

Article

Development and Evaluation of Occupancy-Aware HVAC Control for Residential Building Energy Efficiency and Occupant Comfort

Christina Turley¹, Margarite Jacoby¹, Gregory Pavlak² and Gregor Henze^{1,3,4,*}

- ¹ Department of Civil, Environmental, and Architectural Engineering, University of Colorado, Boulder, CO 80309, USA; christina.turley@colorado.edu (C.T.); margarite.jacoby@colorado.edu (M.J.)
- ² Department of Architectural Engineering, Pennsylvania State University, University Park, PA 16802, USA; gxp93@psu.edu
- ³ National Renewable Energy Laboratory, Golden, CO 80301, USA
- ⁴ Renewable and Sustainable Energy Institute, Boulder, CO 80309, USA
- * Correspondence: gregor.henze@colorado.edu; Tel.: +1-303-492-1094

Received: 25 August 2020; Accepted: 2 October 2020; Published: 15 October 2020

Abstract: Occupancy-aware heating, ventilation, and air conditioning (HVAC) control offers the opportunity to reduce energy use without sacrificing thermal comfort. Residential HVAC systems often use manually-adjusted or constant setpoint temperatures, which heat and cool the house regardless of whether it is needed. By incorporating occupancy-awareness into HVAC control, heating and cooling can be used for only those time periods it is needed. Yet, bringing this technology to fruition is dependent on accurately predicting occupancy. Non-probabilistic prediction models offer an opportunity to use collected occupancy data to predict future occupancy profiles. Smart devices, such as a connected thermostat, which already include occupancy sensors, can be used to provide a continually growing collection of data that can then be harnessed for short-term occupancy prediction by compiling and creating a binary occupancy prediction. Real occupancy data from six homes located in Colorado is analyzed and investigated using this occupancy prediction model. Results show that non-probabilistic occupancy models in combination with occupancy sensors can be combined to provide a hybrid HVAC control with savings on average of 5.0% and without degradation of thermal comfort. Model predictive control provides further opportunities, with the ability to adjust the relative importance between thermal comfort and energy savings to achieve savings between 1% and 13.3% depending on the relative weighting between thermal comfort and energy savings. In all cases, occupancy prediction allows the opportunity for a more intelligent and optimized strategy to residential HVAC control.

Keywords: HVAC control; occupancy prediction; energy consumption; thermal comfort

1. Introduction and Background

The finite quantity of fossil fuels and the mounting concern of climate change makes reducing energy use a global necessity. Buildings are major consumers of energy worldwide, and used around 3060 million tons of oil equivalent (Mtoe) in 2018 according to the International Energy Agency (IEA) [1]. In the United States, heating, ventilation, and air conditioning (HVAC) systems account for 50% of all building energy consumption [2], while U.S. homes alone are responsible for the use of approximately 4.7 quadrillion British thermal units (Btu) for space heating and air conditioning per year [3]. Therefore, reducing energy



consumption associated with residential heating and cooling has the potential to result in large energy savings when applied across the sector.

Traditionally, heating and cooling in residential buildings has been controlled by a thermostat that has a single setpoint temperature, which keeps the indoor temperature constant whenever the thermostat is in use. Over time, different technologies have been added to HVAC systems to improve temperature control and reduce energy use, one of which is occupancy-based control. This method controls the indoor temperature to provide thermal comfort only when the building is believed to be occupied, and turns the HVAC system off when it is vacant. This typically results in reduced energy use during unoccupied hours. Previous studies have estimated that the potential savings when using these systems is between 5–23%. The magnitude of savings depends on various factors, such as climate, building vintage, and occupant behavior [4]. While occupancy-based HVAC control has potential benefits, questions still remain on how best to detect occupancy and implement control decisions. To provide the best experience for occupants, thermal comfort standards should be met during all occupied hours. Thus, a good control strategy needs to not only know when a building is currently occupied, but also needs to accurately predict occupancy ahead of an occupants arrival. This allows the space to be appropriately conditioned in advance of the arrival.

1.1. Historical Trends in U.S. Housing

An understanding of how buildings are changing is critical to reducing energy use in the built environment. Looking at how buildings and control systems have performed historically can highlight opportunities for improvement and can indicate the ways in which current trends may shape the future. The 2015 Residential Energy Consumption Survey found that the United States residential sector is comprised of 118.2 million homes, totaling 223 billion square feet of floor space [3]. Residential buildings currently use 22% of U.S. annual energy, and in the three decades from 1980 to 2009 residential building site energy use increased by 8.9% [5]. This growth can be attributed to increases in three factors: home size, number of homes, and appliance use. For instance, the number of households in the U.S. increased by 33%, while the average size of a single-family detached home also increased from 2100 square feet to 2688 square feet, as depicted in Figure 1. This led to a 52% increase in total floor space [5]. Additionally, appliance electricity consumption during the same time period increased by 30.6%, with the largest increases being from microwave ovens, personal computers, air-conditioners, and clothes dryers [5].



Figure 1. Size of residences by home type for 1980 and 2009 (ft²) [5].

At the same time that energy use was increasing due to changes in housing characteristics, other factors were leading to decreases in energy use. These decreases are attributable to: (1) population shifts in the U.S., as large numbers of people moved from the Northeast and Midwest to the less heating-intensive regions of the South and West; (2) changes in weather patterns, both heating-degree days and cooling degree-days were lower across the nation in 2009 than in 1980 [5]; and (3) a decline in energy intensity, led by advances in engineering and a promotion of energy efficiency standards for household appliances. For example, the annual fuel utilization efficiency (AFUE) of a standard furnace increased from 78% to 97%, leading to decreased energy consumption for the same heating output. The largest change in consumption occurred from 1990–2001, which coincides with an era when federal efficiency programs, like ENERGY STAR, were enacted. The combined effect of all contributors in the 30-year period was an increase in energy consumption over time, with U.S. homes consuming 9.1 quads per year [5].

1.2. Temperature Control in Buildings

Despite changes in the housing sector, occupants' desires for thermal comfort have remained constant. One method of measuring occupant comfort is the predicted mean vote (PMV), first developed by Povl Ole Fanger [6], which predicts the average comfort level of a hypothetical group of people in a space. ASHRAE Standard 55, first published in 1966, specifies the fraction of occupants that find a space comfortable using the predicted mean vote must be at least 80% [7]. PMV, which ranges from -3 (too cold) to +3 (too hot), is based on the combined effects of air temperature, mean radiant temperature, relative humidity, air speed, metabolic rate (based on activity), and occupants' clothing levels.

In buildings, comfort requirements are met by using a thermostat and control system to maintain a setpoint temperature. The thermostat measures the indoor air temperature and compares it to the setpoint temperature, while the control system manages how the HVAC system tracks the indoor air temperature, attempting to keep it within small deviations of the setpoint. The interaction of these two components, and the programming of the system, determines how well the setpoint is tracked and how effectively the system achieves thermal comfort conditions.

Manual, programmable, and "smart" thermostats are the three main categories of thermostats in-use today. In manual thermostats, the setpoint is a single temperature that the system always tries to maintain when it is on. To change the temperature of the space, you must manually change the setpoint. Programmable thermostats are similar, but with different temperature setpoints for different times of the day or days of the week that can be programmed by users [8]. This allows temperatures to be setback during nighttime hours or during daytime vacancies, and automatically adjusts to more comfortable temperatures when people are frequently home. These thermostats often have modes for different days (weekday and weekend), and modes for different times of day (e.g., morning, day-time, evening, and night). Endorsed by Energy Star at their 1995 release, initial demonstrations showed that programmable thermostats could reduce heating and cooling bills by 10%–30%. However, the U.S. Environmental Protection Agency ultimately suspended Energy Star certification of programmable thermostats in 2009 since a lack of undisputed energy savings materialized [8]. Investigations revealed that 30% of households had failed to set them properly, and over 89% had not set separate weekend and weekday schedules [9]. Due to their complexity of operation, most programmable thermostats were operated manually, negating their energy savings potential.

Connected, or "smart", thermostats have emerged in response to consumer aversion to programmable thermostats. Like programmable thermostats, connected thermostats create a setpoint schedule but the operation is designed to be user-friendly and may change over-time, given occupancy patterns. Products, like the Nest Learning Thermostat, Honeywell, or Ecobee thermostat, are internet or "cloud" connected and can be controlled by phone, web interface, or a touchscreen. The system comes with a preset schedule

4 of 30

that modifies itself based on the user's manual adjustments during initial use [10]. First introduced in 2011, the Nest Learning Thermostat catalyzed the market, with over 100% market growth per year in the first three years. In 2015, 40% of the thermostats sold were connected thermostats [10]. The main features of most connected thermostats include extensive data tracking, remote accessibility, local sensors to track occupancy, and web-enabled weather forecast data. While all connected thermostats are designed to enhance temperature control, they also create an opportunity to save energy through the use of setback temperatures that are automatically programmed though the initial use.

1.3. HVAC Control Strategies

The control strategy utilized by an HVAC system determines how the heating and cooling is managed. The two main controller types use in residential buildings are discrete controllers and continuous controllers. A discrete controller, in its simplest form, simply turns devices on and off, allowing only two states of operation. For example, when indoor temperatures are below a certain threshold (often a degree below the setpoint temperature) the heater is actuated. The heater is then turned off once the setpoint temperature is reached. These controllers offer ease of installation and operation, however, they often suffer from overshoot and undershoot, making it difficult to maintain a precise indoor temperature. This occurs because of the large thermal inertia of buildings and their engineered systems, which can result in large deviations from setpoint temperatures [11]. If temperatures are more precisely maintained, then the system will be frequently cycling on and off. Frequent cycling can be damaging to equipment and annoying to occupants.

Continuous controllers, on the other hand, modulate heating and cooling to provide heat transfer at the rate that it is needed to reject disturbances and track the setpoint. Although many continuous controllers only use current indoor temperature as an input, they can provide much more precise setpoint tracking, as the building response dynamics are accounted for in the set-up (i.e., tuning) process. Continuous control is normally provided by proportional-integral (PI) or proportional-integral-derivative (PID) local loop feedback controllers, which attempt to minimize undershoot, overshoot, rise time, settling time, and steady-state error [9].

More advanced control systems can take additional inputs, such as future building occupancy and predicted weather. One such strategy considered in this work, called Model Predictive Control (MPC), predicts the future state of the building by incorporating weather forecasts and current indoor temperature. These inputs are fed into a model to predict how the building will change under a variety of different HVAC actions. An optimization is then performed to determine which action will achieve the required temperature while minimizing energy use. The optimal control action is then sent to and implemented by the HVAC system [12,13], and the cycle will be performed again. By predicting future states and correcting for state prediction errors at every time interval, MPC acts as closed loop, real-time building controller. Known as receding horizon control, a newly optimized control strategy is determined as temperature and weather forecasts are updated [14].

MPC frequently utilizes a reduced order linear dynamic model that represents the building as an equivalent circuit of thermal resistances and capacitances (RC) [13]. This means that the heat transfer in and out of the building is simplified to linear expressions, making the optimization problem convex and easier to solve. Buildings, however, do not always act linearly, which leads to modeling errors and mismatch [13]. The extensive time and effort required for properly calibrating a model for individual buildings has kept MPC from widespread adoption [15]. Yet, MPC continues to show promise and is predicted to gain traction with research showing residential energy savings of 28% on average and cost savings of 16% [15,16].

1.4. Potential of Occupancy-Aware HVAC Systems

When occupancy is included in HVAC control, the system behaves conventionally when the building is believed to be occupied, attempting to meet a pre-programmed temperature setpoint. In contrast, when the space is believed to be vacant, the HVAC system allows the indoor temperature to drift somewhat by using a more liberal setback temperature setpoint, minimizing total energy use. Occupancy-based controls can be either reactive or predictive. In reactive control, the system detects an occupant in the space and then turns on the system. This can lead to uncomfortable temperatures when an occupant first enters a space, as the system may not be able to immediately reach the new setpoint. In predictive control, the arrival time of the occupant is predicted and the system preheats or precools the space so that indoor temperature reaches the setpoint just before the occupant arrives, minimizing energy while maintaining comfortable temperatures during all occupied hours [17]. However, correctly predicting when an occupant will arrive is challenging, as the behavior of individuals is difficult to model.

Industry professionals have been working to ascertain the energy savings potential for occupancy-aware HVAC control. In 2014, Nagele et al. conducted a survey of 30 households in southern Germany over a period of 14 months [9]. They then used the data collected, such as temperature setpoints and house characteristics, to calculate energy use under eight different control strategies for ten simulated households. Using a constant temperature on/off controller as the reference case, they showed that PID controllers, setback temperatures, model predictive control (MPC), and occupancy-based HVAC control all have the ability to reduce energy use, when implemented correctly. See Figure 2 for a comparison of the results.



Figure 2. Boxplot of potential savings by control strategy for ten households [%]; adapted from Reference [9].

The reactive strategy of simple occupancy detection offers the largest potential energy savings, but can also increase unmet comfort hours. If a space is routinely uncomfortable when an occupant comes home, then they are likely to turn the detection control off. Thus, to gain consumer adoption, unmet hours need to be low enough that consumers use the functionality. This makes occupancy prediction the preferred control choice. In simulation studies of occupancy-prediction control, savings are estimated to be between 6%–48%, and depend on factors, such as climate, insulation levels, and occupancy schedules [17,18]. Beyond simulations, utilities have measured the energy savings of connected thermostats, which often employ occupancy-aware controls [19]. In reviewing 35 studies from 2007 to 2016, the U.S. Department of Energy reported energy savings that ranged from 1% to 15% [10]. Definitive values for savings are hard to determine due to variances in hardware, software, buildings, occupant behavior, and local weather.

1.5. Occupancy in Buildings

In order to effectively incorporate occupancy information into building controls, occupancy patterns and their impacts on energy use must first be understood. A 2017 international survey of building energy professionals and researchers listed occupant behavior as the largest contributing factor to energy modeling errors [20]. This is because of the varied and stochastic nature of human behavior, which changes dramatically from person-to-person and from day-to-day. This makes correctly predicting occupant behavior, and therefore its effect on buildings, extremely difficult.

In 1978, Robert Socolow published a 5-year observational study on occupant behavior where they tracked gas consumption of 205 identical townhouses in Twin Rivers, New Jersey, finding a 33% variation in consumption [21]. This revealed that seemingly identical buildings can vary due to factors, such as occupants' setpoint temperatures and hot water use. Similarly, a study in Kuwait showed that residents used setpoint temperatures that varied between 19 °C and 25 °C for air-conditioning, with electricity use increasing 21% with a 2 °C change [22]. In Denmark, Rune Andersen collected four years of annual heating data from 290 identical townhouses. Again, a wide variation was found with annual heating consumption ranging from 9.7 kWh/m² to 197 kWh/m², a ratio of 20 to 1 [23]. These studies, carried out in different climates, across continents and several decades, show the significant variation of occupant behavior and the impact it has on energy consumption.

1.6. Modeling Occupant Presence

Modeling the impact of occupancy behavior on energy consumption is comprised of two steps. First, a researcher must create a reasonably accurate model of occupancy. Second, this model is incorporated into a building performance simulation (BPS). Research in the past decades has investigated the best method for performing each step, both of which are necessary to understand the impact of occupancy-based HVAC control on energy use.

While an accurate occupancy model is critical to understanding the impact of occupant behavior, a 2017 industry survey showed that industry professionals believe current models over-simplify real behaviors, leading to inaccurate predictions [20]. A model can be either overly optimistic, in which case actual energy consumption in a building performs below expectations, or overly conservative, which leads to oversized mechanical equipment.

While many occupant models have been published in scientific papers, an industry consensus on what the best model is has not yet been achieved. Occupancy can be modeled and predicted at two levels: group or individual. In the group level, one model is created for the entire group occupying a building. In this method, which is currently the most widely used, the building is essentially the entity being modeled. At the individual level, a separate model is created for each occupant of a building. Some of the most common occupancy models are described below [24].

- Schedules are the current industry standard for modeling occupancy presence. A predetermined
 fraction of occupancy is multiplied by the space density to determine the number of people during
 each hour.
- Deterministic models use a rule-based approach to represent occupancy behavior. Unlike schedules, deterministic models incorporate environmental triggers that can affect actions.
- *Non-probabilistic models* use historical data to create a model. The aggregated data is averaged to create a probability profile, with each time interval having a probability between 0 and 1. If the probability is above a threshold, the building is predicted to be occupied; below the threshold, vacant. Because the profile is created from a training set, the accuracy of the model depends highly on the data used. The model created does not include a stochastic term.

- *Probabilistic or stochastic models* incorporate the variability of human behavior by using randomization. Like non-probabilistic models, stochastic models use historical data to create a model. A probability profile is created and compared to a randomly generated number to classify the space as occupied or vacant. Because a random number is used, a different profile will result each time the model is generated. Stochastic models require multiple runs to achieve reliable results.
- *Agent-based models* model occupants individually, aggregating multiple prediction models to create a full building model. Because modeling is done on an individual basis, the complexity is extremely high.

1.7. Modeling Building Performance

The second step to incorporating occupancy is loading the model into the building simulation, for which there are many simulation programs available. The International Building Performance Simulation Association (IBPSA) lists sixty-seven whole building energy simulation programs [25]. EnergyPlus, developed and distributed by the U.S. Department of Energy (DOE), is used most commonly in occupancy research [20]. EnergyPlus is a compiled physical model, which means the characteristics of the building, such as insulation values, window size, and orientation, are built into the model itself [17], while the mechanical equipment and schedules, such as occupancy, are included as inputs to the building operation. When executed, the simulation calculates the heat and mass transfer for each time step [4]. Simulations are normally performed per annum to integrate both heating and cooling seasons [17]. ASHRAE occupant schedules are embedded within the example EnergyPlus models but different occupant models can also be incorporated in custom models.

1.8. Review of Commonly Used Occupant Models

Past studies have sought to answer the question of which occupant model works best to predict occupancy [14,18,26–34]. Since the published studies were conducted using individually collected data, such as occupancy and climate, and utilized building-specific performance simulations, it has been difficult to directly compare different occupant models [24]. Individual analyses have sought to solve this by comparing different occupancy models made with the same occupancy data. A review of occupant presence comparison studies is summarized in this section.

A study by Mahdavi and Tahmasebi [35] compared three models: two probabilistic models from literature (Reinhart 2004 and Page 2008) and a non-probabilistic model the authors developed. Using data from eight workspaces, the three models were created using four weeks of training data to predict building occupancy over the next 90 working days. Predicted occupancy was compared to measured ground-truth occupancy to analyze the prediction model's capability. The model was evaluated by comparing the arrival time, departure time, duration of occupancy, fraction of correct occupancy state, and number of transitions to the ground-truth data. Analysis showed that the two stochastic models performed similarly, while the non-probabilistic model performed best. Mahdavi and Tahmasebi conclude that, while probabilistic models are suitable for annual simulations, non-probabilistic models are more effective in providing short-term occupant presence predictions [35].

Following their 2015 study, Tahmasebi and Mahdavi [36] input a variety of occupancy models into a building simulation program to determine the effect of the occupancy prediction on building performance. The first model used the ASHRAE 90.1 office schedule. The second used the average group occupancy data for the year, while the third used the average individual occupancy data for the year. A stochastic model for each of the previous three was created to generate a total of six models. An EnergyPlus performance simulation was executed to calculate energy use under each occupancy model. Stochastic models were executed using 100 Monte Carlo runs to find the average performance. The performance of the models

was evaluated using the key performance indicators of annual and peak heating and cooling loads; see Figure 3. It was observed that the ASHRAE schedules performed poorly in all metrics. The stochastic models of individual and grouped occupancy performed better when simulating heating loads, while the non-probabilistic models performed better when simulating cooling loads. Tahmasebi and Mahdavi conclude that known occupancy data is critical for accurate building performance simulation, while stochastic models are not [36].

Duarte et al. [26] performed an occupancy study on a multi-tenant 11 story office building in Boise, Idaho. Using data from 223 private offices over two years, probabilistic models and ASHRAE 90.1 schedules were compared to a non-probabilistic model. Comparing the different occupancy models, the ASHRAE 90.1 schedule overestimated occupancy by as much as 46%. Using data from ten offices for training, the stochastic model matched the training data but not the overall measured occupancy. The authors recommend using a low and high non-probabilistic model because it represents occupancy well without increasing modeling complexity [26].

In all comparison studies, the authors agree that the best model is case specific [24]. Most models are developed using single data and building sets and do not transfer effectively to different building types or occupant behaviors [37]. Despite this, some general conclusions can be drawn:

- ASHRAE occupancy schedules are not reflective of actual behavior.
- Model complexity, such as stochasticity, does not always improve results.
- Models perform best when applied to the case study used to derive them [24].

Since there is no universal occupant prediction model, the IEA Annex 66 consortium recommends choosing a model that matches the complexity levels of the occupant model to the case study. The study presented in this paper aimed to evaluate the possible energy savings on short-term occupancy-based HVAC predictive control. Thus, a non-probabilistic model, which was shown to have the best short term presence prediction, was used [26,35].



Figure 3. Results of building performance simulation (BPS) model accuracy from 2017 study. Adapted from Reference [38].

2. Occupancy Model Generation and Discussion

The goal of this work was to evaluate the impact of occupancy-based HVAC control on residential building models and associated potential energy savings. To facilitate this, actual residential occupancy had to first be determined. Real occupancy data was collected from multiple homes located in Boulder, Colorado, and used to create non-probabilistic occupancy prediction models. The effectiveness of the model was determined by comparing the predicted occupancy against actual occupancy. The details of this process are explained in following subsections.

2.1. Ground Truth Data Collection

Occupancy and eight different physical modalities (e.g., temperature and CO₂) data was collected from six homes for a period of 4–9 weeks each. Occupancy data was collected using a geo-fencing application installed on the users' cell phones, as well as with a paper sign-in sheet by the front door. The two collection methods were cross-referenced by the researchers to confirm correctness. Through both methods, occupant arrival and departure times were recorded for each person residing in the home. Individual occupant data was combined in order to determine the binary occupancy state of the residence. General occupancy information for each residence is shown in Table 1.

House #	Occupant Count	House Type	Days Measured	Avg. Occupancy
1	4	house	64	86%
2	1	apartment	45	56%
3	3	house	71	75%
4	3	apartment	29	82%
5	2	apartment	27	81%
6	1	apartment	63	52%

Table 1. Summary of residences measured during study.

Residences used for the study were chosen from volunteer participants at a university, and consisted of graduate and undergraduate students, a post-doctoral researcher, and a university professor. Since most of the participants were students of some type, their occupancy patterns may be quite different from those collected in a different segment of the population. For instance, the fact that several of the homes contained multiple graduate students with a variety of class and work schedules meant that someone was nearly always home (yielding occupancy rates of 82% and 86%). Furthermore, none of the homes contained children, which might have led to different occupancy patterns. Additionally, occupancy was impacted by extended absences during spring and fall breaks for a few of the homes studied.

While the average occupancy for the testing period ranged from 52% to 86%, the daily occupancy of each home varied. Distributions of daily occupancy rates for each residence are shown in Figure 4. Home 6 (a young graduate student who lived alone) had the largest distribution in daily occupancy, while home 5 had the lowest (married postdoctoral researchers, one of whom worked from home).



Figure 4. Daily occupancy rate for each measured day by house [%].

Beyond daily occupancy, arrival time, departure time, and number of occupancy state transitions were also analyzed. Arrival is defined as the transition of the residence from an unoccupied state to occupied state; departure is the inverse transition. In residences with more than one occupant, the arrival and departure times identified only indicate cases where the residence transferred to or from vacant. Distributions in daily occupancy, arrivals, departures, and occupancy state transitions demonstrate the stochasticity of human behavior. For all residences measured, the arrival and departure times for weekdays differed significantly from those of weekends.

2.2. Occupancy Model Generation

Based on the literature review, a non-probabilistic model was chosen to model occupancy. Since non-probabilistic models use past data to create an occupancy probability, the model can be optimized by establishing what training data to include. Model optimization was done by splitting the collected occupancy data into separate training and testing data sets. Models were trained using the first set, and performance was evaluated by testing the trained model on the second dataset. The only exception to this is the case of the moving training mode, which used a receding horizon. In this case, the model was continually being trained, then tested on unseen data, and then the model was updated after comparison to ground truth.

The percentage of data used for training versus testing was varied, along with three other parameters. Up to 96 different non-probabilistic models were created for each house, with each model using a different subset of training data to create the occupancy profiles. Table 2 shows the different values for each parameter that was used when creating the models.

Day Categorization	Training Time	Training Mode	Time Resolution
day of week	1	fixed	1 min
week/end	2	moving	5 min
mfweekend	3	-	15 min
	4		60 min

Table 2. Variables used for occupancy model creation.

The four group parameters are defined as follows:

• Day categorization: This determines how each day of the week is categorized. For example, in *day of week*, only training data that matches the day being predicted is used. In *week/end* categorization,

all Mondays, Tuesdays, Wednesdays, Thursdays, and Fridays are used to predict occupancy for weekdays. Finally, in *mfweekend* categorization, Tuesday, Wednesday, and Thursday are used to predict weekdays with Monday and Friday kept as separate individual days.

- Training time: This determines how many weeks are used when training the model, ranging from one week to four weeks. Up to 50% of the collected data was used for training, meaning that in residences where only four weeks of total occupancy data were recorded, only two weeks were available for training. In these cases, only 48 total models were generated because the 3-week and 4-week training time was unavailable.
- Training mode: This determines whether the training time is used in a fixed mode (static training set) or in a moving mode (where a trailing horizon is used). For example, in a 1 week moving mode, only the last seven days are used to predict occupancy for that day.
- Time resolution: This determines how often occupancy is sampled. Time is shown in minutes.

After a training set was generated from the collected data, the average occupancy was determined for each interval of the day, resulting in an occupancy probability between 0 and 1. To create a binary occupancy schedule, a threshold probability was set for each day. This threshold was determined by finding what value produced the same minutes of predicted occupancy as the summed occupancy minutes in the input data. An example of a single day for House 1 is shown in Figure 5. Occupancy prediction models were created for a time period of two to five weeks, depending on the length of total measured data.



Figure 5. Occupancy probability, threshold, and resulting model for single day, by time.

2.3. Occupancy Model Accuracy

Once the non-probabilistic models were generated for each home, the resulting predicted occupancy state was compared against the actual occupancy state. To evaluate the effectiveness of the occupancy prediction models, three metrics were used:

- **False negative rate:** Percentage of minutes that the model incorrectly predicted the house was vacant when it was occupied.
- **False positive rate:** Percentage of minutes that the model incorrectly predicted the house was occupied when it was vacant.
- **State matching error:** Percentage of minutes that the model incorrectly predicted occupancy. This is the inaccuracy of the model. The state matching error is the sum of the false negative and false positive rate.

All the metrics are error rates, and so should be minimized. The best models had low values for false negative rate, false positive rate, and state-matching error. False negative errors and false positive errors have different impacts. When a false negative occurs, the house is actually occupied when predicted vacant. In this case, indoor temperature may not be in the comfort range because the control has been changed to a setback temperature. When a false positive occurs, the house is actually vacant when predicted occupied. In this case, the HVAC system may be running to maintain an unnecessarily tight temperature setpoint range, resulting in higher energy use.

Training and evaluation times varied by house due to differences in data collection periods. Table 3 lists how the collected data was used in generating the prediction models. **Training period** shows how many weeks were set aside for determining occupancy probability, and **evaluation period** shows how many weeks were used to evaluate the generated model. The training period is the maximum number of weeks available. Training time, used as a variable, determined how many of the weeks were used in training the model. During the evaluation period, the predicted occupancy was compared against the measured occupancy at one-minute intervals.

House #	Models Created (Count)	Training Period (Weeks)	Evaluation Period (Weeks)
1	96	4	5.1
2	96	4	2.4
3	96	4	6.1
4	48	2	2.1
5	48	2	1.9
6	96	4	5.3

Table 3. Summary of model training and evaluation for each house.

The model configuration with the lowest state matching error is shown for each house in Table 4. House 5 had the lowest state matching error with an error rate of 8%, while House 6 had the highest at 35%. Each house, due to its occupancy pattern, had a different optimal occupancy model, which indicates the value in tailoring the model to the specific use case.

House	Day Categori-Zation	Training Time	Training Mode	Time Resolution	False Negative	False Positive	State Matching Error
1	mfweekend	4	fixed	15	12%	4%	16%
2	weekend	3	moving	5	13%	13%	26%
3	weekend	1	moving	1	28%	6%	35%
4	weekend	1	moving	60	8%	2%	10%
5	weekend	1	moving	60	5%	3%	8%
6	day of week	2	moving	15	7%	30%	37%

Table 4. Best occupancy prediction model for each residence.

To understand the effect of each parameter on the resulting prediction model, the performance results were compiled and analyzed. The parameter with the largest effect on state matching error was the occupancy pattern, shown in Figure 6. The results show that each house, with its different occupancy patterns, has a strong influence on the effectiveness of creating a prediction model. As previous studies showed, the behavior of people has a large variance and can drastically affect prediction models.



Figure 6. State matching error for each house [%].

In contrast to the differences in house occupancy patterns, the parameters that were used to develop the training set each had a small influence on the prediction error. The results of all four parameters are discussed below.

Day categorization On houses with fewer weeks of collected data (House 2, 4, and 5), using *day of week* categorization resulted in the least accurate prediction models. This is likely due to the extremely limited training data for each day. In a *day of week* model, each weekday is treated individually. Thus, if only two weeks of training data are used, then there would be only two instances of each day. In contrast, in House 6, where the occupant had a part-time job that she attended three days a week, the *day of week* method increased the prediction accuracy.

Training time: As would be expected, using more training data improved the accuracy of models for most of the houses. House 3, the exception, had a shift in occupancy patterns halfway through data collection, when some of the occupants went on vacation and extended visitors arrived. This indicates that when new occupants join a household, the previous training data will not effectively predict the new occupancy pattern. To explore this theory further, Figure 7 shows the resulting state-matching error when the training time is extended to seven weeks. With additional training weeks, the error is reduced, indicating that the longer the training data is accumulated, the more the error can be reduced.



Figure 7. State matching error for House 1, by number of training weeks.

Training mode: The moving training mode had improved prediction accuracy for five of the six houses. By allowing the model to adjust over time, the moving training mode adapts to shifting behavior. House 1 had the best prediction accuracy with the fixed mode. The difference between the highest performing fixed mode and highest performing moving mode for House 1 was 0.3%. The higher performance using fixed mode indicates the initial data reflected the general behavior more than later weeks.

Time resolution: Results are nearly identical for each time resolution variable signaling that sampling time does not play a large role in increasing prediction accuracy.

While the lowest state matching error can be achieved by optimizing a non-probabilistic occupancy prediction model to a specific house, a universally applicable model is desirable. This would allow a single model to be deployed in different houses, without the need for preliminary data gathering to determine the parameters. To find the best occupancy model for all the houses surveyed, the state matching error results were normalized by dividing the results by the lowest state-matching error achieved by that house. The lowest error for each house was used to ensure that all of the houses were considered equally. Figure 8 shows a parallel category plot of the results. Occupancy models that were within 5% of the best model for that house are shown in green, while other models are shown in red.



Figure 8. Parallel category plot of occupancy models.

Results show that the *mfweekend* day categorization does not work well for the houses surveyed. *Day of week* and *week/end* both work well, with *week/end* performing better for most homes, suggesting the six houses surveyed do not have distinct Monday or Friday schedules. Four and one week training times were the best, with two and three week training times performing slightly worse. Houses with consistent schedules benefited from the increased data of a longer training time. Houses in which occupants were absent for days benefited from the faster reaction of the shorter training time adjusting to their absence and return. The moving training mode produced the highest number of low state matching errors, which is likely due to its ability to continuously adjust to occupants' behavior changes over time. As seen in the individual evaluations, different time resolutions produced equivalent results, although the 15-min time resolution yielded a slightly higher number of low normalized state matching errors. Based on these results, the optimal universal model across all houses was a *one week* training time *moving* model that uses *week/end* day categorization and a 15-minute time interval. The state matching error for each house with the universal prediction model is shown in Table 5.

Table 5. Results of universal occupancy prediction model for each residence.	

House	Day Categori-Zation	Training Time	Training Mode	Time Resolution	False Negative	False Positive	State Matching Error
1	weekend	1	moving	15	15%	2%	17%
2	weekend	1	moving	15	22%	10%	32%
3	weekend	1	moving	15	28%	6%	35%
4	weekend	1	moving	15	8%	2%	10%
5	weekend	1	moving	15	5%	3%	8%
6	weekend	1	moving	15	11%	33%	44%

3. Building Simulation Setup

Building performance simulations were conducted in EnergyPlus (Version 9.1, National Renewable Energy Laboratory, Golden, CO, USA) to understand the impact of residential HVAC control on energy use for a number of representative home scenarios. Multiple home types, climates, seasons, and occupancy patterns were used to more globally represent the breadth of scenarios, as well as to understand the range of possible outcomes.

3.1. Building Performance Simulation Settings

The six previously collected occupancy data sets were used as the possible occupancy scenarios. Data from these homes were assumed to represent occupancy patterns over all climates and seasons for which the simulations were performed. Five prototype home styles were used, as provided by National Renewable Energy Laboratory (NREL) for the building models [39]. Each home had a different climate and building construction that was representative of the national housing stock. The five locations used were: Boston, Phoenix, Atlanta, Seattle, and Houston. House sizes averaged 2000 ft² with typical home construction and vintages for each region. Details on the home and construction parameters are shown in Table 6.

The occupancy prediction models developed in Section 2.3 ranged from 13 to 36 days depending on the house. A two-week period with one-minute timestep intervals was used for the building performance simulations. Building simulations were performed for two different seasons using the first two weeks of January and the first two weeks of July. Including both winter and summer runs allowed the impact of HVAC control to be ascertained for both heating and cooling modes.

Boston and Atlanta have cold, near-freezing winters, and hot summers. In both of these locations, the outdoor air temperature was well outside of the comfort range for the majority of the simulations. Houston has moderately cold winters and hot summers, while Phoenix has mild winters and extremely hot summers. Both Houston and Phoenix require significant cooling in the summer. During winter in Phoenix, the outdoor air temperature oscillates within the comfort range. Seattle, in contrast, experiences both cool winters and cool summers, requiring some heating year-round.

Heating and cooling temperature setpoints were established using the method of predicted mean vote (PMV). ISO EN 7730 establishes three comfort categories using operative temperature. These categories are shown in Table 7. Class A and B were used as the defined comfort range to maintain a PMV within 0 \pm 0.5.

	Boston, MA	Phoenix, AZ	Atlanta, GA	Seattle, WA	Houston, TX
Climate	Cold	Hot-Dry	Mixed-Humid	Marine	Hot-Humid
	5A	2B	3A	4C	2A
Vintage	<1950s	1970s	1970s	<1950s	1970s
House Size	2589 ft ²	2013 ft ²	2013 ft ²	1938 ft ²	2013 ft ²
Envelope					
Attic	Uninsulated	Ceiling R-13, Vented	Ceiling R-19, Vented	Ceiling R-13, Vented	Ceiling R-13, Vented
Wall Cavity	Uninsulated	Uninsulated	Uninsulated	Uninsulated	Uninsulated
Foundation	Uninsulated	Uninsulated	Uninsulated	Uninsulated	Uninsulated
Windows	Clear, Double, NM, Air	Clear, Double, Metal, Air	Clear, Single, Metal	Clear, Single, Metal	Clear, Double, Metal, Air
Air Leakage	15 ACH50	15 ACH50	15 ACH50	15 ACH50	15 ACH50
HVAC					
Heating	Gas Boiler, 80% AFUE	Gas Furnace, 80% AFUE	Gas Furnace, 80% AFUE	Gas Furnace, 80% AFUE	Gas Furnace, 80% AFUE
Cooling	Room AC, EER 10.7	Central, SEER 13	Central, SEER 13	None	Central, SEER 13

Table 6. Summary of house model constructions.

	Table 7.	Three categories	of thermal comfort	(ISC	EN 7730	, 2005).
--	----------	------------------	--------------------	------	---------	----------

	Thermal State of the Body as a Whole		Operative Tem	perature °C
Category	PPD %	PMV	Summer (0.5 clo)	Winter (1 clo)
А	<6	-0.2 < PMV < +0.2	23.5-25.5	21.0-23.0
В	<10	-0.5 < PMV < +0.5	23.0-26.0	20.0-24.0
С	<15	-0.7 < PMV < +0.7	22.0-27.0	19.0–25.0

While operative temperature, as used in the calculation for PMV, defines thermal comfort, HVAC systems are controlled by measuring zone air temperature. The ambient air temperature setpoints that were used to control the heating and cooling systems in the simulations are given in Table 8.

Table 8. Zone air temperature setpoints and setback temperatures used in building simulations.

	Setpoint Temperature	Setback Temperature
Heating	22.0 °C	18.0 °C
Cooling	24.5 °C	28.0 °C

Three HVAC control scenarios were modeled in the building performance software: conventional operation, occupancy-based HVAC control, and occupancy-based MPC. Conventional operation used a constant heating and cooling setpoint, and was used as the baseline. Results for each strategy are discussed in the subsequent sections.

3.2. Conventional Control (Baseline) Results

When all home scenarios were operated under conventional control, Boston winters showed the highest amount of energy use among all scenarios. This is logical due to the cold ambient environment, older vintage house, and larger size.

In addition to energy use, thermal comfort was evaluated by calculating operative temperatures. Temperatures between 20 °C and 26 °C were considered comfortable to accommodate 0.5 to 1.0 clo clothing levels (typical of a person inside their home). By using a constant setpoint temperature that was within the comfort range, it was expected that the percentage of time temperatures were within the comfort range would be high. Comfort was high except for Phoenix and Seattle in summer. Further examination shows Phoenix achieved a constant internal air temperature of 24.5 °C in summer. However, the operative temperatures were higher, which created uncomfortable conditions. This case demonstrates the impact of using air temperature, rather than operative temperature to drive system controls. In Seattle, both the

indoor air and operative temperature were too high for comfort. High temperatures were caused by the hot outdoor temperatures which could not be mitigated without an air conditioner.

4. Occupancy-Based HVAC Control Results

Occupancy information can be incorporated into HVAC control in a multitude of ways. In all cases simulated in this study, knowledge of building occupancy, either through prediction or detection methods, was used to establish the setpoint temperature for HVAC control. When the space is believed to be vacant, the temperature is allowed to drift to a more relaxed setback temperature, reducing use of the HVAC system when unoccupied. However, when the space is either believed to be occupied or predicted to soon be occupied, the space is maintained at the setpoint temperature, which ensures that temperatures are within the comfort zone.

4.1. Occupancy Control Schemes

Five total occupancy detection models were considered and simulated, which fall under three main strategies:

- 1. **Reactive:** Occupancy is detected and setpoint temperatures are adjusted accordingly. In this case, occupancy is sensed and no prediction is used.
- 2. **Predictive:** Occupancy is predicted using two different non-probabilistic models, as developed in Section 2.3.
 - *Universal model*: This is the prediction model that performed best for all houses and used a one week, 15 min, week/end categorization, moving training set.
 - *Individually tuned model:* This is the prediction model that performed best for the specific house. The models used are listed in Table 4.
- 3. **Hybrid:** A hybrid of predictive and reactive occupancy models. Occupancy is first predicted using the non-probabilistic models developed in Section 2.3. During operation, if an occupancy change from vacant to occupied is detected that was not predicted, the control will react and reset the temperature control to occupied settings. In order to maintain the predictive aspect of the model, this control method does not react to changes from occupied to vacant states, which would have made for purely reactive control.
 - *Universal hybrid:* This is the prediction model that performed best for all houses, and used the same universal model as described above, but with the reactive component.
 - *Individually tuned hybrid:* This is the prediction model performed best for the specific house, with the reactive component. Models used are listed in Table 4.

4.2. Results

Building simulations were conducted for all five occupancy-aware control methods for all homes, climates, and seasons. The energy savings of each method, relative to the conventional baseline model, are shown in Figure 9, with the average savings for that model shown in Table 9. While all of the control strategies reduced the total energy used during the simulation, the two pure predictive models (universal and individually tuned) had the largest energy savings potentials, with 10.9% and 9.6% savings, respectively. The reactive control model has similar energy savings to the prediction models, with an average of 9% of total energy consumption. This method was particularly helpful in homes where long periods of unpredictable vacancy, such as a vacation, occurred. The hybrid approaches, which used both predictive and reactive occupancy, saved the least energy with an average of 3 %–5% savings.



Table 9. Average energy savings by control method.

Figure 9. Energy savings and discomfort percent for simulation period by control method [%].

Of equal importance to energy savings are the comfort of occupants under each control method, which was assessed using building operative temperature. Table 10 shows the percentage of time that the occupants were predicted to be uncomfortable, as classified by time spent occupying the residence when it was not within Class A or B comfort temperatures. The hybrid models achieved the highest levels of comfort, with time spent in discomfort similar to that seen with conventional setpoint control. In contrast, the purely predictive models, which did not react to incorrect predictions, led to the largest discomfort percentages. Figure 9 shows the savings over conventional, compared with the percentage of occupied time in the uncomfortable range.

Control Method	Unmet %
Conventional	2.4%
Reactive	2.6%
Universal Model	7.3%
Individually Tuned	6.9%
Universal Hybrid	2.0%
Individual Hybrid	2.1%

Achieving energy savings without disrupting occupant comfort is the primary goal of effective HVAC control. Since energy savings are achieved by allowing the temperature to drift to uncomfortable conditions when the space is believed to be unoccupied, occupancy-aware controls can only be truly effective if occupancy is accurately predicted and detected. Parameters, such as city, season, and occupancy patterns, all affect comfort, according to the simulations. Phoenix and Seattle have high unmet comfort ratios in summer, as discussed in the conventional control.

The fact that energy savings are dependent on city indicates that either vintage of the home or climate (and likely both) determine relative savings potential. Since home styles vary according to the region, it is difficult to decouple these two effects. In Figure 10, energy savings are shown by city and season simultaneously. Energy savings are highest in Atlanta, Boston, and Houston, especially during summer. All three climates are extremely hot and humid during the summer months, thus reducing unnecessary air-conditioning results in large energy savings.



Figure 10. Unmet comfort and energy savings for simulation period, by city and season [%].

Both energy savings potential and discomfort were affected by occupancy patterns. Houses 2 and 6, which had the highest vacancies, also had the highest median energy savings. This indicates that the higher the vacancy rate, the higher the energy savings potential. This is not surprising, as more unoccupied hours means more opportunities for the temperatures to drift outside of comfort. In the cases of predictive models, the energy savings potential is also dependent on how well the vacancy is predicted. The savings can only be realized when the house is both vacant and correctly predicted to be so. Despite low vacancy rates, House 5, which had the most accurate model, also had the highest energy savings in some simulations. This indicates that the better the prediction model is, the higher the possible energy savings can be, as time spent heating or cooling a house that has been incorrectly predicted as being occupied is reduced. In analyzing comfort, House 2 had the highest discomfort portion and a high prediction inaccuracy, signalling the importance of accurate prediction to comfort. Overall, these figures indicate that although large energy savings are possible, the prediction model needs to be accurately calibrated to achieve energy savings and comfort.

Due to the large ratio of discomfort, prediction-only models were not an effective HVAC control strategy. Both hybrid models were able to achieve comfort at the same level of conventional control. Since occupant comfort is not degraded, these methods are more likely to be used by occupants. Conventional, reactive, and hybrid methods all achieve discomfort below 3% on average. Reactive control has the highest average energy savings at 9.1% but at the cost of reduced comfort in comparison to conventional control. Conventional control, which is the baseline, has no energy savings. The two hybrid controls, universal and individual, achieve an average energy savings of 4.3% and 5.7%, respectively. Individual hybrid is able to achieve the highest energy savings while maintaining comfort levels at or below conventional control. Therefore, the individual hybrid control is the recommended occupancy-based HVAC control.

5. Model Predictive HVAC Control Results

Model predictive control (MPC) was the final HVAC control scheme considered. In MPC, an algorithm is used to predict and proactively react to upcoming temperature disturbances or setpoint changes. MPC has been used in the past to optimize a number of parameters in building control, from incorporating weather forecasts for temperature control to shifting peak loads for the power grid [11,13,15,16,40–50]. MPC has the advantage of being a proactive rather than reactive control strategy. For example, by predicting the effect of an increase in outdoor temperature before it occurs and overheats the space, indoor temperature can be gradually reduced, and the amount of energy used by the HVAC system can be minimized.

In this work, MPC was used in conjunction with weather and the occupancy-prediction models from Section 2.3 to optimize the temperature setpoints. An optimization algorithm was executed to find the setpoint temperature that minimizes both energy use and discomfort. In Section 4, occupancy prediction models were used to change the setpoint temperature. In that case, four total temperature setpoints were allowed: the heating setpoint temperature, the cooling setpoint temperature, the heating setback temperature, and the cooling setback temperature (Table 8). MPC optimization considered not only those four temperatures, but also temperatures within those bounds, to find the optimal solution. In order to run MPC, parameters, such as the optimization algorithm, cost function, the prediction horizon, the execution horizon, and building model, all had to be determined.

5.1. Model

Commonly, MPC is performed utilizing reduced-order linear system models. This allows the optimization to be performed more quickly and easily. However, whole building energy simulations, like EnergyPlus, allow the calculation of radiant heat balances and non-linear part-load system performances, which simplified models cannot capture [44]. In cases where thermal comfort is being evaluated, these calculations are essential, and so EnergyPlus was chosen as the modeling engine to perform the task. However, to reduce computation time, the model was reduced as much as possible. These reductions were achieved by hard-sizing the HVAC equipment using TMY3 data for the simulation period, increasing the simulation timestep from one to fifteen minutes, and reducing the numbers of reported variables [44,51].

5.2. Optimization Parameters

Model predictive control requires a prediction horizon to designate how far into the future the model is predicting and optimizing. In this study, a 24-hour prediction horizon was used to account for diurnal temperature swings and internal gains from daily occupancy patterns and equipment use. The execution period, which dictates how often the optimization is conducted, used a one-hour horizon to adjust to actual occupancy and indoor temperature values. Once an optimum temperature setpoint was found, it was implemented in the model and the simulation via EnergyPlus was stepped forward one hour in time. A new optimization process was then started with the current state values and a new 24-hour prediction horizon.

The MPC utilized a Matlab-based particle swarm optimization algorithm to determine the optimal control actions. Particle swarm optimization (PSO) uses a group of candidate solutions as beginning values. Simulations using these seeds are executed, yielding initial results for the cost function. As the particles are evaluated and move throughout the decision space, they *swarm* towards the optimum solution. By using a swarm, the possibility of finding a local minimum, rather than global minimum, is reduced [52].

5.3. Objective Function

The cost function combines all the factors that are to be evaluated and optimized into a single formula. How the cost function is configured determines which values are considered most important in finding the best solution. In this study, the cost function minimized energy use and occupant discomfort. Optimization of this function was constrained by the allowable temperature band. Discomfort was calculated using predicted mean vote (PMV). Since thermal comfort was not always achieved in the baseline model, the PMV from the baseline was used as the maximum allowed PMV in the optimization run, which prevented the optimization algorithm from penalizing solutions that provided comfort performance equivalent to (or better than) the baseline model. By minimizing energy use and occupant discomfort concurrently, MPC could reduce energy use without sacrificing occupant comfort using Equation (1).

$$\min\left(\sum_{k} E_k + P\right) \tag{1}$$

subject to: $T_{lower,k} \leq T_{optimal,k} \leq T_{upper,k}$,

where *P* is the occupied discomfort, E_k is hourly HVAC consumption at timestep *k*, *T* is the temperature setpoints at timestep *k*, and *k* is the number of timesteps in the evaluation. Occupant discomfort is calculated as shown in Equation (2).

$$P = C \sum_{k} (|PMV_{k}| - PMV_{max}), \qquad (2)$$

where *C* is the comfort penalty slope, PMV_k is predicted mean vote during timestep *k*, and PMV_{max} is the PMV comfort threshold. Since the PMV during the baseline run may exceed 0.5, such as in the case of Seattle in summer, the threshold is adjusted to allow the optimized MPC to use an equitable PMV during optimization. This threshold is calculated with Equation (3).

$$PMV_{max} = \max(0.5, |PMV_{base,k}|), \tag{3}$$

where PMV_{max} takes the higher value between 0.5 and the PMV from the baseline run at timestep k.

The goal of MPC is to optimize for all the factors within the cost function, of which there are two in this case: energy consumption and thermal discomfort. The relative importance of the two factors is controlled by *C*, the comfort penalty slope, which determines the scaling of discomfort costs. With a smaller *C* value, setpoint temperatures leading to uncomfortable hours do not increase the cost function as much, allowing some thermal discomfort to occur in favor of energy savings. Thus, the comfort slope allows flexibility in the cost function and can be tuned to meet the individual goals of the occupant, depending on how much comfort they are willing to sacrifice.

To determine an appropriate comfort slope, MPC simulations were completed for two days in House 1. During the modeled time period, the house was vacant for approximately six hours of each day, allowing the comfort slope to be evaluated during both occupied and unoccupied states. Figure 11 shows the summarized results from this experiment. Discomfort was evaluated by classifying hours within ± 0.5 PMV as comfortable, ± 0.7 PMV as Class C discomfort, and beyond ± 0.7 PMV as excessive discomfort. As the comfort slope value increased, hours of discomfort decreased, while energy use increased. A comfort slope was chosen to allow only a few hours of Class C discomfort. For all simulations in this study, a comfort penalty slope of 1000 was used, which allowed some energy savings, while still maintaining comfort most of the time.

5.4. Simulation

MPC was used to simulate three different scenarios, in order to understand the factors that make it more or less effective. The individually tuned hybrid models found in Section 2.3 were used in two different locations (Atlanta and Houston) in the summer, and the Atlanta results were compared to those found from a model that used perfect occupancy forecasting. Atlanta and Houston were chosen due to their relatively average energy use and temperatures in Section 4. By using a climate that was neither mild nor extreme, the MPC results should apply to more regions.

In each scenario, two different homes' occupancy schedules were utilized in the simulations to show the effects of differing occupant profiles. The three scenarios yielded different results and insights, which are discussed in the following subsections.



Figure 11. Resulting comfort and energy use, by comfort slope value.



In this simulation, MPC was used with the individually tuned hybrid models found in Section 2.3 to model the cooling requirements for a home in Houston in summer. Occupancy profiles for Houses 2 and 5 were chosen, as these exhibited very different patterns. A summary of the EnergyPlus model settings, optimization parameters, and objective function is provided in Table 11. The maximum computation time allowed for each optimization was 30 min, which yielded a minimum of 300 function evaluations for each execution horizon. Overall, the MPC simulation for each house took 60 h to complete the five-day run period.

Parameter	Value
City	Houston
Season	Summer
Houses	2 & 5
Prediction model	Individual Hybrid
Run period	5 days
Timestep	15 min
Planning horizon	24 h
Execution horizon	1 h
Occupied allowed temperatures	22 °C was $\leq T_{optimal,k} \leq$ 24.5 °C
Unoccupied allowed temperatures	$18 \ ^{\circ}\text{C} \leq T_{optimal,k} \leq 28 \ ^{\circ}\text{C}$
Temperature increments	0.5 °C
Comfort penalty slope (C)	1000
Optimization time per execution horizon	30 min

Table 11. Settings used for Model Predictive Control (MPC) optimization in Case 1.

Results for Case 1 are shown in Table 12. Discomfort is measured by exceedance above Class A and B comfort. This value was calculated by summing the operative temperature deviations above 26.0 °C or below 23.0 °C for all occupied hours. Results are measured in Kelvin-hours (Kh). For both houses simulated, the energy saved is very low, with an average savings of 1%. However, little to no thermal discomfort was achieved. The MPC optimization results found that the highest allowed temperatures provided the lowest resulting cost. Although any temperature within the band was allowed, the optimal temperature ending up matching the values used in occupancy-based setpoint control, signaling that for Case 1 setpoint control and MPC optimization yielded the same temperatures.

Table 12. Results for Case

House	Energy Savings	Discomfort
2	2.1%	3.7 Kh
5	0.2%	0 Kh

Most hours of the simulation are within the comfort region, leading to high comfort for the occupants, but low energy savings. Hours in which the temperature was allowed to drift above comfortable temperatures were few due to the small number of hours when the prediction model accurately predicted the house to be vacant.

5.4.2. MPC Case 2A: Atlanta with Occupancy Prediction

In the second set of simulations, MPC was used with the individually-tuned hybrid models for Houses 1 and 2 in Atlanta in the summer. In this case, the maximum allowed temperatures during occupied hours were kept at the same values used during all unoccupied hours. By allowing a large temperature band at all times, the constraints within the cost function were reduced, and temperatures which produced the smallest cost function were used, rather than the being restricted by the temperature band. A summary of all the settings used for the simulation is shown in Table 13.

Results of these simulations are summarized in Table 14. For this simulation, the two houses saved an average of 9.0% in energy use. Exceedance is higher than Case 1 with an average of 35.3 Kh. Total occupied hours for both houses for the week was 185 h, with House 1 being occupied for 100 h and House 2 being occupied for 85 h. With the average exceedance of 35.3 Kh, an average distribution of thermal discomfort would yield 0.4 °C above the ideal temperatures.

Parameter	Value
City	Atlanta
Season	Summer
Houses	1 & 2
Prediction model	Individual Hybrid
Run period	1 week
Timestep	15 min
Planning horizon	24 h
Execution horizon	1 h
Occupied allowed temperatures	$19 ^{\circ}\mathrm{C} \leq T_{optimal,k} \leq 27 ^{\circ}\mathrm{C}$
Unoccupied allowed temperatures	$19 ^{\circ}\mathrm{C} \leq T_{optimal,k} \leq 27 ^{\circ}\mathrm{C}$
Temperature increments	0.5 °C
Comfort penalty slope (C)	1000
Optimization time per execution horizon	30 min

Table 13. Settings used for MPC optimization in Case 2A.

Table	e 14.	Resul	ts for	Case	2A.

House	Energy Savings	Discomfort
1	7.5%	30.8 Kh
2	10.4%	39.8 Kh

Figure 12 shows temperatures for House 1 for two days of the simulation. Chosen temperatures ranged from 19 °C to 27 °C, with the average setpoint temperature at 25.4 °C and 26.2 °C for House 1 and 2, respectively. With the expanded temperature range, temperature values selected did not always conform to setpoint temperatures as seen in Case 1. While setpoint values allowed an 8 °C range, ambient air temperatures occurring within the building had a 4.5 °C to 4.7 °C range. More extreme setpoint temperatures only lasted for an hour, preventing temperature within the building from reaching the setpoint and maintaining a comfortable space despite the setpoints used.



Figure 12. Case 2A temperatures for House 1 using individualized hybrid prediction model.

Figure 13 shows the electricity consumption resulting from the setpoint temperatures used. Due to changing setpoint temperatures, electricity consumption jumped in hours using low setpoint temperatures

as more cooling occurred. In other hours, however, electricity was significantly less than the conventional constant temperature. Over the week simulated, electricity consumption was reduced to allow 7.5% and 10.4% energy savings for House 1 and 2, respectively.



Figure 13. Case 2A electricity consumption for House 1 using individualized hybrid prediction model.

Figure 14 shows the duration curve of operative temperatures for House 1. The figure shows that allowed temperature deviation was higher for unoccupied hours. In hours that the house was occupied, operative temperature was kept closer to the center point temperature of 23.5 °C. With the chosen cost function, some temperature deviation was allowed to achieve higher energy savings. Unlike Case 1, which had tight occupied temperature constraints, temperature deviation in Case 2A is higher. Changes to the comfort penalty slope would change how much deviation is allowed and, in result, how much energy was saved.



Figure 14. Case 2A duration curve of deviation from 24.5 °C operative temperatures.

5.4.3. MPC Case 2B: Atlanta with Perfect Occupancy Forecasting

In a third scenario, all settings used from Case 2A were kept the same except for the occupancy prediction model. In this scenario, actual occupancy data was used in Houses 1 and 2 to imitate perfect occupancy forecasting. This allows for an exploration of how imperfections in the occupancy prediction impact MPC results. A summary of all used settings are shown in Table 15.

Results from the MPC optimization with perfect occupancy predictions are shown in Table 16. Energy savings for Houses 1 and 2 increased by 5.4% and 2.9%, respectively, while comfort exceedance in both homes decreased (Figure 15). Like Case 2A, setpoint temperatures range from 19 °C to 27 °C. However, unlike Case 2A, low setpoint temperatures are used less often to achieve a quick temperature change. The resulting internal air temperature ranged from 22 °C to 27 °C.

Parameter	Value
City	Atlanta
Season	Summer
Houses	1 & 2
Prediction model	Perfect forecasting
Run period	1 week
Timestep	15 min
Planning horizon	24 h
Execution horizon	1 h
Occupied allowed temperatures	$19 ^{\circ}\text{C} \leq T_{optimal,k} \leq 27 ^{\circ}\text{C}$
Unoccupied allowed temperatures	$19 ^{\circ}\mathrm{C} \leq T_{optimal,k} \leq 27 ^{\circ}\mathrm{C}$
Temperature increments	0.5 °C
Comfort penalty slope (C)	1000
Optimization time per execution horizon	30 min

Table 15. Settings used for MPC optimization in Case 2B.

House	Energy Savings	Discomfort
1	12.9%	21.0 Kh
2	13.3%	21.0 Kh

Table 16. Results for Case 2B.

This experiment shows that with an accurate occupancy forecast, MPC optimization is able to allow less energy-intensive temperatures during vacant periods, without the penalty from discomfort when an occupant unexpectedly returns. This allows improvements in both energy savings and thermal comfort. Thus, with accurate occupancy forecasting, energy savings above 10% are possible with the occupancy patterns recorded. Accurate occupancy prediction, therefore, is essential in improving HVAC control.



Figure 15. Case 2B duration curve of deviation from 24.5 °C operative temperatures.

6. Summary and Conclusions

Residential heating and cooling accounts for a large portion of annual energy consumption in the United States. Reducing energy use can contribute to furthering the goals of the Paris Agreement by reducing the burning of fossil fuels and thus reducing CO_2 emissions. Within the residential energy sector, energy savings can be realized by accounting for occupancy in HVAC control. This allows energy reduction without negatively impacting thermal comfort of occupants, which is essential to the widespread adoption of new HVAC control technologies.

A literature review revealed that non-probabilistic models historically have performed best for short-term occupancy prediction. By collecting real occupancy data from six different homes, individual non-probabilistic models were created and evaluated. Prediction inaccuracy in the models, termed state-matching error in the study, ranged from 8.0% to 48.7%. Model training data that used a moving, multi-week training set worked the best for all homes, with differences in occupancy patterns being the highest contributor to prediction inaccuracy. An examination of increased training time indicates that models can improve over time as more data is collected and included into the prediction model. Once the occupancy prediction models were generated, they were then incorporated in occupancy-based setpoint control and occupancy-based model predictive control in a building performance simulation. Five EnergyPlus home models were used to simulate the energy use and indoor temperatures for two-week periods in summer and in winter. Occupancy-based setpoint control showed possible energy savings from 0% to 50.0% over control methods that used a constant setpoint temperature depending on climate, occupancy pattern, and control strategy. Non-probabilistic prediction models achieved the highest energy savings, with an average of 10.0%, but with the disadvantage of high thermal discomfort for the occupants. By including an override, in which the occupancy prediction model can sense the actual occupant presence and react to it, thermal discomfort was reduced. In these hybrid occupancy models, the energy savings averaged 5.0%, while the number of hours that the space was deemed uncomfortable were low.

Model predictive control showed that energy savings is highly dependent on how the cost function and constraints are parameterized. In Case 1, where the temperature constraints were much stricter during occupied hours, little to no energy savings was achieved. However, in Case 2A, where temperature constraints were relaxed during occupied hours, energy savings increased with only a slight impact on discomfort. In Case 2B, where occupancy was perfectly predicted, both energy savings and thermal comfort improved, leading to two conclusions: First, a cost function that combines both energy consumption and thermal discomfort allows for flexibility for the user to determine what trade-off between energy savings and discomfort is appropriate for the them. Second, accurate occupancy prediction improves both performance aspects in the cost function. This shows that, as occupancy prediction improves, the ability for occupancy-aware HVAC control to maintain comfort and increase energy savings improves.

Author Contributions: Conceptualization, G.H.; methodology, G.H. and C.T.; software, C.T. and G.P.; validation, C.T., M.J., G.P. and G.H.; formal analysis, C.T.; investigation, C.T.; resources, M.J.; data curation, C.T. and M.J.; writing–original draft preparation, C.T.; writing–review and editing, C.T., M.J., G.P. and G.H.; visualization, C.T.; supervision, G.H.; project administration, G.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research and the APC was funded by DOE ARPA-E grant number DE-AR0000938.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Tracking Buildings 2020—Analysis. 2020. Available online: https://www.iea.org/reports/tracking-buildings-2020 (accessed on 12 October 2020).
- 2. Pe, L. A review on buildings energy consumption information. Energy Build. 2008, 40, 394–398. [CrossRef]
- Residential Energy Consumption Survey (RECS)-Data—U.S. Energy Information Administration (EIA). 2019. Available online: https://www.eia.gov/consumption/residential/data/2015/ (accessed on 19 September 2019).

- 4. Sun, K.; Hong, T. A simulation approach to estimate energy savings potential of occupant behavior measures. *Energy Build.* **2017**, *136*, 43–62. [CrossRef]
- 5. U.S. Energy Information Administration. *Drivers of U. S. Household Energy Consumption, 1980–2009;* U.S Department of Energy: Washington, DC, USA, 2015.
- 6. Fanger, P.O. *Thermal Confort Analysis and Applications in Environmental Engineering;* Mcgraw-Hill: New York, NY, USA, 1972.
- 7. ASHRAE. ASHRAE Standard 55-2010 Thermal Environmental Conditions for Human Occupancy; ASHRAE Inc.: Atlanta, GA, USA, 2010; Volume 2010, p. 42; ISSN 1041-2336.
- Malinick, T.; Wilairat, N.; Holmes, J.; Perry, L. Destined to Disappoint: Programmable Thermostat Savings are Only as Good as the Assumptions about Their Operating Characteristics. In ACEEE Summer Study on Energy Efficiency in Buildings; American Council for an Energy-Efficient Economy: Washington, DC, USA, 2012; pp. 162–173.
- Nägele, F.; Kasper, T.; Girod, B. Turning up the heat on obsolete thermostats: A simulation-based comparison of intelligent control approaches for residential heating systems. *Renew. Sustain. Energy Rev.* 2017, 75, 1254–1268. [CrossRef]
- 10. U.S. Department of Energy. *Overview of Existing and Future Residential Use Cases for Connected Thermostats;* Technical Report; Energetics Incorporated, Washington, DC, USA, 2016.
- 11. Afram, A.; Janabi-Sharifi, F. Theory and applications of HVAC control systems—A review of model predictive control (MPC). *Build. Environ.* **2014**, *72*, 343–355. [CrossRef]
- 12. Jia, M.; Srinivasan, R.S.; Raheem, A.A. From occupancy to occupant behavior: An analytical survey of data acquisition technologies, modeling methodologies and simulation coupling mechanisms for building energy efficiency. *Renew. Sustain. Energy Rev.* **2017**, *68*, 525–540. [CrossRef]
- Aftab, M.; Chen, C.; Chau, C.K.; Rahwan, T. Automatic HVAC control with real-time occupancy recognition and simulation-guided model predictive control in low-cost embedded system. *Energy Build.* 2017, 154, 141–156. [CrossRef]
- 14. Chen, Z.; Soh, Y.C. Comparing occupancy models and data mining approaches for regular occupancy prediction in commercial buildings. *J. Build. Perform. Simul.* **2017**, *10*, 545–553. [CrossRef]
- 15. Lindelöf, D.; Afshari, H.; Alisafaee, M.; Biswas, J.; Caban, M.; Mocellin, X.; Viaene, J. Field tests of an adaptive, model-predictive heating controller for residential buildings. *Energy Build.* **2015**, *99*, 292–302. [CrossRef]
- Afram, A.; Janabi-Sharifi, F. Supervisory model predictive controller (MPC) for residential HVAC systems: Implementation and experimentation on archetype sustainable house in Toronto. *Energy Build.* 2017, 154, 268–282. [CrossRef]
- 17. Lu, J.; Sookoor, T.; Srinivasan, V.; Gao, G.; Holben, B.; Stankovic, J.; Field, E.; Whitehouse, K. The smart thermostat: using occupancy sensors to save energy in homes. In Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems—SenSys '10, Zurich Switzerland, 2–5 November 2010; p. 211. [CrossRef]
- 18. Kleiminger, W.; Mattern, F.; Santini, S. Predicting household occupancy for smart heating control: A comparative performance analysis of state-of-the-art approaches. *Energy Build.* **2014**, *85*, 493–505. [CrossRef]
- 19. Southern California Edison. *Residential Smart Communicating Thermostat;* Work Paper SCE17HC054 Revision 0; California Public Utilies Commission: San Francisco, CA, USA, 2017.
- 20. O'Brien, W.; Gaetani, I.; Gilani, S.; Carlucci, S.; Hoes, P.J.; Hensen, J. International survey on current occupant modelling approaches in building performance simulation. *J. Build. Perform. Simul.* **2017**, *10*, 653–671. [CrossRef]
- 21. Socolow, R.H. The twin rivers program on energy conservation in housing: Highlights and conclusions. *Energy Build.* **1978**, *1*, 207–242. [CrossRef]
- 22. Al-Mumin, A.; Khattab, O.; Sridhar, G. Occupants' behavior and activity patterns influencing the energy consumption in the Kuwaiti residences. *Energy Build.* **2003**, *35*, 549–559. [CrossRef]
- Andersen, R. The influence of occupants' behaviour on energy consumption investigated in 290 identical dwellings and in 35 apartments. In Proceedings of the Abstract from 10th International Conference on Healthy Buildings, Brisbane, Australia, 8–12 July 2012; pp. 1–3.

- 24. Gaetani, I.; Hoes, P.J.; Hensen, J.L. Occupant behavior in building energy simulation: Towards a fit-for-purpose modeling strategy. *Energy Build*. **2016**, *121*, 188–204. [CrossRef]
- Best Directory Building Energy Software Tools. 2019. Available online: https://www.buildingenergysoftwaretools. com/ (accessed on 19 September 2019).
- 26. Duarte, C.; Van Den Wymelenberg, K.; Rieger, C. Revealing occupancy patterns in an office building through the use of occupancy sensor data. *Energy Build.* **2013**, *67*, 587–595. [CrossRef]
- 27. Page, J.; Robinson, D.; Morel, N.; Scartezzini, J.L. A generalised stochastic model for the simulation of occupant presence. *Energy Build.* **2008**, *40*, 83–98. [CrossRef]
- 28. Tahmasebi, F.; Mahdavi, A. Exploring the Predictive Potential of Probabilistic Occupancy Models. In Proceedings of the BauSIM, Aachen, Germany, 22–24 September 2014. [CrossRef]
- 29. Richardson, I.; Thomson, M.; Infield, D. A high-resolution domestic building occupancy model for energy demand simulations. *Energy Build*. 2008, 40, 1560–1566. [CrossRef]
- D'Oca, S.; Fabi, V.; Corgnati, S.P.; Andersen, R.K. Effect of thermostat and window opening occupant behavior models on energy use in homes. *Build. Simul.* 2014, 7, 683–694. [CrossRef]
- 31. Feng, X.; Yan, D.; Hong, T. Simulation of occupancy in buildings. Energy Build. 2015, 87, 348–359. [CrossRef]
- 32. Hong, T.; Taylor-Lange, S.C.; D'Oca, S.; Yan, D.; Corgnati, S.P. Advances in research and applications of energy-related occupant behavior in buildings. *Energy Build*. **2016**, *116*, 694–702. [CrossRef]
- 33. Li, X.; Yao, R. A machine-learning-based approach to predict residential annual space heating and cooling loads considering occupant behaviour *Energy* **2020**, *212*. [CrossRef]
- 34. Lee, Y.S.; Malkawi, A.M. Simulating multiple occupant behaviors in buildings: An agent-based modeling approach. *Energy Build*. **2014**, *69*, 407–416. [CrossRef]
- 35. Mahdavi, A.; Tahmasebi, F. Predicting people's presence in buildings: An empirically based model performance analysis. *Energy Build.* **2015**, *86*, 349–355. [CrossRef]
- 36. Tahmasebi, F.; Mahdavi, A. The sensitivity of building performance simulation results to the choice of occupants' presence models: A case study. *J. Build. Perform. Simul.* **2017**, *10*, 625–635. [CrossRef]
- 37. Yan, D.; Hong, T.; Dong, B.; Mahdavi, A.; D'Oca, S.; Gaetani, I.; Feng, X. IEA EBC Annex 66: Definition and simulation of occupant behavior in buildings. *Energy Build*. **2017**, *156*, 258–270. [CrossRef]
- 38. Mahdavi, A.; Taheri, M. An ontology for building monitoring. J. Build. Perform. Simul. 2017, 10, 499–508. [CrossRef]
- Speake, A.; Wilson, E.; Salzman, M.; Hasz, A. Cost Benchmarks for Residential 50% Energy Savings Packages; National Renewable Energy Lab: Golden, CO, USA, 2012.
- Oldewurtel, F.; Parisio, A.; Jones, C.N.; Gyalistras, D.; Gwerder, M.; Stauch, V.; Lehmann, B.; Morari, M. Use of model predictive control and weather forecasts for energy efficient building climate control. *Energy Build*. 2012, 45, 15–27. [CrossRef]
- 41. Roy, A.; Bai, L. Control of HVAC System in Residential Buildings using Model Predictive Control: Plant Model Selection. In Proceedings of the IISE Annual Conference, Pittsburgh, PA, USA, 20–23 May 2017; pp. 1538–1544.
- Široký, J.; Oldewurtel, F.; Cigler, J.; Prívara, S. Experimental analysis of model predictive control for an energy efficient building heating system. *Appl. Energy* 2011, *88*, 3079–3087. [CrossRef]
- 43. Bălan, R.; Cooper, J.; Chao, K.M.; Stan, S.; Donca, R. Parameter identification and model based predictive control of temperature inside a house. *Energy Build*. **2011**, *43*, 748–758. [CrossRef]
- 44. Corbin, C.D.; Henze, G.P.; May-Ostendorp, P. A model predictive control optimization environment for real-time commercial building application. *J. Build. Perform. Simul.* **2013**, *6*, 159–174. [CrossRef]
- 45. Pavlak, G.S.; Henze, G.P.; Cushing, V.J. Evaluating synergistic effect of optimally controlling commercial building thermal mass portfolios. *Energy* **2015**, *84*, 161–176. [CrossRef]
- 46. Henze, G.P.; Yuill, D.P.; Coward, A.H. Development of a Model Predictive Controller for Tankless Water Heaters. *HVAC R Res.* 2009, 15, 3–23. [CrossRef]
- 47. Tang, R.; Wang, S.; Xu, L. An MPC-Based Optimal Control Strategy of Active Thermal Storage in Commercial Buildings during Fast Demand Response Events in Smart Grids. *Energy Procedia* **2019**, *158*, 2506–2511. [CrossRef]

- 48. Woo, D.; Junghans, L. Framework for model predictive control (MPC)-based surface condensation prevention for thermo-active building systems (TABS). *Energy Build* **2020**, *215*. [CrossRef]
- 49. Coffey, B.; Haghighat, F.; Morofsky, E.; Kutrowski, E. A software framework for model predictive control with GenOpt. *Energy Build.* **2010**, *42*, 1084–1092. [CrossRef]
- 50. May-Ostendorp, P.; Henze, G.P.; Corbin, C.D.; Rajagopalan, B.; Felsmann, C. Model-predictive control of mixed-mode buildings with rule extraction. *Build. Environ.* **2011**, *46*, 428–437. [CrossRef]
- 51. Lawrence Berkeley National Laboratory; United States Department of Energy. *EnergyPlus Run Time Analysis*; Lawrence Berkeley National Laboratory: Oak Ridge, TN, USA; Environmental Energy Technologies Division: Berkeley, CA, USA, 2008.
- 52. Olsson, A.E. *Particle Swarm Optimization: Theory, Techniques and Applications;* Nova Science Publishers: New York, NY, USA, 2011.

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).