Purdue University
Purdue e-Pubs

International High Performance Buildings Conference | School of Mechanical Engineering

July 2018

Identifying Peer Groups in a Multifamily Residential Building for Eco-Feedback Design

Sang woo Ham
Purdue University, West Lafayette, IN, USA, ham9@purdue.edu

Panagiota Karava
Purdue University, West Lafayette, IN, USA, pkarava@purdue.edu

Follow this and additional works at: https://docs.lib.purdue.edu/ihpbc

Ham, Sang woo and Karava, Panagiota, "Identifying Peer Groups in a Multifamily Residential Building for Eco-Feedback Design" (2018). *International High Performance Buildings Conference.* Paper 319.
https://docs.lib.purdue.edu/ihpbc/319

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
Complete proceedings may be acquired in print and on CD-ROM directly from the Ray W. Herrick Laboratories at https://engineering.purdue.edu/Herrick/Events/orderlit.html

Identifying Peer Groups in a Multifamily Residential Building for Eco-Feedback Design

Sang woo Ham[1, 2], Panagiota Karava[1, 2*]

[1]School of Civil Engineering, Purdue University, West Lafayette, Indiana, USA

[2]Center for High Performance Buildings, Ray W. Herrick Laboratories, Purdue University, West Lafayette, Indiana, USA

* Corresponding Author

ABSTRACT

Heating and cooling energy consumption in residential buildings is the result of complex interactions of occupant behavior, weather, and building characteristics. Energy saving strategies require residents' participation because they control energy consuming devices and pay the utilities. An effective way to increase this participation is to provide information on their household energy use, potential benefits, and incentives associated with the acceptance of energy-conserving behaviors. Energy comparison in a peer group, often called normative feedback, has been utilized in eco-feedback research to motivate people to reduce their energy use. Due to the variation of building characteristics, even among units in a multifamily residential building that are exposed to the same weather, it is difficult to make a fair comparison of the energy use attributed to behavior. In this paper, we present a Bayesian mixture model that is used to identify building groups with similar thermal characteristics. The model is developed using disaggregated energy use and temperature data from Wi-Fi-enabled power meters and smart thermostats collected in 31 apartments of a multi-unit residential building.

1. INTRODUCTION

Heating and cooling systems for residential buildings account for 20% of the primary energy consumption in the U.S. (U.S. Energy Information Administration, 2018). While various advanced building control and energy management systems for commercial buildings have been developed and implemented in practice, research in the residential sector has been mainly focused on retrofits or energy benchmarks for home improvement and asset rating (Bourassa, Rainer, Mills, & Glickman, 2012). An expensive mechanical system with advanced controls is not a common choice for a residential building, while control interfaces, e.g., thermostats, need to be simple and intuitive so that residents can operate the system without difficulties (Peffer, Pritoni, Meier, Aragon, & Perry, 2011). As a result, sensors, data acquisition systems, and connected-control devices are rarely available in a residential house. However, the recent development of smart devices and people's interests on the devices provide new opportunities for energy management in residential buildings (Ford, Pritoni, Sanguinetti, & Karlin, 2017; Karlin, Ford, et al., 2015).

A major advantage of using smart devices is that those devices do not require a large investment in communication infrastructure for data collection and system control. Many smart devices provide open application programming interface (API) and work through the wireless internet (Wi-Fi). Thus, data acquisition and connected control interfaces can be established without using dedicated software and communication protocols.

In general, residents make decisions on how to use their heating/cooling system by controlling their thermostat according to their comfort preference and utility bills. In this context, two research directions have actively emerged: how to design and provide eco-feedback that could change residents' energy use behavior (Froehlich, Findlater, & Landay, 2010; Karlin, Zinger, & Ford, 2015) and how to automate home energy systems to save energy (Pisharoty, Yang, Newman, & Whitehouse, 2015). Although many field studies have demonstrated promising outcomes (Ayres, Raseman, & Shih, 2009; Rotondo et al., 2016), limitations were also raised with regards to the effectiveness of feedback over long periods of time and the need to consider personalization in the design (Buchanan, Russo, & Anderson, 2015; Khosrowpour, Xie, Taylor, & Hong, 2016). Furthermore, the benefits of home automation enabled by smart thermostats can be realized with improved end-user interaction systems that include eco-feedback (Yang & Newman, 2013; Yang, R., Newman, M. W., & Forlizzi, J., 2014).

Based on the feedback intervention theory (FIT), when feedback directs the locus of the individual's attention to the gap between pre-existing or intervention-provided standards and the current behavior (i.e., *a feedback-standard gap*), the behavior is regulated (Karlin, Zinger, et al., 2015). The standards have various formats such as goal setting, historical comparison, normative comparison, etc., but the realization of behavioral change depends on personal characteristics and motivation, and the feedback information (Buchanan, Russo, & Anderson, 2015). Designing feedback mechanisms that incorporate accurate information with meaningful comparisons is essential for the end-users to develop trust in the system. In the case of feedback based on historical comparison (i.e., self-comparison), the weather impact can be removed using normalization (Energy Star, 2017) or recent data, i.e., from the previous week (Jain, Taylor, & Peschiera, 2012). For normative feedback, the energy consumption per floor area is typically used to make a comparison between residences that have different building characteristics (Dong, Li & Mcfadden, G. 2015). Also, the energy consumption of different buildings with similar floor area in close proximity is used in the case of detached houses (Ayres et al., 2009). In a multifamily residential building, energy consumption has been compared among units on the same floor (Ma, Lin, Li, & Zhou, 2017) or end-users had the option to manually add their peers, e.g. friends (Jain et al., 2012). However, the household units in a multifamily building can have substantially different energy use despite the similarities in some building characteristics, such as geometry, orientation, building envelope, HVAC system, etc. (Rouleau, Gosselin, & Blanchet, 2018).

Previous benchmark studies (Arambula Lara, Pernigotto, Cappelletti, Romagnoni, & Gasparella, 2015; Gao & Malkawi, 2014) suggested the use of k-means clustering to find representative buildings that have a similar relationship between energy consumption and building characteristics. The relationship was characterized using a multivariate regression model, and the parameters of the model were used for the clustering. However, multivariate regression does not include the complex interactions between different building parameters, weather, and human behavior. Also, the spatial and temporal dependencies of parameters cannot be captured when using clustering methods.

Ideally, we can evaluate the true impact of residents' behavior (e.g. setpoint temperature) on energy use by comparing the consumption of different households under the same conditions such as weather and building characteristics, but this is rare in field experiments. In this paper, we use the heat balance equation for a building zone to derive lumped parameters that represent the zone's thermal chracteristics. Then, we use monitored data from smart energy meters and smart thermostats in 31 units of a multifamily residential building to identify peer groups, i.e. units with similar lumped parameters, through a Bayesian mixture model.

2. FIELD STUDY

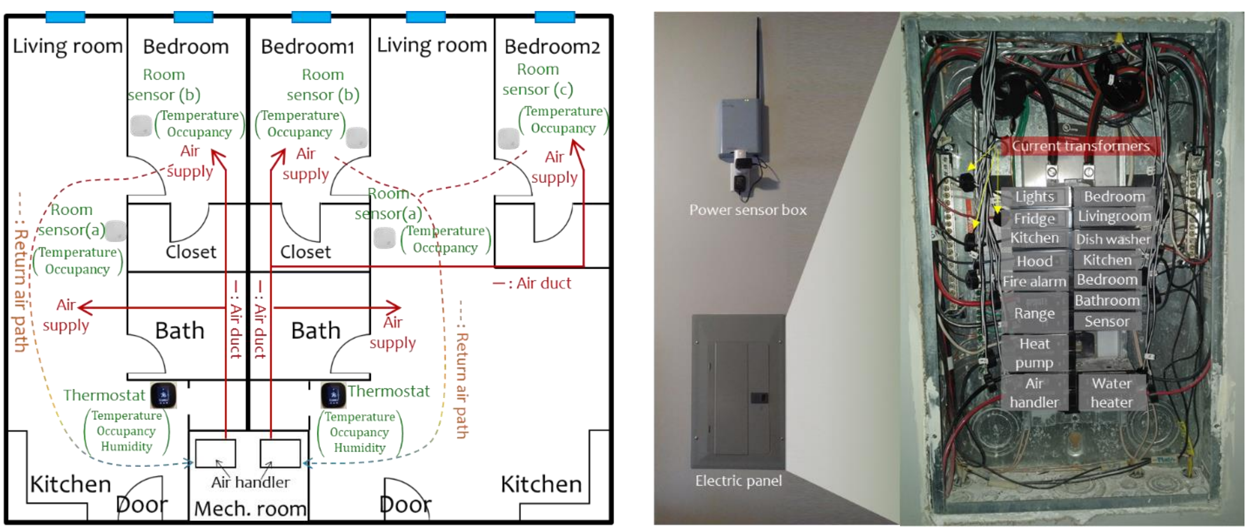
2.1 Building Overview

Our test-bed is a fully-remodeled multifamily residential building, located in Indiana, United States. Construction details are summarized in Table 1. The building has 49 occupied units ($40\times$1-bedroom and $9\times$2-bedroom units) located on the 2nd, 3rd, and 4th floors while one 2-bedroom unit and amenities (multi-purpose room, laundry, PC room, storage rooms) are located on the 1st floor. All building materials were replaced during the remodeling except for the main concrete floor slabs, columns, and the south wall façade. The external walls and inter-unit walls include 6-inch fiberglass insulation (R19). The roof has 5-inch polyiso insulation (R30), and there is no insulation in the main concrete floor slabs. The restored south wall façade is composed of old brick without additional insulation. The apartments are aligned along the west and east side of the building and the units have windows facing west or east. Units on the west side have a balcony with sliding doors in the living room. The balconies on the 2nd floor are located on the ground and look like a backyard (since the 1st floor is underground on the west side) while the balconies on the 3rd and 4th floors are non-protrusion type, and the units have smaller floor area. The units on the east side have operable awning windows in the living room with vinyl frame. Also, units on both east and west side have operable awning windows in bedrooms with vinyl frame.

Each unit is conditioned by a dedicated heat pump system (i.e., indoor air handler unit with auxiliary heating coil and a heat pump outdoor unit). The conditioned air is delivered to all rooms via air ducts and diffusers. The air handler and heat pump water heaters are installed next to the household units in separate mechanical rooms, and the heat pump outdoor units are placed on the rooftop. The unit air handler with a heat pump is controlled by a thermostat, located near the return duct and the entrance door of the unit (Figure 1 (a)). Each unit has typical appliances such as refrigerator, dishwasher, heat pump water heater, range, ceiling lights and fans. The energy consumption in all units is 100% electricity with individual utility meters. 17 units house low-income families that receive utility support.

Table 1: Building overview

Housing type	Fully-remodeled multifamily residential building
Location	Indiana, Unites States
Household unit	49 Units (40 × 1-bed, 9 × 2-bed units) 2nd, 3rd, and 4th floors, 1× 2-bed unit on 1st floor
Bldg. material	• 6-inch Fiberglass for exterior wall and inter-unit wall (R19) • 5-inch Polyiso roof insulation (R30) • One-side old brick south wall façade (4 × 2-bed units are exposed to this wall)
Glazing	• Vinyl frame operable awning window • Sliding door in a balcony
HVAC	• Dedicated air-handler and heat pump with an auxiliary heating coil for each unit • Room air diffusers with air duct • Programmable thermostat • Exhaust fans in restroom and range hood • Conditioned hallway and common spaces
Built-in appliance	Refrigerator, dishwasher, heat pump water heater, range with oven and hood, ceiling light with fan
Utility	All electricity, unit-level billing, 17 units get utility support



1-bedroom unit 2-bedroom unit

(a) Floor plan with sensor locations

Power sensor box and electric panel

Electric panel configuration and current transformers

(b) Power meter and electric panel

Figure 1: Building monitoring details

2.2 Sensor Installation and Data Collection

For the purose of this study, the programmable thermostats were replaced with Wi-Fi-enabled smart thermostats[1] but all the smart features were disabled, and thus, there is no functional difference besides the ability to collect measured data (e.g., temperature, occupancy, etc.) and user's input (e.g., heating and cooling setpoint) via web API. Remote room sensors that measure temperature and proximity (occupancy) were also installed in bedrooms and the living room of each unit. These sensors are used only for data collection purposes and not for setpoint control. In addition, electric current transformers were wired to a sensor box and installed on every circuit inside of the electric panel

[1] Ecobee3 thermostat (https://www.ecobee.com) (accessed 05/01/2018)

(Figure 1 (b)). The power meter[2] can have up to 32 channels. This installation enables dissagregation of the power consumption by device and category as shown in Figure 1 (b). Also, remote power plugs and larger Wi-Fi antennas were installed to improve the Wi-Fi connection and provide the ability to turn on/off the meter remotely.

A Wi-Fi network that already installed for the building security system was leveraged for this study. To reduce Wi-Fi signal interference, we relocated the Wi-Fi APs on the 2nd, 3rd, and 4th floors and modified the Wi-Fi channel allocations. Once the sensors are connected to the network, the data is sent to two cloud servers. The thermostats data is sent to the sensor's cloud server while the power sensors directly send the data to a dedicated cloud server developed for this study using the HTTP GET method with a security token. For thermostats, we extract the data from the cloud server every 5 minutes via API. The power sensors send the data every 30 seconds, and the small-time interval is used to interpolate missing data, while 5-minute data is stored after pre-processing. The study was approved by the Institutional Review Board (IRB Protocol #: 1702018811).

2.3 Observations

In this analysis, data collected in Jan-Feb 2018 from 31 units are used. Figure 2 shows the total energy consumption for each category in the 31 units. The household numbers are encoded for privacy purposes. For better readability, we rearranged the electricity consumption data using seven categories: hvac=energy consumption of the air handler and heat pump; light=energy consumption of all ceiling lights; room1, room2, and living=plug loads of bedroom1, bedroom2, and living room; waterheater=energy consumption of the heat pump water heater; kitchen=energy consumption of dishwasher, range, range hood, and plugs. Due to the cold weather in this period, the highest percentage of energy consumption is used for heating (hvac). Although the units have similar floor area, heating system (though the capacities are different in some units), appliances, etc. and were exposed to the same weather, the differences in their heating energy consumption are large. This is anticipated to different building characteristics and energy-related human behavior, and thus, a direct energy use comparison among units is not suitable for eco-feedback design.

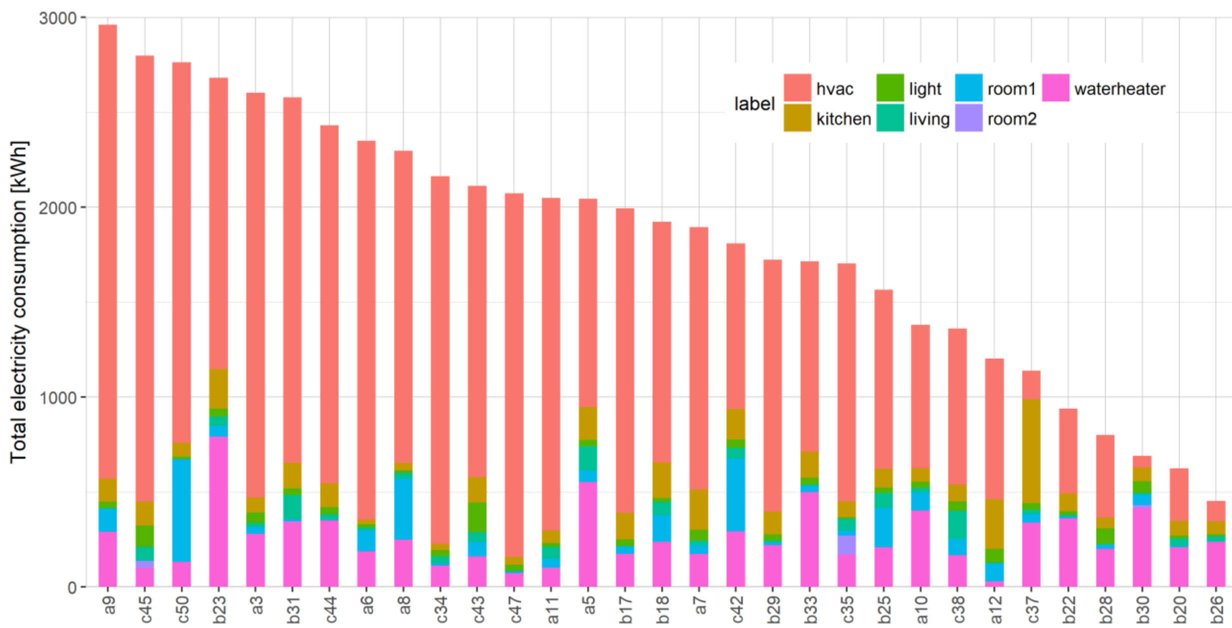


Figure 2: Total electricity consumption (Jan-Feb 2018)

[2] GreenEye Monitor (https://www.brultech.com/greeneye) (accessed 05/01/2018)

3. MODELING

The core idea of peer group identification is to group units with similar thermal characteristics. Using the heat balance equation, we decompose the effect of building characteristics and human behavior to represent non-behavioral thermal characteristics of a building using lumped parameters. The sensible heat balance in a typical building zone can be written as (Mitchell & Braun, 2012):

$$C_h \frac{dT_{\mathrm{m},h,t}}{dt} = \dot{Q}_{\mathrm{ex},h,t} + \dot{Q}_{\mathrm{win},h,t} + \dot{Q}_{\mathrm{in},h,t} + \dot{Q}_{\mathrm{hc},h,t} + \dot{Q}_{\mathrm{vent},h,t} + \dot{Q}_{\mathrm{inf},h,t} + \dot{Q}_{\mathrm{lg},h,t} + \dot{Q}_{\mathrm{ig},h,t} \tag{1}$$

where $T_{\mathrm{m},h,t}$ is the temperature of a thermostat (m) in a household unit (h) at time (t); C_h is lumped capacitance of a building zone (i.e., household unit); $\dot{Q}_{\mathrm{ex},h,t}$ and $\dot{Q}_{\mathrm{win},h,t}$ is the rate of heat transfer to the zone air through the exterior surfaces (ex) and glazing (win), including the affect of solar radiation modelled using the sol-air temperature for simplicity; $\dot{Q}_{\mathrm{in},h,t}$ is rate of heat transfer from neighboring spaces (e.g., household units, storages, amenities) to the zone air through internal partition walls. $\dot{Q}_{\mathrm{hc},h,t}$ is the rate of heat transfer from the heating system and portable heater (if any); $\dot{Q}_{\mathrm{vent},h,t}$ and $\dot{Q}_{\mathrm{inf},h,t}$ is the rate of heat transfer due to ventilation and infiltration; $\dot{Q}_{\mathrm{lg},h,t}$ is the heat gain from measurable large-load appliances such as range; $\dot{Q}_{\mathrm{ig},h,t}$ is the heat gain from occupants and small appliances.

The terms in Eq. (1) can be divided into two categories; weather and building characteristics-related ($\dot{Q}_{\mathrm{ex},h,t}$, $\dot{Q}_{\mathrm{win},h,t}$, $\dot{Q}_{\mathrm{in},h,t}$, and $\dot{Q}_{\mathrm{inf},h,t}$) and behavior-related ($\dot{Q}_{\mathrm{hc},h,t}$, $\dot{Q}_{\mathrm{vent},h,t}$, $\dot{Q}_{\mathrm{lg},h,t}$, and $\dot{Q}_{\mathrm{ig},h,t}$). The models for these terms consist of linearized equations of weather variables (i.e., dry bulb temperature, etc.), building material properties, and indoor temperatures. For example, the rate of heat transfer through the exterior wall is a product of the overall heat transfer coefficient, wall area, and the difference between the sol-air temperature and indoor air temperature. All units are exposed to the same weather and weather variables have temporal dependency. The building material properties are assumed to be space-dependent since different units have different exterior wall, partition wall, and glazing area. Different units have different indoor temperatures so this parameter is time- and space-dependent.

There are two types of behavior-related terms: measurable and non-measurable. For the measurable terms, we can estimate $\dot{Q}_{\mathrm{hc},h,t}$ and $\dot{Q}_{\mathrm{lg},h,t}$ by considering the coefficient of performance of the heat pump and heat fraction from the measured power data. However, in a field experiment, the measurement of $\dot{Q}_{\mathrm{vent},h,t}$ and $\dot{Q}_{\mathrm{ig},h,t}$ is not available in general since we need to monitor all human actions such as windows opening, ventilation fan control, appliance usage, and human occupancy, along with the associated heat gains. When the heat gains from the non-measurable terms in a household unit are quantified in short time intervals (e.g., daily, etc.), their differences can be significant due to schedule variations. However, when the heat gains are integrated for a week, the differences would be reduced since people usually have weekly periodic schedules. Thus, the non-measurable terms can be treated as a unit-specific stochastic noise when considering a weekly time-step. On the other hand, the integration of the left-hand-side term ($C_h\, dT_{\mathrm{m},h,t}/dt$) of Eq. (1) would be small compared to other terms because the increase and decrease of temperature is offset for a period of one week.

The design of normative feedback requires comparison of the energy use among different units in a building at the same period, so variables that only depend on time such as weather can be treated as time-specific constants. Summarizing all information provided above, we can rewrite Eq. (1) by parameterizing the material properties and time-dependent variables (weather), integrating the linear equations in the weekly interval, treating non-measurable terms as stochastic noise, and rearranging all terms in the linear equations with respect to measurable variables (i.e., $\dot{Q}_{\mathrm{hc},h,t} + \dot{Q}_{\mathrm{lg},h,t}$ and $T_{\mathrm{m},h,t}$):

$$y_{h,week} = \beta_{0,h,week} + \beta_{1,h,week} x_{h,week} + \varepsilon_{h,week} \tag{2}$$

where $y_{h,week}$ is the average $\dot{Q}_{\mathrm{hc},h,t} + \dot{Q}_{\mathrm{lg},h,t}$ value of a household for a week; $\beta_{0,h,week}$ is a lumped parameter that includes the effect of weather, building material properties, and temperature of neighboring spaces; $\beta_{1,h,week}$ is a lumped parameter that includes building material properties and infiltration terms that are related to $T_{\mathrm{m},h,t}$, and depends on both space and time (due to the infiltration); $x_{h,week}$ is the average $T_{\mathrm{m},h,t}$ value of a household for a week; $\varepsilon_{h,week}$ is stochastic noise term that represents unobserved heat gains from $\dot{Q}_{\mathrm{vent},h,t}$ and $\dot{Q}_{\mathrm{ig},h,t}$.

In Eq. (2), the two lumped parameters ($\beta_{0,h,week}$ and $\beta_{1,h,week}$) determine the unit's overall thermal characteristics in a specific week because the stochastic human behaviors are treated as noise parameter, and the two lumped parameters represent the relationship between indoor temperature and heat supply. In other words, different units that have similar overall thermal characteristics would have similar lumped parameters, and we can express the parameters using a group assignment (i.e., $\beta_{0,g_h,week}$ and $\beta_{1,g_h,week}$), where, g_h is the group assignment for a unit (h). Now, the group identification becomes a mixture of linear regression problem, which is solved using a Bayesian approach (Eq. (3)) to include the spatial (between units) and temporal (within a unit) characteristics of lumped parameters.

$$
\begin{aligned}
\boldsymbol{\alpha}_{1..K} &\sim \text{Gamma(shape=1, rate=1)} \\
\boldsymbol{\pi}_h &\sim \text{Dirichlet}(\boldsymbol{\alpha}_{1..K} + \mathbf{1}) \\
g_h &\sim \text{Categorical}(\boldsymbol{\pi}_h) \\
\beta_{0,g_h,week} &\sim \text{Normal}(\mu=0, \sigma^2=3) \\
\boldsymbol{\mu}_{\beta_1,1..K} &\sim \text{Normal}(\mu=0, \sigma^2=3) \\
\beta_{1,g_h,week} &\sim \text{log-Normal}(\mu=\mu_{\beta_1,g_h}, \sigma^2=1) \\
\sigma_h^2 &\sim \text{Inv-Gamma(shape=0.1, rate=0.1)} \\
\mu_{h,week} &= \beta_{0,g_h,week} + \beta_{0,g_h,week} x_{h,week} \\
y_{h,week} &\sim \text{Normal}(\mu=\mu_{h,week}, \sigma^2=\sigma_h^2)
\end{aligned}
\tag{3}
$$

where $\boldsymbol{\alpha}_{1..K}$ is a prior for the Dirichlet distribution; K is number of groups; $\boldsymbol{\pi}_h$ is a prior for the distribution of group assignment; μ_{β_1,g_h} is prior of log-Normal distribution of $\beta_{1,g_h,week}$; σ_h^2 is unit-specific variance of $y_{h,week}$, which is the square of $\varepsilon_{h,week}$.

The intercept parameter ($\beta_{0,g_h,week}$) includes both weather and building characteristics, so a time- and group-independent prior is used. The slope term ($\beta_{1,g_h,week}$) mostly consists of time-independent building characteristics but is also slightly affected by the weather due to infiltration. Thus, the slope term is less sensitive to time variation since it is governed by building characteristics such as the *U*-value, and we can use a shared hyper-prior (μ_{β_1,g_h}) to consider this effect. Also, the building characteristics are material properties, which are positive, so we use log-Normal distribution. Finally, as we discussed earlier, the unobserved heat gains can be treated as unit-specific noise, and a unit-specific variance term (σ_h^2) is used.

All the measured data ($y_{h,week}$ and $x_{h,week}$) is standardized (i.e., mean=0, variance=1) during MCMC. In this model, we need to calculate the total heat supply from the heating device ($\dot{Q}_{\text{hc},h,t}$), i.e. the heat pump system and portable heater, and the heat gain from high-load appliances such as range ($\dot{Q}_{\text{lg},h,t}$). Since we do not measure the heat supply from the heat pump system, we estimate it from the power consumption data using information from the manufacturer's catalog for the coefficient of performance (COP) as a function of the outdoor air temperature. For the auxiliary heating coil and portable heater, a coefficient of 0.9 is used for the duct losses and efficiency. For the heat gain from range, a load fraction of 0.4 is used to estimate the sensible heat gain (Wilson, Metzger, Horowitz, & Hendron, 2014).

All parameters are estimated using the Markov-chain Monte Carlo (MCMC) method. Specifically, the JAGS (Plummer, 2003) sampler is used. The model is selected based on widely applicable information criteria (WAIC) and leave-one-out cross-validation (LOO) score (Vehtari, Gelman, & Gabry, 2016) that are used to compare different models. Both metrics are derived from the prediction accuracy using a fully Bayesian approach.

$$
\text{lppd}(y_{h,week}) = \log \Pr\left(y_{h,week} | \beta_{0,g_h,week}^{(s)}, \beta_{1,g_h,week}^{(s)}, x_{h,week}, \sigma_h^{(s)}\right) \tag{4}
$$

To calculate the score, the log-pointwise predictive density (lppd) needs to be calculated (Eq. 4), where superscript (s) indicates MCMC samples. From the model, we compared WAIC and LOO scores of different numbers of groups from 3 to 6. When the number of groups (k) is 3, we obtain the best scores for both metrics.

4. RESULTS

Figure 3 shows the group identification results obtained using data from 31 units for 9 weeks. Group numbers have been assigned to units based on the most frequent group values from MCMC samples. In each graph, the relationship between the weekly average unit temperature ($T_{\mathrm{m},h,t}$) and weekly average rate of heat transfer from the heat pump system and portable heater ($\dot{Q}_{\mathrm{hc},h,t}$), and large appliances ($\dot{Q}_{\mathrm{lg},h,t}$) is shown (i.e., the relationship between $x_{h,week}$ and $y_{h,week}$ in Eq. (2)). Three groups are identified, which are encoded as group 1 (red), 2 (green), and 3 (blue) in Figure 3. Each dot indicates each household unit, and the corresponding unit identifiers are marked next to the dots.

As discussed above, the overall thermal characteristics in a unit are represented by two lumped parameters: the intercept ($\beta_{0,g_h,week}$) and slope ($\beta_{1,g_h,week}$). The model identified groups have similar lumped parameters. When the average room temperature is similar, the average rate of heat transfer from the heating device ($\dot{Q}_{\mathrm{hc},h,t}$) plus the heat gain from large appliances ($\dot{Q}_{\mathrm{lg},h,t}$) is larger in group 1 compared to 2 and 3, and the rate in group 2 larger than 3.

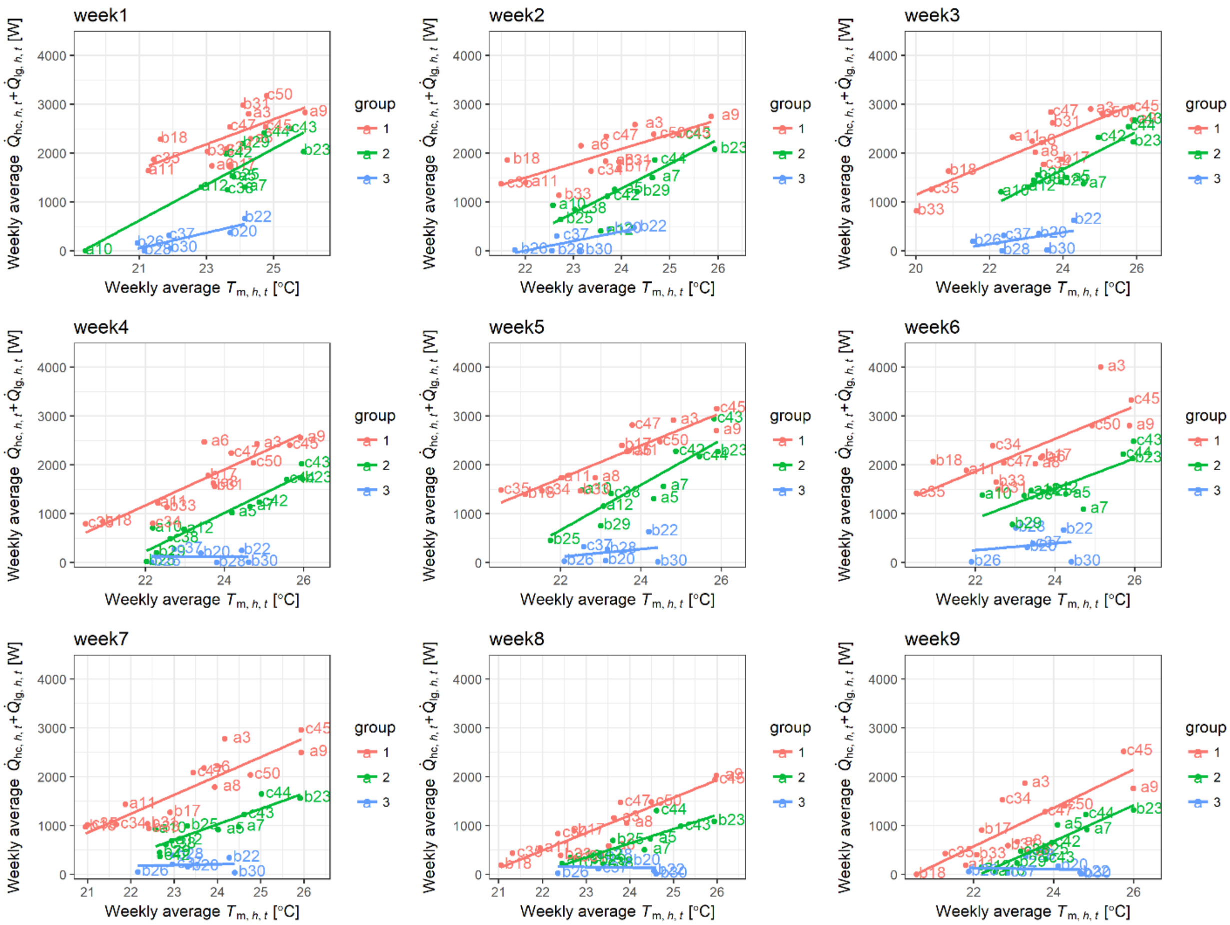


Figure 3: Group identification results

Figure 4 shows a comparison of the estimated (sampled) intercept ($\beta_{0,g_h,week}$) values and outdoor air temperature. The intercept parameter is multiplied by -1 for better visualization and comparison. The input data have been normalized, so the absolute scales of the estimated parameter values do not have physical meaning. As expected, the increase/decrease of intercept parameter is similar to the variation of the outdoor air temperature because the weather-related information is included in this term although the magnitudes are different among groups. This observation implies that the intercept parameter successfully captures the outdoor air temperature information.

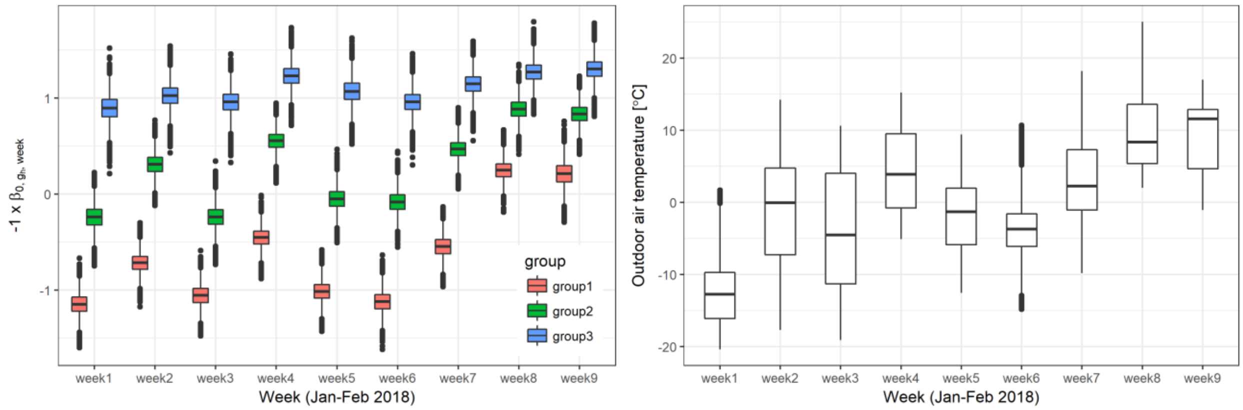


Figure 4: Estimated (sampled) intercept ($\beta_{0,g_h,week}$) values and outdoor air temperature variation

5. CONCLUSIONS

In this study, we presented a new approach to identify peer groups that have similar building thermal characteristics using data from thermostats and energy meters collected during the heating period in 31 apartments in a multifamily residential building located in Indiana. We derived a simple linear model with two lumped parameters that represent the zone thermal characteristics. The model is expressed as a mixture of linear regression and solved through a Bayesian approach to take the spatial and temporal variations of the lumped parameters into account. In the future, we will extend this work to design feedback mechanisms that motivate and incentivize energy-conserving behaviors towards energy-aware residential communities.

REFERENCES

Arambula Lara, R., Pernigotto, G., Cappelletti, F., Romagnoni, P., & Gasparella, A. (2015). Energy audit of schools by means of cluster analysis. *Energy and Buildings*, *95*, 160–171. https://doi.org/10.1016/j.enbuild.2015.03.036

Ayres, I., Raseman, S., & Shih, A. (2009). Evidence from Two Large Field Experiments that Peer Comparison Feedback Can Reduce Residential Energy Usage. *National Bureau of Economic Research Working Paper Series*, *No. 15386*. https://doi.org/10.3386/w15386

Bourassa, N. J., Rainer, L., Mills, E., & Glickman, J. (2012). The Home Energy Scoring Tool: A Simplified Asset Rating for Single Family Homes. *2012 ACEEE Summer Study on Energy Efficiency in Buildings*, (12), 34–46.

Buchanan, K., Russo, R., & Anderson, B. (2015). The question of energy reduction: The problem(s) with feedback. *Energy Policy*, *77*, 89–96. https://doi.org/10.1016/j.enpol.2014.12.008

Dong, B., Li, Z., & Mcfadden, G. (2015). An investigation on energy-related occupancy behavior for low-income residential buildings. Science and Technology for the Built Environment, 21(6), 892–901. https://doi.org/10.1080/23744731.2015.1040321

Energy Star. (2017). Technical Reference: Climate and Weather. Retrieved from https://www.energystar.gov/sites/default/files/tools/Climate_and_Weather_August_2017_EN_508.pdf

Ford, R., Pritoni, M., Sanguinetti, A., & Karlin, B. (2017). Categories and functionality of smart home technology for energy management. *Building and Environment*, *123*, 543–554. https://doi.org/10.1016/j.buildenv.2017.07.020

Froehlich, J., Findlater, L., & Landay, J. (2010). The Design of Eco-feedback Technology. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1999–2008). New York, NY, USA: ACM. https://doi.org/10.1145/1753326.1753629

Gao, X., & Malkawi, A. (2014). A new methodology for building energy performance benchmarking: An approach based on intelligent clustering algorithm. *Energy and Buildings*, *84*, 607–616. https://doi.org/10.1016/j.enbuild.2014.08.030

Jain, R. K., Taylor, J. E., & Peschiera, G. (2012). Assessing eco-feedback interface usage and design to drive energy efficiency in buildings. *Energy and Buildings*, *48*, 8–17. https://doi.org/10.1016/j.enbuild.2011.12.033

Karlin, B., Ford, R., Sanguinetti, A., Gannon, J., Rajukumar, M., & Donnelly, K. (2015). Characterization and

Potential of Home Energy Management (HEM) Technology. San Francisco, CA: Pacific Gas and Electric.

Karlin, B., Zinger, J. F., & Ford, R. (2015). The effects of feedback on energy conservation: A meta-analysis. *Psychological Bulletin*. Karlin, Beth: University of California, Irvine, 5548 Social & Behavioral Science Gateway, Irvine, CA, US, 92697-7075, bkarlin@uci.edu: American Psychological Association. https://doi.org/10.1037/a0039650

Khosrowpour, A., Xie, Y., Taylor, J. E., & Hong, Y. (2016). One size does not fit all: Establishing the need for targeted eco-feedback. *Applied Energy*, *184*, 523–530. https://doi.org/10.1016/j.apenergy.2016.10.036

Ma, G., Lin, J., Li, N., & Zhou, J. (2017). Cross-cultural assessment of the effectiveness of eco-feedback in building energy conservation. *Energy and Buildings*, *134*, 329–338. https://doi.org/10.1016/j.enbuild.2016.11.008

Mitchell, J. W., & Braun, J. E. (2012). *Principles of Heating, Ventilation, and Air Conditioning in Buildings* (1st ed.). Wiley.

Peffer, T., Pritoni, M., Meier, A., Aragon, C., & Perry, D. (2011). How people use thermostats in homes: A review. *Building and Environment*, *46*(12), 2529–2541. https://doi.org/10.1016/j.buildenv.2011.06.002

Pisharoty, D., Yang, R., Newman, M. W., & Whitehouse, K. (2015). ThermoCoach: Reducing Home Energy Consumption with Personalized Thermostat Recommendations. In *Proceedings of the 2Nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments* (pp. 201–210). New York, NY, USA: ACM. https://doi.org/10.1145/2821650.2821671

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd international workshop on distributed statistical computing* (Vol. 124, p. 125). Vienna, Austria.

Rotondo, J., Johnson, R., Gonzalez, N., Waranowski, A., Badger, C., Lange, N., … Foster, R. (2016). Overview of Existing and Future Residential Use Cases for Connected Thermostats. Report prepared by Energetics Incorporated and Vermont Energy Investment Corporation (VEIC) for the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy Building Technologies Office.

Rouleau, J., Gosselin, L., & Blanchet, P. (2018). Understanding energy consumption in high-performance social housing buildings: A case study from Canada. *Energy*, *145*, 677–690. https://doi.org/10.1016/j.energy.2017.12.107

U.S. Energy Information Administration. (2018). Annual Energy Outlook 2018 with projections to 2050. U.S. Energy Information Administration.

Vehtari, A., Gelman, A., & Gabry, J. (2016). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 1–20. https://doi.org/10.1007/s11222-016-9696-4

Wilson, E., Metzger, C. E., Horowitz, S., & Hendron, R. (2014). 2014 Building America House Simulation Protocols. National Renewable Energy Laboratory.

Yang, R., & Newman, M. W. (2013). Learning from a learning thermostat: Lessons for Intelligent Systems for the Home. *UbiComp 2013 - Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. https://doi.org/10.1145/2493432.2493489

Yang, R., Newman, M. W., & Forlizzi, J. (2014). Making Sustainability Sustainable: Challenges in the Design of Eco-interaction Technologies. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 823–832). New York, NY, USA: ACM. https://doi.org/10.1145/2556288.2557380

ACKNOWLEDGEMENT

This work was funded by the National Science Foundation under Grant No. 1737591 and the Big Ideas Challenge program at Purdue University. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.