tenants reported use of portable heaters in their suite, an important consumption of energy. Other suites were observed with tenant-applied obstructions to vent systems either as a direct effort to improve comfort or as secondary to fit out that is not coordinated with building design. Similarly, adaptive attempts to improve lighting conditions / control glare often resulted in blinds or other shading devices drawn and not being reset. Fit out also played a role in occupant-led lighting management when light fixtures, some of which were in need of maintenance, were misaligned and not located over work activity locations or partitions obstructed windows.

Table 1. Summary of Observations, Tenant Occupied Spaces

Topic	***Example***	***Design Note***
Adaptive actions in office: Temperature management	• Vents blocked with boxes or carpet tiles • Portable heaters in use • Blocking vents with materials	• Vents located close to seated occupant • Closed offices have cold exposed brick finish, cool lobby with extensive glass and older windows • No shut off to local heating cabinets
Adaptive actions in office: Lighting management	• Blinds down in the South, up in the North • Overhead lights dim • Lights on in one office without occupant • Task lighting not used • Sheets of paper hanging from window to block glare	• Blinds behind partitions are difficult to access • Diffusers need cleaning, bulbs out • No override 'off' switches • Task / cabinet lighting not placed over work area • Blinds cover only portions of windows
Exposure	Glare	Office affected can vary with season, time of day
Fit-Out	• Support for daylighting into workspace • Obstructions to daylighting • Multiple accessary task lighting • Obstructions to HVAC performance	• Panels on partition tops • Lower partitions facing windows • Tall partitions in front of windows • Blinds down to control glare & infrequently readjusted • Overhead lighting not placed over desk work activity area • Hydronic heat cabinets are used as work surfaces • File cabinets placed over floor vents • Materials placed on heat cabinet vents

Table 2. Summary of Tenant Reported Features

Topics	***Example***	***Design Note***
Workplace schedule	Most are full-time employees	Majority of occupants there every day
Adaptive actions in office: Temperature management	• Portable fans in summer • Portable heaters in winter • Extra clothing, jacket continuously, blankets, long underwear • Cool in summer &/or winter • Cardboard used on window seams • Vents blocked to reduce temp • Windows opened • Cold in early am but thermostat does not report accurately • Some tenants work with blinds and sunlight schedules • Contact with management was every day, now less frequent	• Possible inadequate / excess summer cooling and inadequate heating in winter depending on tenant location; vacancy may affect thermal comfort performance • Windows allow drafts on coldest days • Space is small & intended as open area, downsized overhead duct space with no returns for air circulation • Tenants do not perceive thermostats as being functional & some thermostats are controlled by other offices • Blinds are non-motorized • Management easily accessible / readily available
Adaptive actions in office: Lighting management	• Lighting discomfort, headaches so will sometimes work without lights • Lights stay on in unoccupied spaces • No use of task / cabinet lighting • Sheets of paper hanging in front of windows to stop glare	• Limitations with lighting switches (allows either all on or all off) or do not exist • Lighting schedules can vary by office needs • Placement of lighting not coordinated with desk work area • Blinds cover only certain parts of windows
Adaptive actions in office: use of appliance	Energy intensive appliances	• Refrigerator water filter feature not present or inoperable • Infrequent use of common space energy-intensive kitchen appliances
Exposure	• Glare management • View	• Shades often not readjusted when down • Views overlooking parade grounds are important
Fit-Out	Layout coordination with building design & objectives	• Planned design of coordinated building / fit-out is often missing • Offices have grown with different fit-out needs
Availability of Energy Saving Policy / Information	• Communications about energy saving behavior • When questions / problems are logged they are addressed quickly • Workshops and sessions	• None specific to energy savings from either building or employer • Communications with facilities appear fluid • Periodic workshops are not well advertised

Plug load metering. Substantial energy consumption was also detected on some common appliances where energy performance might be improved. A total of 33 meters were installed, primarily on pantry or kitchen appliances. One meter was used to measure a task light but removed because of negligible kWh readings. Because of the contracted period for completing the study, plug load meter installations provided for approximately 2 weeks' worth of data collection. For one appliance (industrial refrigeration case shown in Figure 3) a meter was installed for the day of the field site visit. The data retrieved and reported here is preliminary but offers some information that suggests potential areas of additional and readily achievable opportunities for increased energy savings.

Figure 3: Underutilized refrigerator

Methods

EnergyPlus building physics model & calibration process

EnergyPlus is a US DOE supported energy analysis and thermal load simulation program (see http://apps1.eere.energy.gov/buildings/energyplus/energyplus_about.cfm). The program is capable of calculating and integrating details of heating and cooling loads, conditions from HVAC and coil loads, and energy consumption of primary plant equipment in text format. EnergyPlus is a stand-alone simulation program that is available in a number of graphical user interfaces.

Xu (2012) provides a detailed, step-by-step account of the data collection, model preparation, and calibration process used to create the EnergyPlus model of Building 101, summarized in Figure 4.

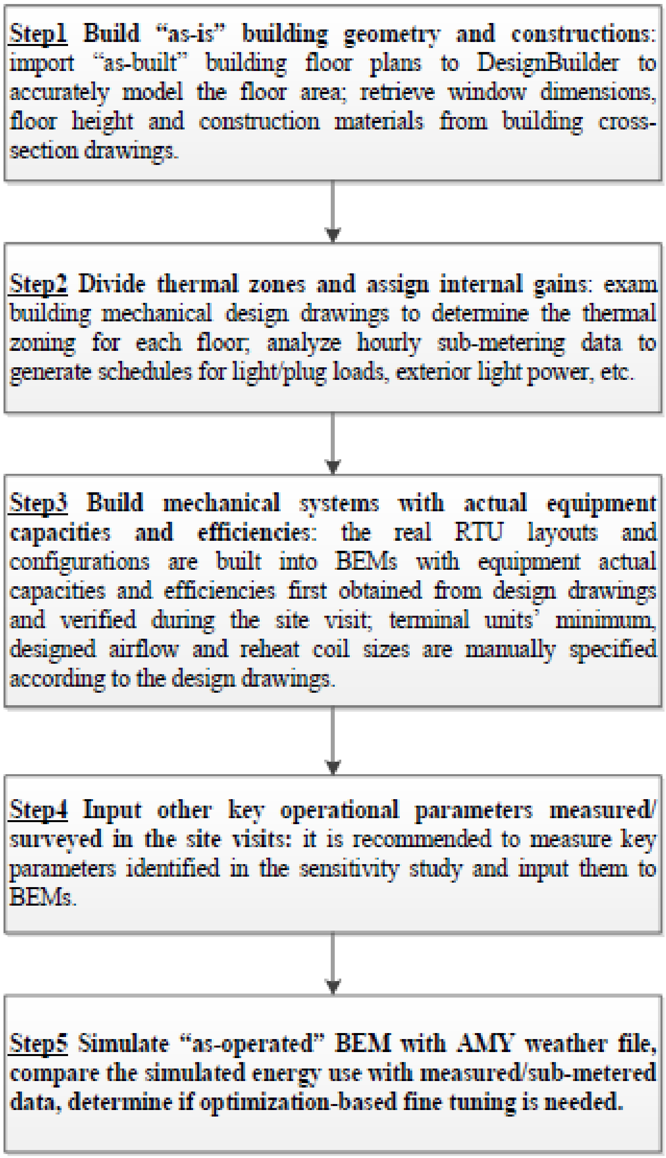


Figure 4: Procedure to build an "as-operated" building energy model
Source: Xu (2012, pg., 81)

Markov Chain Models

We will cover some of the basic concepts of Markov Models theory along with some current efforts found in the literature to make use of Markov Models to model building occupancy behavior and building energy usage estimation.

Review of Markov Models

Markov Models are a family of models that rely in the assumption of the verification of the Markov Property. One way to understand this property is to think of a stochastic system evolving in continuous or discrete time stages, jumping at each time epoch from one system state to another with some probabilities. The set of system states has to be finite (or at least countable). Then, given a present state of the system, the probability of jumping to another state in a non-Markov system is in general dependent on all the previous state-history, while in a Markovian system this probability depends only on the present state. This property can be then stated as:

"All the information for estimating the system's future behavior is available (stored) at the present time."

Or in an equivalent form:

"The system's future behavior is only dependent on the present time, and it is independent of the past."

In a more rigorous way, if $S = \{Set\ of\ states\ of\ the\ system\}$, and $X(n) \in S$ is the system's state at discrete time n, then

$$Prob(X(n+1) = j \mid X(n) = i, X(n-1) = k, \dots) = Prob(X(n+1) = j \mid X(n) = i) = p_{ij}(n+1)$$

An example could be if in an office the probability of turning off the lights at a certain hour depends only on the light condition at the previous hour, and not on the whole day light-history.

Considering a finite set of states, $S = \{0, 1, 2, \dots, N\}$, and a Markovian system evolving on discrete time (this is called a Markov Chain, or Markov Process). The set of *transition probabilities*

$$p_{ij}(n) = Prob(X(n) = j \mid X(n-1) = i), \qquad i, j \in S,$$

is usually arranged in matrix form in the *transition matrix*

$$P = \begin{pmatrix} p_{00} & \cdots & p_{0N} \\ \vdots & \ddots & \vdots \\ p_{N0} & \cdots & p_{NN} \end{pmatrix}.$$

These probabilities, assuming that the true values are available, are sufficient to describe all the dynamics of the system and its stochastic evolution.

The transition matrix gives at row i the probability distribution of the next state if the system is currently in state $X(n) = i$, since this is a distribution, it is required that the row-sums of the P matrix are one (this is called a *stochastic matrix*). The individual probabilities p_{ij}, however, can be zero if certain state-transitions are not allowed, or 1 if there is only one transition permitted.

For example, if we have three states $S = \{1, 2, 3\}$, and the transition matrix is the following:

$$P = \begin{pmatrix} 0{,}1 & 0{,}5 & 0{,}4 \\ 0{,}0 & 0{,}0 & 1{,}0 \\ 0{,}3 & 0{,}1 & 0{,}6 \end{pmatrix}$$

Then, looking at the third row, the probability that the next state will be state 1 given that we are in state 3 is $p_{31} = 0{,}3 = 30\%$, the probability to jump to state 2 from state 3 is $p_{32} = 0{,}1 = 10\%$, and the probability of remaining in state 3 is $p_{33} = 0{,}6 = 60\%$. Also, the probability of leaving state 3 is $p_{31} + p_{32} = 0{,}3 + 0{,}1 = 0{,}4 = 40\%$. From the second row we see that if the system is in state 2, then it cannot go to state 1 in the next step, nor can it stay in state 2, the only possibility is to transition to state 3 with 100% probability (deterministically).

Calculating powers of these matrices, one can obtain the transition probabilities for larger numbers of steps, for example $P^2 = P \cdot P$ holds the probabilities for two-step transitions between any two states.

Under certain technical conditions, such as finite number of states and aperiodic systems, the larger powers of this matrices start to converge to a matrix which holds, at every row (all rows are the same), the stationary distribution $\pi = (\pi_1, \pi_2, \dots, \pi_N)$:

$$\lim_{k\to\infty} P^k = \begin{bmatrix} \pi \\ \vdots \\ \pi \end{bmatrix}.$$

The π_i values are interpreted as the long-run average probability of being in state i, or the overall expected long-run proportion of time spent in state i. For the previous example we have

$$\lim_{k\to\infty} P^k = \begin{pmatrix} 5/24 & 4/24 & 15/24 \\ 5/24 & 4/24 & 15/24 \\ 5/24 & 4/24 & 15/24 \end{pmatrix},$$

then, $\pi = \left(\frac{5}{24}, \frac{4}{24}, \frac{15}{24}\right) = (0{,}2084, 0{,}1667, 0{,}6249)$, which are the overall fractions of time spent in each state in the long run.

For an application problem, the transition matrix can be calculated by defining the states of the system, keeping track of the transitions and calculating the proportion of times that any transition happens from a given state. Of course, the longer the longer the period of time that this process is observed, the more precise the matrix will be.

Once that the transition matrix for a system is known, we can simulate the stochastic process using Monte Carlo sampling. For this we calculate the cumulative distribution of each row distribution in the P matrix (cumulative row sum), and add a column of zeros at the beginning to form the P_{cum} matrix:

$$P = \begin{pmatrix} 0{,}1 & 0{,}5 & 0{,}4 \\ 0{,}0 & 0{,}0 & 1{,}0 \\ 0{,}3 & 0{,}1 & 0{,}6 \end{pmatrix} \Rightarrow P_{cum} = \begin{pmatrix} 0{,}0 & 0{,}1 & 0{,}6 & 1{,}0 \\ 0{,}0 & 0{,}0 & 0{,}0 & 1{,}0 \\ 0{,}0 & 0{,}3 & 0{,}4 & 1{,}0 \end{pmatrix}$$

Each row in the P_{cum} matrix then defines a set of intervals each one corresponding to a potential state (in the first row of the example the intervals are $[0\ 0{,}1]$, $[0{,}1\ 0{,}6]$, $[0{,}6\ 1{,}0]$). Then, drawing a random number uniformly in the interval $[0\ 1]$, and selecting the interval in which the random number fits, we effectively select the next state of the system considering the transition distribution of probabilities. If the number falls in a bordering probability shared by two intervals the tie is broken arbitrarily. If, as in the example second row, some intervals are of the type $[0\ 0]$ or $[p\ p]$, then it is assumed that the drawn sample cannot lie in that interval and the next interval is checked.

Continuing with the example, consider that the system is in state 3 at the current state and we want to determine randomly the next state but considering the distribution corresponding to the third state in the third row of P (in which remaining in the state 3 is 6 times more likely than transitioning to state 2, and 2 times more likely than transitioning to state 1). Then, we draw a random number uniformly in the interval $[0\ 1]$, $p = 0{,}5469$, which falls in the third interval $[0{,}4\ 1{,}0]$, and thus the next state for this instance is the third state. Repeating this experiment many times, reveals that the selection of the next state follows indeed the distribution of the third row, that is that state 3 is selected 60% of the times, state 2 30% of the times, and state 1 only %10 percent of the times. Many software packages (such as Excel, R, Stata, Matlab, etc) have random generators of uniformly distributed numbers.

The utility of this type of simulation depends on the granularity of the number of states. Defining two states could give too little information on the evolution if the system but at the same time it could be calibrated with little data. On the other hand, defining ten states requires a long period of simulation/calibration and data to capture all the variations but could yield a more flexible and meaningful model.

There is a lot more information that can be obtained from the P matrix, for example the expected number of time steps before reaching certain specific set (or a set of states), or the probability of ever returning to a specific state. Calculating this measure, however, requires some knowledge on Probability Theory and Linear Algebra. See the book of Kulkarni (2009) [1] for a reference in this subject.

From the previous discussions one can obtain some insights in the fact that models based in Markov Chains are complete enough to convey at the same time average information of the system by means of the long-term distribution calculation and also reproduce the transient stochastic evolution of the system.

In a more complex family of stochastic processes we count the Markov Decision Processes. In this type of processes not only the next state is random, but also the state distribution is randomized. This is similar to having a number of, in general different, transition matrices and using each one at some time step depending on some “distribution of distributions”. This type of modeling is a standard approach in the field of Stochastic Dynamic Programming and makes use of the Markov Property that many systems have to efficiently solve large-scale optimization problems even without the need to enumerate all states of the system.

Markov and Probabilistic Models in Occupant Scheduling and Energy Performance

Wilke et al. (2013), France [2], are recently working on developing residential building occupancy profiles for time-dependent activities, as parts of the inputs needed for other dynamic models of energy performance in residential buildings. They count with very precise and specific time-use survey data on occupant behaviors in France for the 1998-1999 year period.

The data is based on three time-dependent measures:

i. Probability of being in the residency
ii. Conditional probabilities of engaging on an activity while present in the residency
iii. Probability distributions of activity duration

Processing this information they produce a full occupant behavior probability schedule as shown in Figure 5.

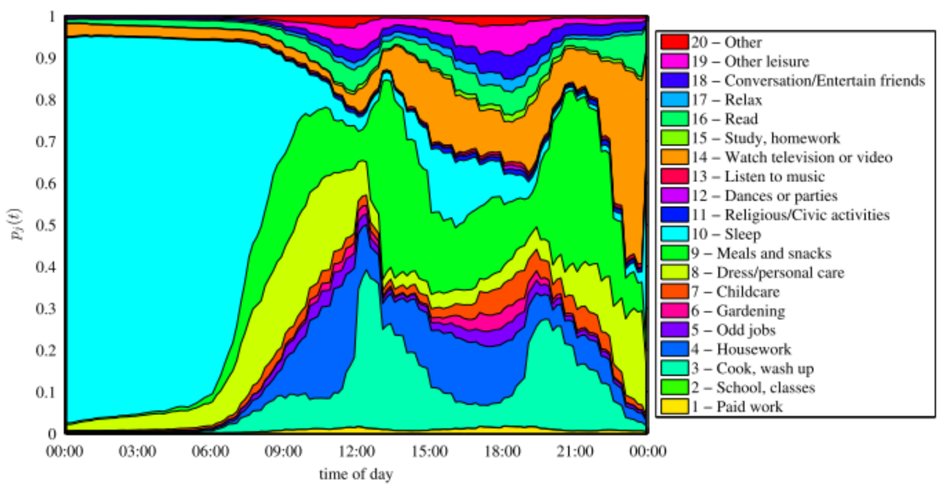


Figure 5: hourly activity proportion schedule, Wilke et al [2]

The data is then analyzed using Bayesian theory, studying the activities duration distributions and normalized to obtain hourly probability distributions or occupancy schedules as shown in Figure 6.

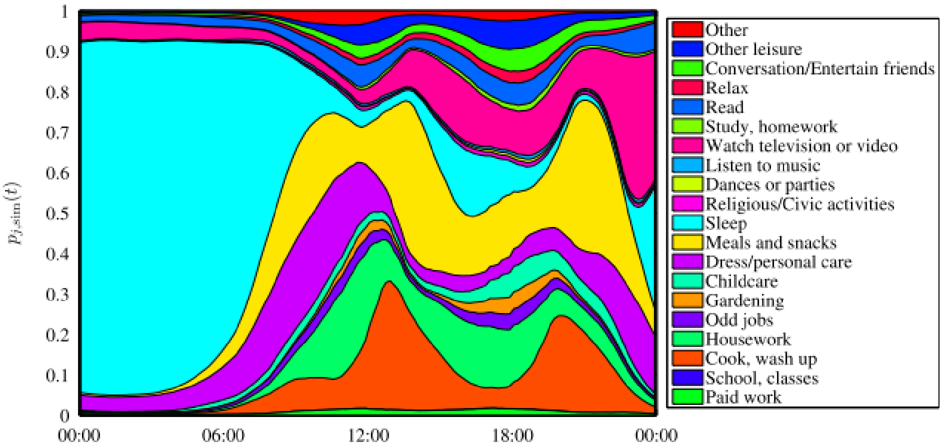


Figure 6: hourly activity probability schedule, Wilke et al [2]

The individuals are assumed at first to be of only one kind, but then the model is refined to be able to incorporate variations in activity behaviors or "sub-populations". The transitions between to activities are modeled by means of Markov processes. Wilke et al, present also calibration and validation techniques with which they refine and test the predictive capabilities of their schedule processing models.

Dodier et al. (2006) [3], make use of Belief Networks, which are a class of graphical probability models to answer, from occupancy data, simple and complex queries such as: number of people expected in the building in one day, sensor malfunctioning, conditional information as presence in an office given that a specific sensor is malfunctioning, average conditional information such as typical readings of over-sensitive sensors.

Given some events A, B, C, D, E (which can represent: presence in an office, malfunctioning of a sensor, etc.) they place them in a graph according to their dependencies (see Figure 7).

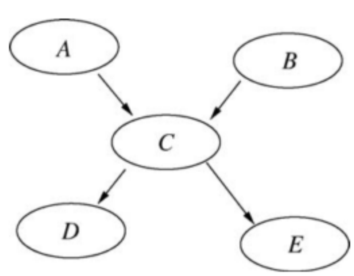


Figure 7: Belief network, Dodier et al [3]

Then, the probabilities of occurrence of each one of these (maybe conditional) events can be calculated directly from the data using counts and conditional probability definitions:

$$p(A,D|B) = \frac{p(A,B,D)}{p(B)} = \frac{\sum_C \sum_E p(A,B,C,D,E)}{\sum_A \sum_C \sum_D \sum_E p(A,B,C,D,E)}$$

This approach has the weakness that it requires large quantities of data to construct a model sensitive enough to distinguish the dependencies on all the defined events.

In a different two-level probabilistic approach, Page et al. (2008) [4] develop an algorithm for the simulation of occupant presence as an input for occupant behavior models.

The model consists of an inhomogeneous Markov chain (probabilities changing in time) for modeling the occupancy interrupted by occasional periods of long absence/presence. The model is capable of accurately reproducing occupant arrival times, departure times, periods of intermediate absence/presence and periods of long absence/presence.

The model works as a Markov process in which each transition is dictated by probabilities that decay or increase through time according to counters that are fed into calibrated probability density functions. These distributions are mostly exponential, which is a classical approach to model expected duration times. The data for the calibration is obtained from movement sensors.

Virote et al. (2012) [5], make direct use of Markov chains to model the agent transition between states in an agent-based model. They define the system's states according to the different user-usage combinations (which tend to be huge) to make direct energy consumption estimates. They describe the chain by means of its transition matrix and make predictions based on the transition probabilities. As mentioned before, these predictions tend to "average out" if the prediction horizon is large enough.

This type of approach has promising results in the sense that the model effectively learns the occupant behavioral patterns from the building and it reliably reproduces them to give accurate predictions of the building energy consumption. It is also possible in this model to identify potential areas of energy waste by studying the average time spent in high-energy usage states.

Lastly, Meidani et al. (2013) [6] make a type of generalization of the pure Markov Chain model motivated by the assertion that some of the uncertainties in these evolving systems cannot be captured in the calculation of the transition rates/probabilities from finite samples (such as data). They account for such variations by considering randomized transition matrices (which yield a type of models called hidden Markov chains). In these models, the transition matrix to be used at each step is not fixed. This type of formalism can capture fluctuations in the environment in which the chain evolves, such as weather variations or the presence of rare events.

Agent-based model

One of the main themes in this paper is to address a need to reflect the response of occupants for better building simulation [Andrews et al 2011; Andrews et al 2012]. After the fieldwork experiments, the lighting system, the HVAC system, and plug load information of the building was modeled and calibrated in EnergyPlus 8.0.0 [USDOE 2012; Ke Xu 2012]. EnergyPlus is widely used Building Information modeling (BIM) that helps designers to understand the physical characteristics of the building. It, however, still lacks of the behavioral element. Recent research tries to integrate the user behavior elements into BIM [Shen 2012; Andrews, Senick, Wener 2012].

This paper modeled occupants' thermal comfort actions (adjusting thermostat set points, turning on/off space heater, opening/closing the windows and door, and changing winter/summer clothes) and their influence on airflow rate entering their thermal zone by using set points and infiltration schedules. Occupants' lighting comfort actions (turning on/off headlights, turning on/off task lights, opening/closing windows blinds) were modeled using equipment schedule.

In modeling the occupant behavior that update the schedule, this study adopts two paradigms to specify theories and processes of human behavior. Agent-Based modeling (ABM) provides a paradigm of simple entities, called by agents that respond respectively to the environment. ABM is widely used in the ecological domain, but not very straightforward in representing human-like behavior [Epstein 2006; Axelrod 1997]. Belief, Desire, Intention (BDI) is a paradigm of agents that are based on a psychological view of how people behave. BDI characterizes the process of human decision-making, such that autonomous agents follow five procedural steps in making behavioral decisions: establishing beliefs, desires, and intentions, developing plans, and deciding to carry out a particular plan of action[Rao and Georgeff, 1998]. NetLogo [Wilensky and Rand 2013] is used to develop an integrated model of the two paradigms. Calibration is done using survey and interview data from individual building occupants, plus building-wide performance data for building 101. The model is validated by using it to predict outcomes (expressed as usability metrics) for an additional building.

The complete modeling logic is summarized in Figure 8 below. It contains a building performance submodel that updates the state of the indoor environment over time. It contains a human agent submodel that simulates individual and shared decisions of occupants as they experience and react to changing environmental conditions. It also includes a file populated with information about the current state of controllable and uncontrollable building features. This modeling framework was introduced in Andrews et al [2011] and extended in Andrews et al [2012].

A building performance submodel has inputs such as building site conditions and design choices. Inputs for human agent model include occupancy schedule, occupant preferences and capabilities. Outputs include the usability measures of effectiveness, efficiency, and satisfaction [Andrews et al 2011].

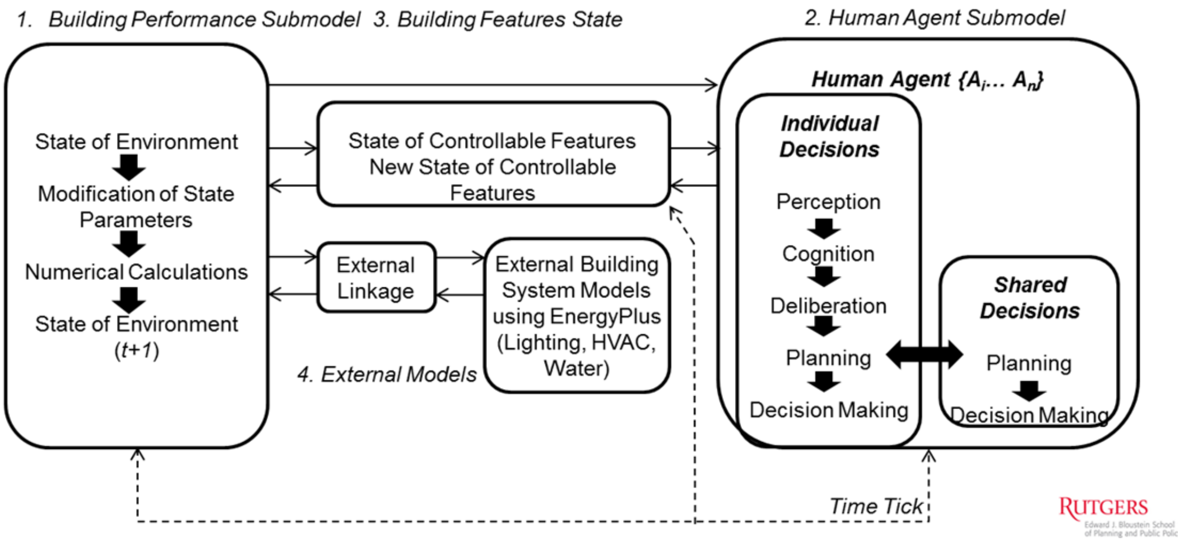


Fig 8: Modeling system logic
Source: Andrews et al 2012

EnergyPlus hot-linked to agent-based model

The Building 101 simulation study consists of three main components: the building energy model, the occupant behavior model, and the integrating model. Each is programmed with a different software application. The building energy model uses the EnergyPlus modeling engine that characterizes the energy performance of the building design (http://apps1.eere.energy.gov/buildings/energyplus/). The occupant behavior model is programmed in the NetLogo agent-based-modeling environment (www.netlogo.org). The integrated model is a model that allows a communication between the occupant behavior model and the building energy model by using Java programming language (http://www.oracle.com/technetwork/java/index.html).

The building energy model, using EnergyPlus, incorporates occupant behavior component within it at a very limited level. The picture of having the building physics and the building occupants to perform an active-reactive interaction drives the overall goal of this simulation study. The building energy model does not allow users to modify the input variables, located in the .idf file, on the fly while running the simulation (Figure 9). In other words, users modify the .idf file prior each overall simulation run. It also does not receive values exogenously for all the input variables. The integrated model runs in two-step for each simulation-hour. The model calls the building energy sub-model and the occupant behavior sub-model alternately. Initially, the integrated model runs the building energy sub-model in order to create the building environment. The model, then, runs the occupant behavior model in order to simulate building occupants' sensation and adaptive behavior towards the surrounding building environment. The occupant behavior model will consider the building environment conditions, resulted from the building energy model run at the previous time period, and the occupants' physiological preference towards the environment (Figure 10). For example, the task requires to run a 24-hour simulation period, the integrated model will call both the occupant behavior model and the building energy model for 24 times. As a result, the overall running time for the building energy model is relatively faster than the integrated model. However, running the building energy model does not reflect the reality of building occupants' adaptive behaviors. Modeling the adaptive behavior requires the occupants to evaluate and make changes of the environment in every cycle of the overall simulation period.

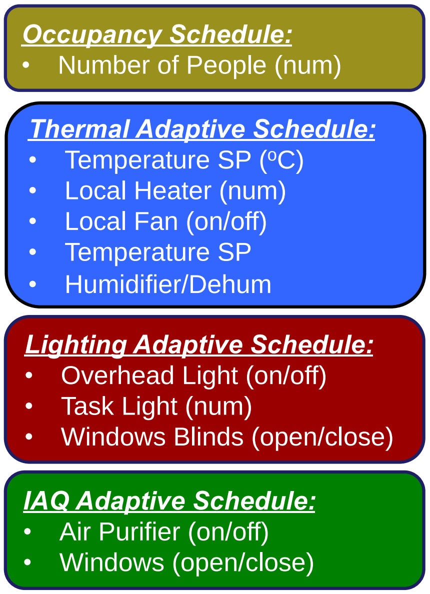


Figure 9. Building energy model parameters

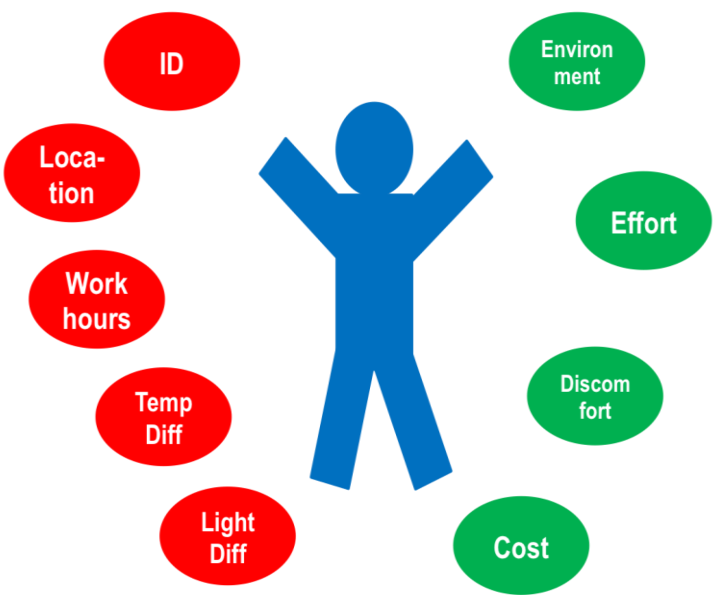


Figure 10. Occupant behavior parameters

How the model simulates comfort and satisfaction.

The integrated model has not yet completed to perform calibration for both comfort and satisfaction simulations on Building 101. The model, however, successfully follows the logic of occupants' comfort and satisfaction (Andrews, Chandra Putra, and Brennan, 2013). In the thermal comfort scenarios, occupants perceive the environment as Too Hot, Thermally Neutral, or Too Cold. The set of adaptive behaviors occupants perform range from Do Nothing, Adjust the Thermostat, Turn On/Off a Personal Fan, Turn On/Off a Personal Space Heater, and Add/Remove Clothing. In experiments simulating illumination levels, occupants perceive Too Bright, Illumination-Neutral, or Too Dim. Occupants can respond such sensations with the following adaptive behaviors: Do Nothing, Adjust Window Blinds, Turn Task Light On/Off, and Turn Overlight On/Off (Table 3, 4)

Perceptions/ Adaptive Behavior	Do Nothing	Thermo stat	Space heater	Blinds	Win-dows	Fan s	Change Clothes	Overhead lighting	Task light	Cookin g app.	Dishw asher	Shower	Faucet	Toilet
Too hot/Too cold	X	X	X				X							
Too bright/Too dark	X			X				X	X					
Too dirty -IAQ-	X				X	X	X							
Dish stock >	X										X			
P body care habits	X						X					X	X	
P sustenance habits	X									X			X	
P eliminate habits	X											X		X

Table 3. Perceptions and adaptive behaviors

	Environmental Impact	Effort	Discomfort	Cost
Do Nothing	Lo	Lo	Hi	Lo
Thermostat	Hi	Lo	Lo	Hi
Space heater	Hi	Lo	Md	Hi
Blinds	Lo	Hi	Md	Lo
Windows	Lo	Hi	Md	Lo
Ceiling fans	Lo	Hi	Md	Md
Overhead lighting	Hi	Lo	Lo	Hi
Task light	Md	Md	Lo	Md

Table 4. Disutility values

3. Results

Fieldwork

The observational and interview data are summarized in Tables 1 and 2. An additional study of unregulated plug loads in common areas yielded interesting results. Figure 11 shows average daily electricity use per tenant by appliance. Figure 12 shows four major appliance types and the variation in their energy use across instances. Interestingly, refrigerator electricity use varied widely across tenants. This could be due to different sizes, vintages, and frequencies of opening.

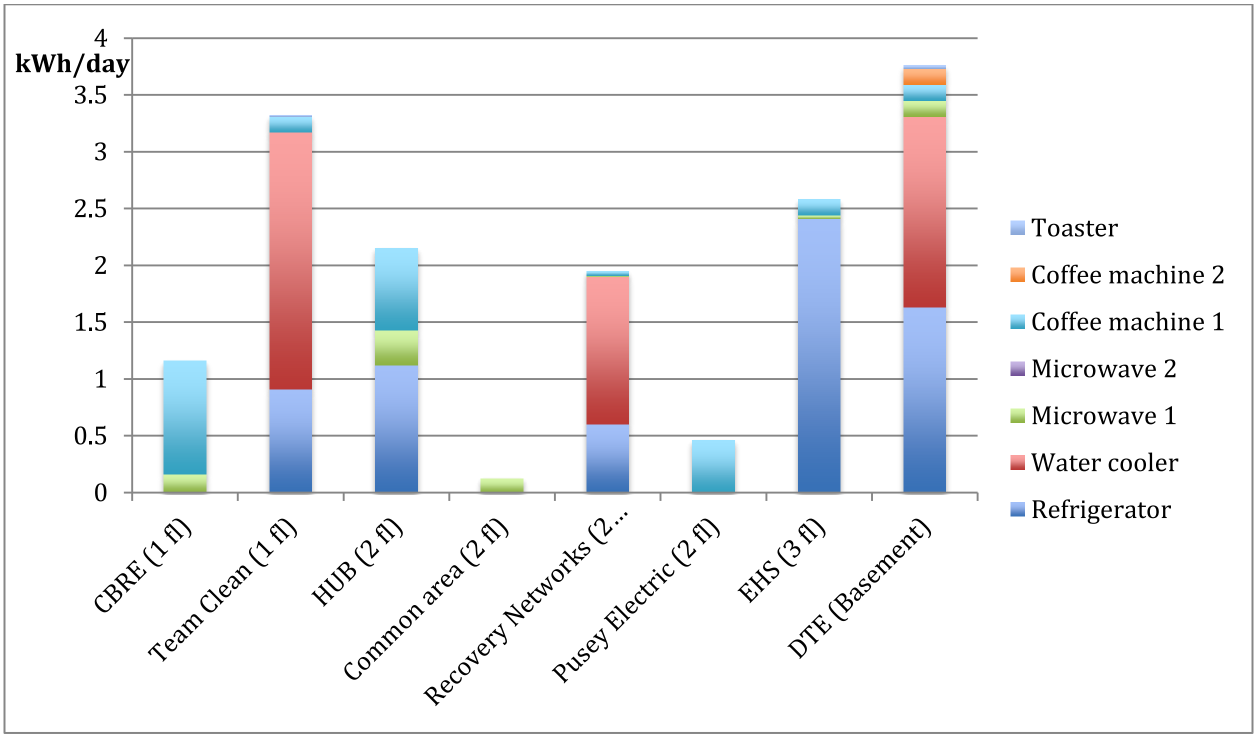


Figure 11: Appliance electricity consumption by tenant in Building 101

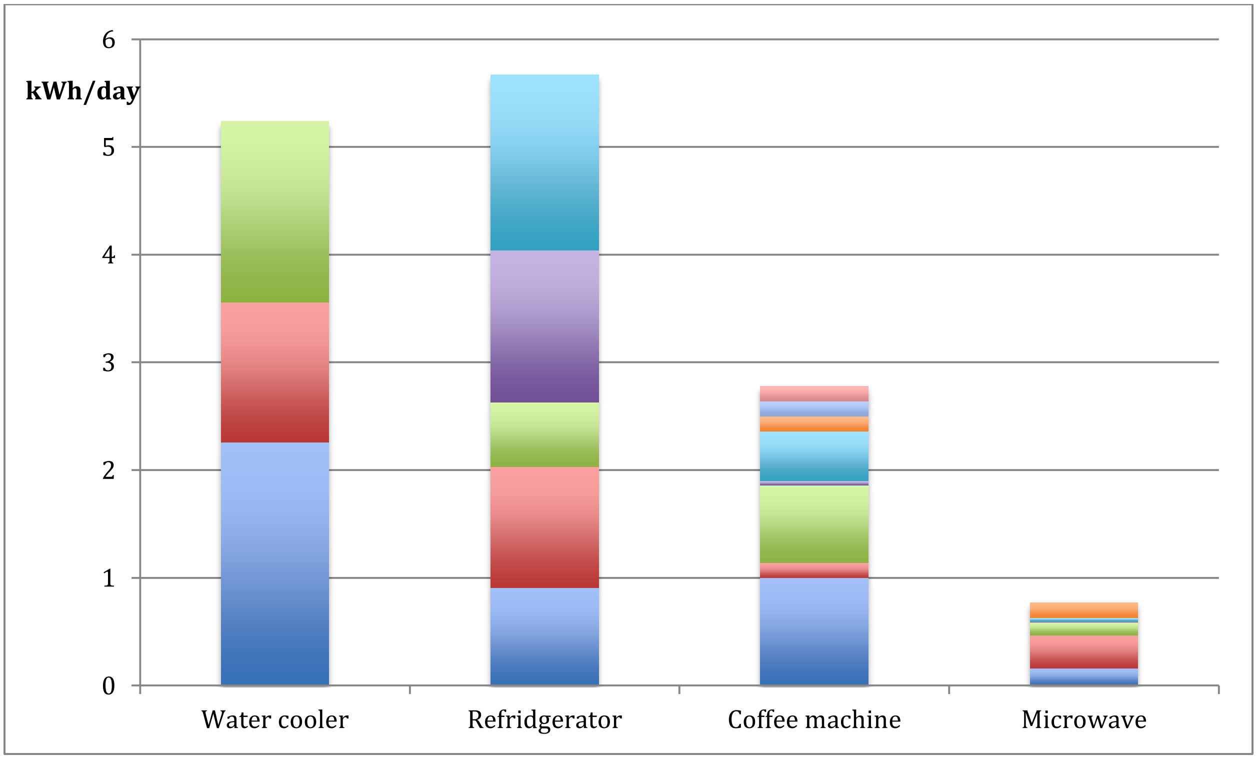
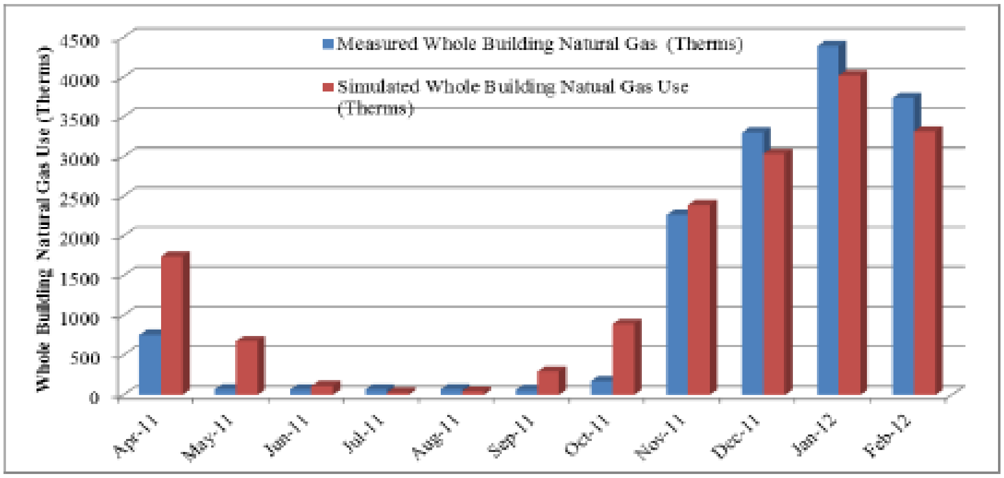


Figure 12: Variation in appliance electricity consumption across instances in Building 101

Conventional EnergyPlus model

Figure 13 shows high-level results of the model calibration process carried out by Xu (2012). Relative to actual utility bills, the model substantially over-estimates natural gas use in the swing months and under-estimates it slightly during the winter months. The model substantially under-estimates electricity consumption during the spring and summer months but more closely approximates it during the winter months. Xu (2012) provides much additional detail on the performance of this model, and concludes that, although this is an envelope-dominated (rather than internal load dominated) structure, unverified equipment efficiencies, usage schedules, and plug loads are among the major contributors to discrepancies between modeled and measured energy performance. This begs the question of whether additional detail on occupant schedules and occupant behavior might be helpful.



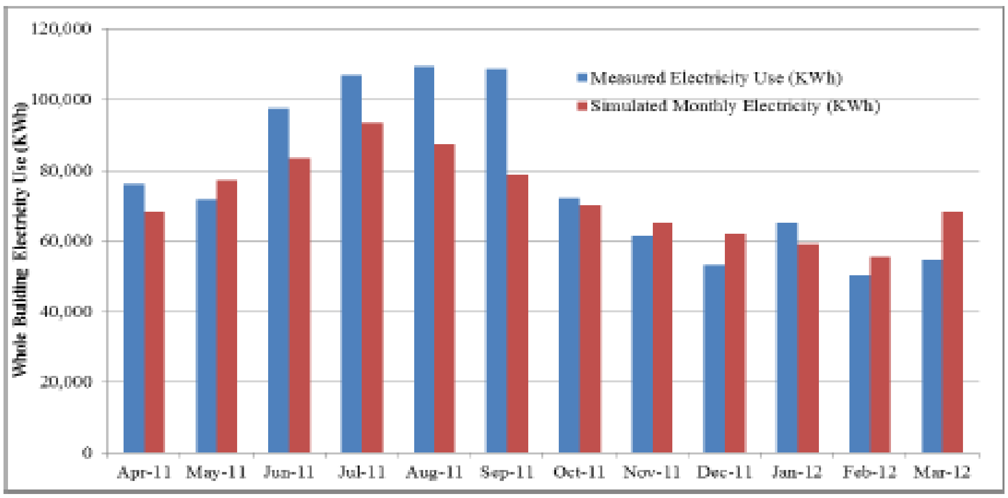


Figure 13: Monthly energy use comparison
Source: Xu (2012, pg. 94)

Occupancy Schedules

Some important inputs of Energy Plus models, on which many of the energy calculations depend, are the people schedules. These schedules define the number of occupants to be expected in the different zones of a building at each time for a full year of simulation. A common approach is to define separate schedules for zones that have very different occupation levels (e.g. offices versus hallways). The schedules are defined in a bottom-up fashion by defining daily schedules that are aggregated into weekly schedules and so on until completing a full year. A daily schedule can be defined by enumerating the proportion or fraction of occupancy out of a maximum occupancy level (which is set separately, as another input) expected at each hour of the day.

It is common to use simple daily schedules in the form of step functions, which would correspond, for example, to schedules like the following: between 0:00 AM and 8:00 AM, expect 10% of occupancy (0,1 times the maximum occupancy level set), between 8:00 AM and 5:00 PM expect 80% of occupancy, and expect 10% for the remaining hours. These types of schedules have the advantage of being very simple and easy to feed into the models, but could be oversimplified.

We propose to enrich these schedules with the people counter data in order to try to better predict energy usage for building 101.

Average Schedules

A first approach would be to use an *average schedule*, constructed by taking the arithmetic average of the weekly people occupancy data (the average of the schedules as the one showed in Figure 2). Figure 14 shows in a black line this average schedule for the retrieved data together with its deviation in red lines. The average occupancy and deviation are calculated for each hour using all the weeks available and the deviation is then added and subtracted.

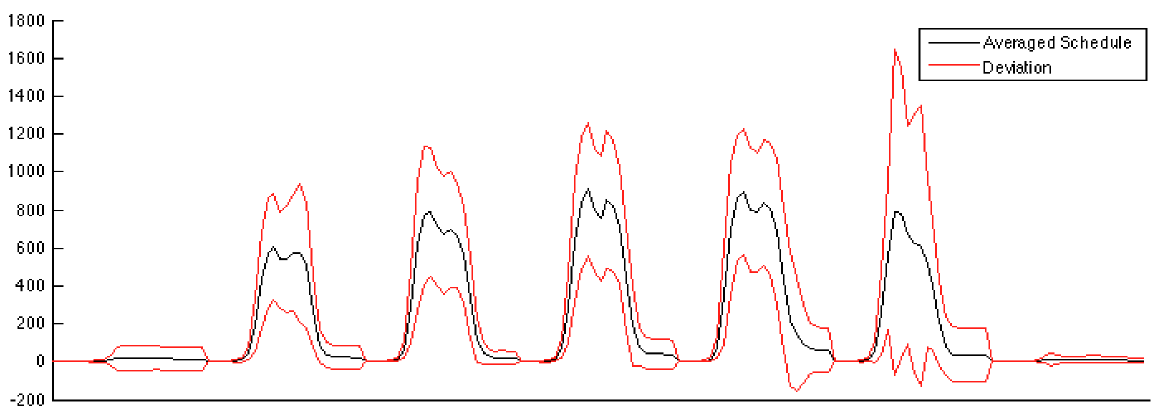


Figure 14: Average Schedule and Deviation

It is clear that, except for hours in the night and maybe weekends, the average schedule is over simplistic in the sense that it does not capture the weekly (or seasonal) variations which cause high and low peaks in the occupancy levels. These peaks could be affecting the comfort/discomfort levels of the occupants, or could render the equipment sizing inadequate.

We will compare later how this average schedule fares on an energy model of Building 101, but first we will get to the task of generating richer schedules.

Parameterized Model

We want to find a way to parameterize all the typical daily schedules that are found throughout the year in order to reproduce then realistic schedules. A preliminary analysis on the data shows that: (1) there is no distinctive "summer" behavior in the sense that the occupancy of the building does not drop significantly during the summer months (2) the "day of the week" effect is very important, meaning that the fact that the occupancy in the building builds up until Wednesday and then it gradually drops until Saturday is present throughout the data (3) weekends are very random and with very low occupancy. These three facts suggest us to model each day of the week separately, and not to make great efforts to model weekends accurately.

For a typical weekday it becomes rapidly apparent that the schedule will have the characteristics of exactly one of four schedules depicted in Figure 15.

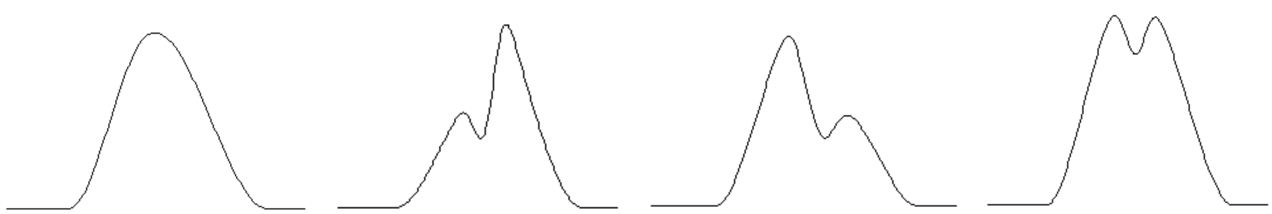


Figure 15: Typical Daily Schedules

The three last daily schedules in Figure 15 show two picks surrounding a lunch period happening at around 12:30 PM. It is typical too to see a big difference in the number of people before and after this lunchtime as the two schedules in the middle show.

A reduced set of parameters describing all of these shapes would allow us to track the schedule changes in time (from the data), and to produce Markov Chain models that capture the average behavior and the variability of the people occupancy (this procedure will be shown in the next section).
The set of parameters has to be small but sufficient to reconstruct the daily occupancy evolution. Functional forms (such as polynomials, trigonometric, and harmonic functions) are not adequate, since too many of them are needed to capture with fidelity the "flat ends" that characterize the

schedules shown in Figure 15 (and Figure 2, Figure 14). A different approach is to capture a small number of points in the shape that are enough to reconstruct the rest of the shape using, for example, cubic spline interpolation. A set of five points/parameters sufficient for this task is the following:

i. Number of local maxima (either one or two)
ii. Maxima (either one maxima or two)
iii. Time at which maxima occurs
iv. Time of lunch occupancy drop
v. Impact of lunch break (depth of the "valley")

Figure 16 shows an example of a shape with two maxima.

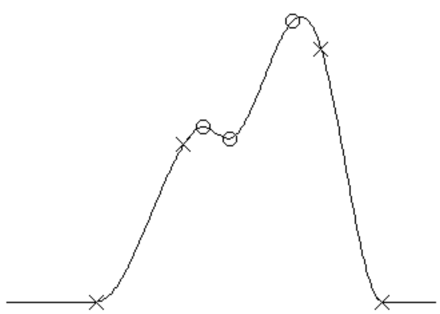


Figure 16: Typical Daily Schedules

The information used for constructing Figure 16 is the fact that there will be two maxima and the points marked with circles, which correspond precisely to the information listed for parameters ii. - v. With that information, and using the facts that (1) in average there are no occupants in the building before 5:00 AM and after 9:00 PM and (2) the drops from the peaks are usually of 10% before the first peak and after the second peak, we can reconstruct the shape using cubic spline interpolation. The last two information sets described are marked in the example in Figure 16 with crosses.

Stochastic Occupancy Schedules

Once we have a way to describe the occupancy evolution for each weekday in the data with a small number of parameters, we can track the changes by separating the parameters in bins and counting the successive changes through time.

For example, consider the parameter "height of the peaks of Wednesdays". Figure 17 shows all the peak-heights available from the data for the different Wednesdays available as they evolve through time from left to right.

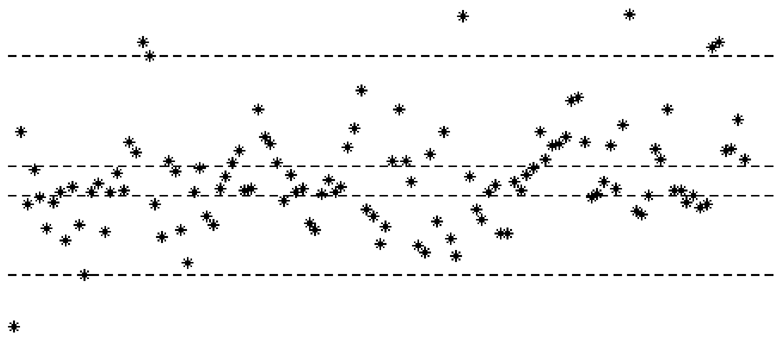


Figure 17: Wednesday Peak-Heights

Then, as shown in Figure 17, the parameters are classified in 3 "equally populated" bins, which define 3 different height levels (or states): low, medium and high. The bins are actually defined in such a way to discard outliers and to define equal percentiles. Then we can count from left to right the number of times that a peak "transitions" from low to medium, or low to high, and every other possible combination for a total of nine. Arranging these transition counts in matrix form and normalizing each row, we obtain a transition matrix as defined before. For the example in Figure 17 the matrix is the following:

$$P = \begin{pmatrix} 0{,}4444 & 0{,}4167 & 0{,}1389 \\ 0{,}3243 & 0{,}4595 & 0{,}2162 \\ 0{,}2105 & 0{,}1316 & 0{,}6579 \end{pmatrix}$$

Since each bin defines a state with a range of heights associated, then as described in methods section knowing the "current" state we can obtain randomly the future state according to the transition matrix distribution and once having the state in hand we can draw uniformly a height from the corresponding range (bin).

Repeating this procedure for every parameter for each day of the week we obtain all the transition matrices needed to simulate the evolution of the schedules through time. The usage of 3 bins is enough to capture the variability of the processes without overcomplicating the model.

Figure 18 shows a random sample constructed with all the transitions matrices calculated for each day of the week. The process is started from the average state for every transition matrix calculated as the most likely state in the stationary distribution (see methods section). Figure 18 shows one of the successive samples.

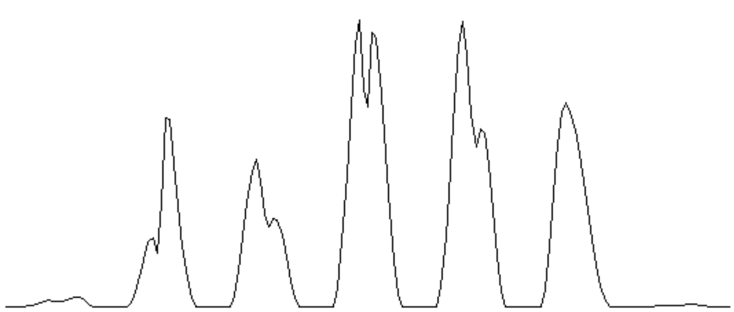


Figure 18: Stochastically Generated Week Schedule

It is clear that, since the sampling is random, that even starting from the same initial set of states every schedule will be in general different.

Energy Estimation

We compare now the daily whole-building energy consumption predicted by the energy plus models using the different types of schedules we have reviewed with actual energy consumption measurements reported in the dataset. We include also the "real" schedule that we have available from the actual occupancy measures, which is fed into the energy model. We call this last schedule the *raw* schedule.

Figure 19 shows the metered energy reported on the CDH website data as the blue line, the simulation with the original step-like occupancy schedules in green, the simulation with the raw occupancy data in red (as in Figure 2), the simulation with the average schedule in cyan, and the parameterized stochastic Markov Chain model in magenta. The invisible portions of the crude step-like schedule simulation (green), the raw data schedule simulation (red), and the average schedule simulation (cyan), are hidden due to a heavy overlap with the richer stochastic schedule simulation (magenta).

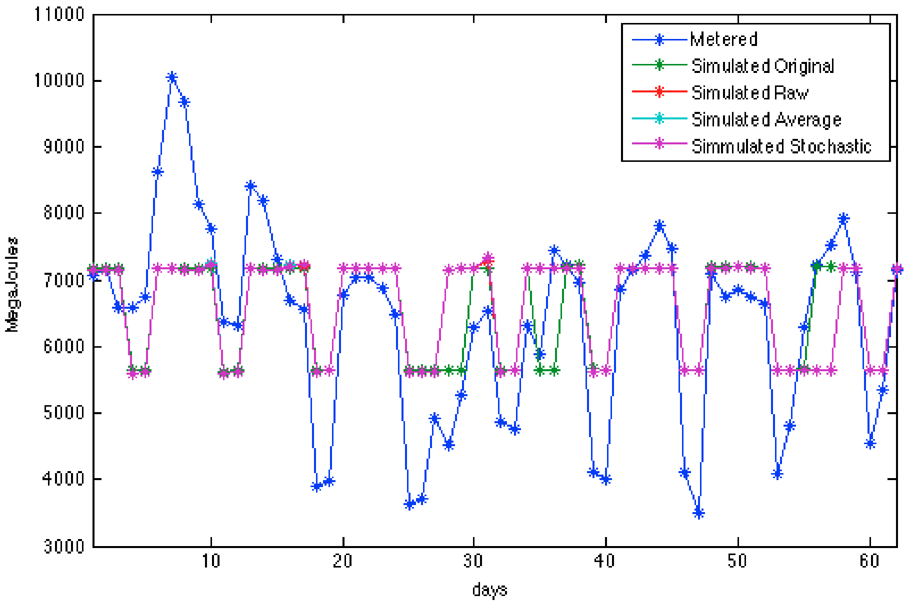


Figure 19: Predicted Energy Consumption versus Metered Energy Usage

It is clear that, quantitatively, there is not a big difference in the energy estimations for all the simulated models, with maybe the exception of the original step-like occupant schedule. It is also clear that the simulations are not matching the reported energy consumptions. Regardless of this we will try to see if any of the models can be seen to have fared better in any sense.

Comparison Raw Data Schedule

First we try to see if there is a quantitative difference in using the raw data, which should be the more precise schedule, versus using one of our proposed simplified schedules. Table 5 shows the errors (Euclidean Distance) of the predicted energy yielded by our schedules versus the raw data fed directly as a schedule.

	Step-like	Average	Stochastic
Error (Mega Joules)	3759.6	66.7	52.9
Relative Error	7.1%	0.13%	0.1%

Table 5: Raw Data Simulation Versus Other Schedules

We see that, if we consider the raw data schedule as the most precise, then the average and fully stochastic schedules report significantly less error than the original step-like schedule. This could mean that the introduction of sophistication in the schedules is justified. On the other hand, there is quantitatively no difference in using the average versus the fully stochastic schedules.

Comparison to Metered Energy

We compare now all four simulated schedules (raw occupancy data, step-like, average, and stochastic) to the metered data.

Table 6 shows the errors (Euclidean Distance) of the metered energy versus the predicted energy consumptions of each model.

	Raw Data	Step-like	Average	Stochastic
Error (Mega Joules)	8935	8150	8940	8939
Relative Error	17.3%	15.7%	17.3%	17.3%

Table 6: Metered Data Versus Simulations

As mentioned before, all the models yield unacceptable error levels of the same order of magnitude. There are no significant differences in this regard.

Agent-based model

The agent-based model is still undergoing calibration using field data on occupant behavior that is being collected through January 31, 2014, the due date of this draft report. Future research will ask how well the agent-based model calibrates to observed occupant behavior, how the EnergyPlus model performs when hot-linked to the agent-based model (compared to observed energy consumption), and how agent-based model/EnergyPlus combo calibrates to reported occupant comfort and satisfaction data.

The calibration of the occupant behavior model of Building 101 will proceed in two steps. First, the measured annual energy consumption in kBtu/sq.ft/year will be compared with the modeled energy use in kBtu/sq.ft/year, yielding a % error calculation. Second, the fieldwork results on occupants' response towards the building environment, will be compared with the modeled occupant behavior to provide a basis for estimating % error in behavioral factors. To test the validity of the modeled Building 101, sensitivity runs will be made for the building.

Comparison across models

As future research, it will be worthwhile to compare observed energy consumption, reported occupant comfort and satisfaction, results from the EnergyPlus model developed by Xu (2012),the Markov chain model, and the Agent-based model.

4. Discussion

Xu (2012) identifies several shortcomings of the conventional EnergyPlus model of Building 101. It mis-predicts central HVAC system energy consumption, mis-predicts plug loads, and the assumed occupancy schedules are not accurate.

The schedules obtained by the Markovian 5-Parameter Model seem visually realistic. The methodology here exposed could be used to simulate, in an at least qualitatively correct way, other random processes to capture variations in time.

Additionally, if the building model is not too sensitive to changes in occupancy which could be due to calibration or the fact that for commercial buildings most of the energy consumption comes from big stationary equipment, then it is possible that changing the schedule to a more sophisticated one will not make an improvement on the energy consumption estimations at least in comparison with the reported metered consumption. If the model is sensitive enough, then the average schedule and the stochastic schedule fare very similar to the actual occupancy data schedule.

The assertion that realistic stochastic schedules (which include variations) are better than averaged schedules or better than step-like schedules is at this point questionable.

The Markovian analysis provides little evidence of seasonality in building occupancy. This work demonstrates that a parsimonious version of Markov chain modeling is feasible in this domain.

Regarding occupants' adaptive responses, we expect future research to show that modeling it using agent-based techniques does not help predict energy consumption in this building. We expect that it will provide good basis for predicting occupant comfort, satisfaction, and maladaptive responses that will need attention from building manager.

We note two limits of this research that extend beyond its incomplete status. First, other buildings might perform differently than the retrofitted commercial building modeled here (especially residential buildings where occupant behavior matters more). Second, the current research does not include attempts to influence occupant behavior (e.g., Energy Chickens, dashboards, 3-net leases), although the modeling framework supports these possibilities once data become available.

5. Conclusions

We offer the following tentative conclusions based on work completed to date:

- Better occupancy data greatly improves energy model accuracy
- Standard assumptions about occupant schedules are often wrong so that a more sophisticated representation is warranted
- Better data about occupants' adaptive responses only marginally improves energy model accuracy
- Yet such data are quite valuable for predicting occupant satisfaction
- EnergyPlus needs additional hooks for incorporating occupant behavior.

6. Acknowledgements

The authors gratefully acknowledge the assistance of colleagues at the Rutgers Center for Green Building and the Energy Efficient Buildings Hub. Funding was provided by National Science Foundation grant #CMS-0725503 and the Energy Efficient Buildings Hub sponsored by the U.S. Department of Energy under Award Number DE-EE0004261.

7. References

Abreu, J.M., Pereira, F.C., and Ferrão, P. Using pattern recognition to identify habitual behavior in residential electricity consumption. *Energy and Buildings* 49(2012): 479-487.

Andrews, C.J., H. Chandra Putra, and C. Brennan. "Simulation Modeling of Occupant Behavior in Commercial Buildings". Prepared by the Center for Green Building at Rutgers University for the Energy Efficient Buildings Hub, Philadelphia, PA. 2013. At http://rcgb.rutgers.edu/wp-content/uploads/2013/09/EEBHub_ModelingReport_2013.pdf

Andrews, C.J., Senick, J.A., and Wener, R.E. Incorporating occupant perceptions and behavior into BIM, chapter 20 (pp. 234-46) in Shauna Mallory- Hill, Wolfgang F.E. Preiser & Chris Watson, eds. *Enhancing Building Performance*, Oxford: Blackwell, 2012.

Andrews, C.J., Senick, J.A., Wener, R.E., and Sorensen-Allacci, M. Investigating Building Performance Through Simulation of Occupant Behavior. *GreenBuild 2012 Proceedings*, pp. 16-21.

Andrews, C.J., Yi, D., Krogmann, U., Senick, J.A., and Wener, R.E. Designing Buildings for Real Occupants: An Agent-Based Approach. *IEEE Transactions on Systems, Man, and Cybernetics--Part A: Systems and Humans* November 2011, 41(6): 1077-1091.

Axelrod, R. *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*, Princeton University Press, Princeton, NJ, 1997.

Azar, E., and Menassa, C.C. A comprehensive analysis of the impact of occupancy parameters in energy simulation of office buildings. *Energy and Buildings* 55 (2013): 841-853.

Blight, T.S., and Coley, D.A. Sensitivity analysis of the effect of occupant behaviour on the energy consumption of passive house dwellings. *Energy and Buildings* 66 (2013): 183–192.

Cao, B., Zhu, Y., Ouyang, Q., Zhou, X., and Huang, L. Field study of human thermal comfort and thermal adaptability during the summer and winter in Beijing. *Energy and Buildings* 43(2011): 1051-1056.

Chen, J., Taylor, J.E., and Wei, H.H. Modeling building occupant network energy consumption decision-making: The interplay between network structure and conservation. *Energy and Buildings* 47(2012): 515-524.

Chiou, Y.S., Carley, K.M., Davidson, C.I., and Johnson, M.P. A high spatial resolution residential energy model based on American Time Use Survey data and the bootstrap sampling method. *Energy and Buildings* 43(2011): 3528-3538.

Choi, J.H., Loftness, V., and Aziz, A. Post-occupancy evaluation of 20 office buildings as basis for future IEQ standards and guidelines. *Energy and Buildings* 46(2012): 167-175.

Davis, J.A., and Nutter, D.W. Occupancy diversity factors for common university building types. *Energy and Buildings* 42(2010): 1543-1551.

Dodier, R.H., G. P. Henze, D. K. Tiller, and X. Guo, "Building occupancy detection through sensor belief networks," *Energy and Buildings*, vol. 38, pp. 1033–1043, Sept. 2006.

Duarte, C., van den Wymelenberg, K. and Rieger, C. Revealing occupancy patterns in an office building through the use of occupancy sensor data. *Energy and Buildings* 67(2013): 587-595.

Epstein, J.M. *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton University Press, Princeton, NJ, 2006.

Frontczak, M., and Wargocki, P. Literature survey on how different factors influence human comfort in indoor environments. *Building and Environment* 46(2011): 922-937.

Fumo, N., Mago, P., and Luck, R. Methodology to estimate building energy consumption using EnergyPlus Benchmark Models. *Energy and Buildings* 42(2010): 2331-2337.

Haldi, F., and Robinson, D. On the behaviour and adaptation of office occupants. *Building and Environment* 43(2008): 2163-2177.

Herkel, S., Knapp, U. and Pfafferott, J. Towards a model of user behaviour regarding the manual control of windows in office buildings. *Building and Environment* 43(2008): 588-600.

Hoes, P., Hensen, J.L.M., Loomans, M.G.L.C., de Vries, B., and Bourgeois, D. User behavior in whole building simulation. *Energy and Buildings* 41(2009): 295-302.

Hopfe, C.J., and Hensen, J.L.M. Uncertainty analysis in building performance simulation for design support. *Energy and Buildings* 43(2011): 2798-2805.

Indraganti, M. Behavioural adaptation and the use of environmental controls in summer for thermal comfort in apartments in India. *Energy and Buildings* 42(2010): 1019-1025.

Indraganti, M., and Rao, K.D. Effect of age, gender, economic group and tenure on thermal comfort: A field study in residential buildings in hot and dry climate with seasonal variations. *Energy and Buildings* 42(2010): 273-281.

Kalmár, F., and Kalmár, T. Alternative personalized comfort. *Energy and Buildings* 65(2013): 37-44.

Karjalainen**, S.** Should it be automatic or manual—The occupant's perspective on the design of domestic control systems. *Energy and Buildings* 65(2013): 119-126.

Kim, G.Y., Kong, H.J., Kim, H., and Kim, J.T. A field survey of visual comfort and lighting energy consumption in open plan offices. *Energy and Buildings* 46(2012): 146-151.

Kulkarni, V.G. *Modeling and Analysis of Stochastic Systems*, Second Edition (Chapman & Hall/CRC Texts in Statistical Science, Dec 18, 2009.

Kyrö, R., Heinonen, J., Säynäjoki, A., and Junnila, S. Occupants have little influence on the overall energy consumption in district heated apartment buildings. *Energy and Buildings* 43(2011): 3484-3490.

LBNL (The Board of Trustees of the University of Illinois and the Regents of the University of California through the Ernest Orlando Lawrence Berkeley National Laboratory), "The Encyclopedic Reference to EnergyPlus Input and Output", 1996-2013.

Liu, J., Yao, R., and McCloy, R. A method to weight three categories of adaptive thermal comfort. *Energy and Buildings* 47(2012): 312-320.

Logic Comput., 8(3): 293–342, Jun. 1998.

Masoso, O.T., and Grobler, L.J. The dark side of occupants' behaviour on building energy use. *Energy and Buildings* 42(2010): 173-177.

Meidani, H., and Ghanem, R. Multiscale Markov models with random transitions for energy demand management. *Energy and Buildings* 61(2013): 267-274.

Page, J., Robinson, D., Morel, N., Scartezzini, J.L. A generalised stochastic model for the simulation of occupant presence. *Energy and Buildings* 40(2008): 83-98.

Pfafferott, J., and Herkel, S. Statistical simulation of user behaviour in low-energy office buildings. *Solar Energy* 81(2007): 676-682.

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

Rao, A. S., and Georgeff, M. P. Decision procedures for BDI logics. *J.*

Richardson, I., Thomson, M., and Infield, D. A high-resolution domestic building occupancy model for energy demand simulations. *Energy and Buildings* 40(2008): 1560-1566.

Ryan, E.M., and Sanquist, T.F. Validation of building energy modeling tools under idealized and realistic conditions. *Energy and Buildings* 47(2012): 375-382.

Saelens, D., Parys, W., and Baetens, R. Energy and comfort performance of thermally activated building systems including occupant behavior. *Building and Environment* 46(2011): 835-848.

Santin, O.G. Behavioural Patterns and User Profiles related to energy consumption for heating. *Energy and Buildings* 43(2011): 2662-2672.

Shen, W., Shen, Q., and Sun, Q. Building Information Modeling-based user activity simulation and evaluation method for improving designer–user communications. *Automation in Construction* 21(2012): 148-160.

Stoppel, C.M., and Leite, F. Integrating probabilistic methods for describing occupant presence with building energy simulation models. *Energy and Buildings* 68(2014): 99-107.
Supplemental Material

Swan, L.G., Ugursal, V.I., and Beausoleil-Morrison, I. Occupant related household energy consumption in Canada: Estimation using a bottom-up neural-network technique. *Energy and Buildings* 43(2011): 326-337.

Tabak, V., and de Vries, B. Methods for the prediction of intermediate activities by office occupants. *Building and Environment* 45(2010): 1366-1372.

Tanimoto, J., Hagishima, A., and Sagara, H. A methodology for peak energy requirement considering actual variation of occupants' behavior schedules. *Building and Environment* 43(2008): 610-619.

U.S. Department of Energy (USDOE). 2012. EnergyPlus.

USDOE. EnergyPlus Weather Data, http://apps1.eere.energy.gov/buildings/energyplus/weatherdata_about.cfm, US Department of Energy.

USDOE. EnergyPlus website, http://apps1.eere.energy.gov/buildings/energyplus/energyplus_about.cfm, US Department of Energy.

Virote, J. and Neves-Silva, R. Stochastic models for building energy prediction based on occupant behavior assessment. *Energy and Buildings* 53(2012): 183-193.

Widén, J., Nilsson, A.M., and Wäckelgård, E. A combined Markov-chain and bottom-up approach to modelling of domestic lighting demand. *Energy and Buildings* 41(2009): 1001-1012.

Wilensky, U., & Rand, W. An introduction to agent-based modeling: Modeling natural, social and engineered complex systems with NetLogo. MIT Press, Cambridge, MA, 2014.

Wilke, U., F. Haldi, J.L. Scartezzini, and D. Robinson, "A bottom-up stochastic model to predict building occupants' time-dependent activities," *Building and Environment*, vol. 60, pp. 254–264, Feb. 2013.

Xu, K. *Assessing the minimum instrumentation to well tune existing medium sized office building energy models*. Doctoral dissertation in Architectural Engineering. Pennsylvania State University. University Park, PA.

Yun, G.Y., Kim, H., and Kim, J.T. Effects of occupancy and lighting use patterns on lighting energy consumption. *Energy and Buildings* 46(2012): 152-158.

Yun, G.Y., Tuohy, P., and Steemers, K. Thermal performance of a naturally ventilated building using a combined algorithm of probabilistic occupant behaviour and deterministic heat and mass balance models. *Energy and Buildings* 41(2009): 489-499.

Zhang, T., Siebers, P.O., and Aickelin, U. Modelling electricity consumption in office buildings: An agent based approach. *Energy and Buildings* 43(2011): 2882-2892.