Development and evaluation of occupancy-aware model predictive control for residential building energy efficiency and occupant comfort

C Turley¹, M Jacoby¹, G Henze¹,²,³, G Pavlak⁴

¹ University of Colorado, Boulder, Colorado, United States
² National Renewable Energy Laboratory, Golden, Colorado, United States
³ Renewable and Sustainable Energy Institute, Boulder, Colorado, United States
⁴ Pennsylvania State University, University Park, Pennsylvania, United States

E-mail: christina.turley@colorado.edu, margarite.jacoby@colorado.edu, gregor.henze@colorado.edu, gxp93@psu.edu

Abstract. The residential sector accounts for 25% of global primary energy consumption. Two methods have previously been proposed to reduce residential energy use associated with the provision of occupant thermal comfort: 1. Occupancy-based HVAC control, operating systems only during confirmed occupancy, and 2. model predictive control (MPC), harnessing a mathematical model and forecasts to find optimal operating strategies. Previous studies estimate the average energy savings of the two methods individually in the range of 21% and 16%, respectively. The research presented herein was carried out to evaluate the energy savings potential in residential buildings by combining both approaches across different climates, house vintages, and occupancy patterns. Occupancy and eight different physical modalities (e.g. CO₂ and VOC) data were collected from five homes for time periods of 4–9 weeks. Collected data sets were used to train occupancy prediction models suggested by an extensive literature survey of occupancy model types. The trained prediction models were combined with MPC and detailed EnergyPlus building simulation models to evaluate residential building performance in terms of annual energy savings and thermal comfort, along with discomfort exceedance metrics. Multiple home types and regions were analyzed to understand regional and climate-dependent potential. Based on actual field data, the occupancy models had a prediction inaccuracy between 8% and 35% across the investigated homes. Average occupancy for the collected data ranged from 56% to 86%, a typical range reported in the literature. Building simulations were conducted for three control scenarios: conventional thermostatic control, occupancy-based, and occupancy-based MPC. The results indicate that all advanced strategies improve upon the conventional control, with some scenarios cutting energy use in half with only occasional incurrence of discomfort. The findings indicate that occupancy-aware model predictive residential building control has the potential to drastically reduce energy use and associated emissions while maintaining occupant comfort for both new and existing buildings.

1. Introduction

A finite supply of fossil fuels and the mounting concern of climate change make reducing energy use a global challenge. Buildings are major consumers of energy worldwide, using around 3,060 million tons of oil equivalent (Mtoe) in 2018 [1]. Heating and cooling constitute a large portion of total energy use in buildings. In the United States, heating, ventilation, and air conditioning...
(HVAC) systems account for 50% of building energy consumption [2]. In 2015, U.S. homes used 4,933,000 Terajoules (TJ) for space heating and air conditioning [3]. Reducing this consumption would support the UN Sustainable Development Goals (SDGs) of using technology innovations to improve urban metabolism and energy systems.

Heating and cooling in residential buildings are conventionally controlled by a thermostat with a single setpoint or dual setpoints, which aims to keep the indoor space at a constant temperature or within a setpoint band, regardless of the outside temperature or the occupancy status of the house. Over time, technologies such as programmable thermostats have entered the market to improve comfort control and energy use. As discussed in detail by [4], the adoption of programmable communicating thermostats (PCT) has not led to the deep energy savings as predicted largely due to occupant behavior and willingness to engage with PCTs, finding that nearly half do not use the programming features.

Two control methods, which are explored in this paper, are occupancy-based setpoint control and model predictive control (MPC). In occupancy-based setpoint control, the temperature setpoint is used only when the building is determined to be occupied. The HVAC system uses a setback temperature when the building is vacant, allowing the indoor temperature to float, and thus reducing energy use during unoccupied hours. Previous studies estimate that this method can result in savings ranging between 5 and 23% depending on climate, building vintage, and occupant behavior [5].

In MPC, an online optimization is conducted based on the predicted occupancy and weather forecast to find the HVAC temperature settings that use the least amount of energy while maintaining thermal comfort. During the optimization, the simulation is run repeatedly, guided by a particle swarm optimization algorithm, exploring a large number of allowed setpoint temperatures. The setpoint that uses the least total energy for the simulation period, without increasing discomfort, is chosen and implemented. MPC is re-run at every new time interval, allowing it to adapt to unexpected disturbances. Previous studies estimate occupancy-based MPC can result in energy savings up to 43% for the climate and occupancy patterns reported in [6].

While both control strategies have potential energy benefits, questions still exist on how best to predict occupancy and incorporate it into HVAC control. Because thermal comfort standards should be met during all occupied hours, a good control strategy needs to accurately anticipate occupancy and condition the space in time for the occupants’ arrival. Conditioning the space too soon will result in unnecessary energy usage, so a careful balance must be found.

This paper describes the work performed in generating an optimally tailored occupancy prediction model, which was subsequently used to simulate implementation of the two control strategies mentioned above for a variety of homes and locations. Energy savings and occupant discomfort of the two strategies were compared against conventional controls.

2. Occupancy Data and Prediction Model

Predicting occupancy is a notoriously difficult task. A research topic since the 1970’s, occupant behavior is varied and stochastic [7], and accurately predicting occupancy continues to challenge researchers and industry professionals [8]. Previous studies have shown that model accuracy depends highly on the circumstances such as building type or geography [9]. Occupancy models are usually based on one of three types: Probabilistic, non-probabilistic, and ASHRAE schedules. ASHRAE schemes use a static schedule to predict occupancy, with one daily schedule defining the whole year.

Non-probabilistic models use past occupancy data to determine the historical probability that an occupant is present [10]. The occupancy from the training data is averaged for each timestep in the forecast period. Probabilities with averages above a certain threshold, are predicted to be occupied. Probabilistic models also use historical data. The probability of occupant arrival
Figure 1. Occupancy probability, threshold, and resulting model for single day, by time

and departure is compared against a randomly generated number to predict occupancy creating stochasticity within the model. There is no model that can be argued to work universally better than others [9]. In studies where the three types were directly compared, non-probabilistic and probabilistic models had similar accuracy results in predicting annual occupancy, performing better than ASHRAE schedules [11, 12, 10]. Non-probabilistic models were found to have better short-term prediction performance [10]. Because the occupancy prediction models will be used to control an HVAC system, non-probabilistic models were chosen to generate models with the best possible short-term accuracy.

To generate realistic occupancy models, actual occupancy data was collected from multiple homes over a period of nine months. The occupancy data was gathered from five different homes in Boulder, Colorado, and was collected over a duration of four to nine weeks in each home. The residents were mixtures of students (both undergraduate and graduate) and professionals, and the number of occupants per home varied from one to five. Arrival and departure time in a home was recorded for each resident via a geo-fence application installed on their mobile phone. Paper sign-in sheets were additionally utilized as a back-up and verification for the geo-fence providing ground truth for occupancy. Data for all occupants was compiled into a binary record of occupied and vacant states for the home.

The collected data was used to train and test non-probabilistic occupancy models in order to determine which occupancy models worked best. The first half of the data was used to train the model by averaging the historical occupancy to obtain an occupancy probability. The occupancy probability was calculated for each 24 hour period. The data was then turned into a binary occupancy by setting a threshold. The threshold value was determined by calculating the value that would cause the area under the binary profile to be equal to the area of the aggregated training area. This set the occupied-to-vacant ratio to remain the same between the prediction model and the training data. Values above the threshold were predicted as occupied; values below, unoccupied. Figure 1 shows this process for one day for House 1.

Ninety-six probabilistic models were exhaustively evaluated for each occupancy data set. The different models were differentiated in terms of the historical training data used. Variables considered how days of the week were grouped, how often the data was sampled, the number of weeks of training data used, and if the training data was fixed or moved as a receding horizon. For example, to predict occupancy on a Monday, all the weekdays could be used or only Mondays, data could be sampled at a rate between one minute and one hour, and one to four weeks of historical data could be used.

The occupancy prediction model was evaluated by comparing the predicted occupancy to the actual occupancy confirmed through the ground truth data. Three metrics were calculated:
Figure 2. Parallel category plot of occupancy models

- False negative rate: Percentage of time intervals (minutes) that model incorrectly predicted house was vacant when it was occupied. This can result in thermal discomfort.
- False positive rate: Percentage of time intervals (minutes) that model incorrectly predicted house was occupied when it was vacant. This can result in unnecessary energy use.
- State matching error: Percentage of minutes that model incorrectly predicted occupancy. This is the inaccuracy of the model and the sum of the false negative and false positive rate.

The three metrics were calculated for each model generated. To find the best occupancy model for all the houses surveyed, the state matching error was normalized for each house to remove the effect of house-to-house accuracy errors. This was done by dividing the state-matching error by the lowest state-matching error for that house, ensuring that all of the houses were considered equally. Figure 2 shows a parallel category plot of the results. Occupancy models that were within 5% of the lowest achieved state-matching error for each house are shown in green, higher values are red. For the houses studied, the results showed that categorizing days by weekday/weekend or by day of the week worked well. Using one week or four weeks of training data, which was only available for houses with eight or more weeks of collected occupancy, had both the highest accuracy. A moving training mode, in which the past data is used in a backward looking receding horizon mode, worked best for most houses. Time resolution did not have a large impact.

Two occupancy prediction models were chosen for each house to control the HVAC system: The model that worked the best for that particular house (individually tuned model) and the model that worked best across all houses (universal model). The universal model used a weekend/weekday categorization with a moving one-week, 15-minute resolution training set. By evaluating a universal model and an individually optimized model, the additional energy savings for individually optimizing models could be ascertained. The model performance for the two prediction models are shown for each house in Table 2.

3. Building Models
To model and understand the impact of occupancy-based setpoint control and MPC, five homes were modeled in EnergyPlus 9.1 [13]. Multiple homes, climates, and seasons were simulated to calculate the range of possible scenarios. By simulating multiple building conditions, the results more adequately represent the achievable outcomes. Five prototype homes and weather
BEYOND 2020 – World Sustainable Built Environment conference
IOP Conf. Series: Earth and Environmental Science 588 (2020) 022043
doi:10.1088/1755-1315/588/2/022043

Table 1. Results of occupancy prediction models for each residence

| House Characteristics | Model Inaccuracy (%) |
|-----------------------|-----------------------|
| House # | Avg. Occupancy | Universal Model | Individually Tuned |
| 1       | 86%           | 17%            | 16%             |
| 2       | 56%           | 32%            | 26%             |
| 3       | 75%           | 35%            | 35%             |
| 4       | 82%           | 10%            | 10%             |
| 5       | 81%           | 8%             | 8%              |

files, provided by the National Renewable Energy Laboratory, were used as the building models [14]. These houses are representative of typical U.S. housing stock, with 55% of homes in the United States built before 1980 [3]. Because many of the homes are older with poor insulation and infiltration, large portions of their energy consumption are used in heating and cooling, making the energy savings potential from HVAC control higher. Construction parameters for each model are shown in Table 2.

Table 2. Summary of house model constructions

| Climate          | Boston, MA | Phoenix, AZ | Atlanta, GA | Seattle, WA | Houston, TX |
|------------------|------------|-------------|-------------|-------------|-------------|
| Climate          | Cold       | Hot-Dry     | Mixed-Humid | Marine      | Hot-Humid   |
| Vintage          | 5A         | 2B          | 3A          | 4C          | 2A          |
| House Size       | 2589 ft²  | 2013 ft²    | 1970s       | <1950s      | 1970s       |
| 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          | Double, NM, Air | Double, Metal, Air | Single, Metal | Single, Metal | 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 |

For occupancy-based setpoint control, a two-week period was simulated in each run at a one-minute timestep. The building simulations were run for two weeks in January and two weeks in July to ascertain the impact on both cooling and heating modes. Heating and cooling setpoints (see Table 3) were set to maintain the predicted mean vote (PMV), a commonly used comfort metric, below ± 0.5 [15]. The conventional HVAC control, which was used as the baseline, used the setpoint temperatures to maintain a constant temperature for the entire simulation.

Table 3. Setpoint and setback temperatures used in building simulation

| Setpoint Temperature °C | Setback Temperature °C |
|-------------------------|------------------------|
| Heating 22              | 18                     |
| Cooling 24.5            | 28                     |
4. Occupancy-Based Setpoint Control

Five occupancy-based setpoint control strategies were evaluated. Each method used occupancy to change the HVAC thermostat between the occupied setpoint temperature and the unoccupied setback temperature. The setback temperature allowed a wider temperature range during vacant periods in order to reduce HVAC use. The five control methods were:

- **Reactive**: Occupancy is detected and setpoint temperatures are adjusted accordingly. In this case, occupancy sensing is purely reactive.
- **Universal prediction model**: Occupancy is predicted using the universal prediction model. Model used a one-week, 15-minute, weekend/weekday categorization, moving training set.
- **Individually tuned prediction model**: Occupancy is predicted using the prediction model that performed best for the specific house.
- **Universal hybrid**: Occupancy is predicted using the universal prediction model. 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.
- **Individual hybrid**: Occupancy is predicted using the individually tuned prediction model. 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.

The building simulations were performed for each occupancy-based setpoint control method for all homes, climates, occupancy sets, and seasons. The energy savings of each method relative to the conventional baseline model is shown in Figure 3. All occupancy-based setpoint control strategies reduced the total energy used during the simulation run regardless of climate, house, or season. Higher energy savings were achieved in houses with high vacancy rates and in humid climates that require both latent and sensible heat extraction. The universal and individually tuned prediction model control strategies offered the highest energy savings potential but at the cost of occupant discomfort. The hybrid control methods achieved energy savings while maintaining thermal comfort equivalent to conventional control, suggesting that hybrid control methods are the best strategy to support both energy savings and comfort goals.

5. Occupancy-Based Model Predictive Control

In the second control strategy, the occupancy prediction models were used in conjunction with model predictive control. In this case, the hourly occupancy prediction and weather data were used to find the optimal setpoint temperature trajectory that would yield the lowest energy use while maintaining comfort at or better than the conventional control. Model predictive control operates by minimizing a cost function, in this case a combination of occupant discomfort and

---

**Figure 3.** Energy savings and discomfort percent for simulation period, by control method [\%]
energy use, for a designated prediction horizon. After a solution is found at the current timestep, the first element of the optimal control solution is implemented and the optimization re-run for the next timestep. Model predictive control has the advantage of predicting and responding to disturbances changes, while developing a control strategy in consideration of system dynamics and state and control constraints.

Because the optimization calculations run during MPC are computationally very expensive, not all 250 scenarios completed during the occupancy-based setpoint control analysis were simulated. Two occupancy data sets, House 2 and 5, were simulated for one week in July in Houston. House 2 had the highest vacancy rate, allowing the largest potential reduction in HVAC use while, House 5 had the lowest prediction error, allowing optimization on the most accurate prediction model. To build an accurate thermal history for the building, the preceding week was run during the EnergyPlus simulation to reduce the effect of the warm-up period inaccuracies on temperatures and heat balance [16]. The simulation was run at a 15-minute timestep. Figure 4 shows the optimized setpoint temperature and resulting air temperature for two days in House 2. In this scenario, the optimal setpoint found was at the upper band of allowable temperatures (Table 3) for all but two hours of the week in the first house and five hours in the second.

The electricity used for the house over two days of the week simulated is shown in Figure 5. The conventional control used a constant setpoint temperature of 24.5 °C, while the MPC optimal control used the setpoint temperatures found during the optimization process. Due to the large difference between indoor and outdoor temperatures during daytime, all control strategies experienced a large electricity demand peak in the early afternoon. The MPC method had reduced electricity use when relaxed setpoint temperatures were used and increased electricity use when the setpoint temperature was tightened to a lower value. Over the one week simulated, MPC resulted in 2% energy savings with 0% discomfort. The second house, which also used the upper allowed setpoint temperatures, resulted in less than 1% savings due to the minimal number of vacant hours. This indicates that MPC is a viable method to achieve energy savings while maintaining comfort.

6. Results and Discussion
The occupancy prediction models created for the five houses studied in this paper were able to achieve between 8 and 35% inaccuracy. This leaves a significant potential for improving prediction accuracy. As occupancy prediction is improved, the energy savings and occupant discomfort can be improved. For example, during the MPC simulation for House 2, 14 hours during the week were predicted to be occupied but were actually vacant. Setback temperatures during this time could have been used to increase energy savings. Similarly, when an occupant returns before they are predicted to arrive, the space may not yet be at a comfortable temperature, leading to an increase in occupant dissatisfaction. Improving prediction models will thus improve energy savings and comfort for all occupancy-based control methods.

Both occupancy-based setpoint control and occupancy-based model predictive control were
able to achieve energy savings without decreasing occupant comfort. For the houses studied, methods that minimized occupant discomfort were able to achieve energy savings up to 33%.

In the Houston summer case, MPC found the optimal temperature to be the same as the setback temperature for the majority of the time. Therefore, setback temperatures could provide the same energy and comfort results as the time-intensive optimization. Because timber frame constructed houses common in the United States do not have a large amount of thermal mass, heat transfer across the façade dominate the total heat transfer occurring. The indoor temperature and resulting HVAC use correlate highly with the difference between indoor to outdoor temperature. Unlike concrete or masonry construction, which can absorb and re-radiate heat hours later, the wood and insulation in homes do not retard heat transfer. This reduces the potential benefit of pre-cooling or pre-heating the structure. These results suggest that prescriptive setpoint values could be developed from model predictive control results rather than creating and implementing computationally demanding real-time model predictive control.

7. Conclusion
Incorporating occupancy-based residential building control is a promising solution to increasing energy savings, in turn furthering the UN goals of upgrading infrastructure through increased resource efficiency. Occupancy-based residential building control is a promising technology that can be applied to existing buildings without large retrofit costs. In older, less insulated houses simulated in this study, energy savings could be up to 33%. This approach, applicable to both new and existing building stock, allows for enhanced occupant comfort, reduced costs, and improved energy performance.

References
[1] TCEP: Buildings, Sep 2019. [Online; accessed 19. Sep. 2019].
[2] Luis Pe. A review on buildings energy consumption information. 40:394–398, 2008.
[3] Residential Energy Consumption Survey (RECS) - Data - U.S. Energy Information Administration (EIA), Sep 2019. [Online; accessed 19. Sep. 2019].
[4] Therese Peller, Marco Pritoni, Alan Meier, Cecilia Aragon, and Daniel Perry. How people use thermostats in homes: A review. Building and Environment, 46(12):2529–2541, 2011.
[5] Kaiyu Sun and Tianzhen Hong. A simulation approach to estimate energy savings potential of occupant behavior measures. Energy and Buildings, 136:43–62, 2017.
[6] Florian Nägele, Thomas Kasper, and Bastien Girod. Turning up the heat on obsolete thermostats: A simulation-based comparison of intelligent control approaches for residential heating systems. Renewable and Sustainable Energy Reviews, 75(January):1254–1268, 2017.
[7] Robert H. Socolow. The twin rivers program on energy conservation in housing: Highlights and conclusions. Energy and Buildings, 1(3):207–242, 1978.
[8] Simona D’Oca, Tianzhen Hong, and Jared Langevin. The human dimensions of energy use in buildings: A review. Renewable and Sustainable Energy Reviews, 81(May 2017):731–742, 2018.
[9] Isabella Gaetani, Pieter Jan Hoes, and Jan L.M. Hensen. Occupant behavior in building energy simulation: Towards a fit-for-purpose modeling strategy. Energy and Buildings, 121:188–204, 2016.
[10] Ardeshir Mahdavi and Farhang Tahmasebi. Predicting people’s presence in buildings: An empirically based model performance analysis. Energy and Buildings, 86:349–355, 2015.
[11] Carlos Duarte, Kevin Van Den Wymelenberg, and Craig Rieger. Revealing occupancy patterns in an office building through the use of occupancy sensor data. Energy and Buildings, 67:587–595, 2013.
[12] Farhang Tahmasebi and Ardeshir Mahdavi. The sensitivity of building performance simulation results to the choice of occupants’ presence models: a case study. Journal of Building Performance Simulation, 10(5-6):625–635, 2017.
[13] DOE. EnergyPlus Simulation Software, Version 9.1.
[14] Andrew Speake, Eric Wilson, Madeline Salzman, and Adam Hasz. Cost Benchmarks for Residential 50 % Energy Savings Packages. page 2012, 2012.
[15] P. O. Fanger. Thermal comfort analysis and applications in environmental engineering. Mcgraw-Hill, 1972.
[16] Charles D. Corbin, Gregor P. Henze, and Peter May-Ostendorp. A model predictive control optimization environment for real-time commercial building application. Journal of Building Performance Simulation, 6(3):159–174, 2013.