Integrating occupants' voluntary thermal preference responses into personalized thermal control in office buildings

Seungjae Lee¹,², Panagiota Karava¹,²*, Athanasios Tzempelikos¹,², Ilias Bilionis³

¹ Lyles School of Civil Engineering, Purdue University, 550 Stadium Mall Dr., West Lafayette, IN 47907, USA
² Center for High Performance Buildings, Ray W. Herrick Laboratories, Purdue University, 140 S. Martin Jischke Dr., West Lafayette, IN 47907, USA
³ School of Mechanical Engineering, Purdue University, 585 Purdue Mall, West Lafayette, IN 47907, USA

* Corresponding author. E-mail: pkarava@purdue.edu

Abstract. A Bayesian modeling approach which allows incorporating voluntary feedback data (comfort-related responses), collected via participatory interfaces, along with requested feedback data, into a thermal preference learning framework. This is achieved by explicitly considering occupant participation, a type of behavior, in the model. Experiments with human subjects were conducted to collect thermal preference datasets, with both participatory and requested setups, which were used to train personalized thermal preference models. The proposed approach allows using the participatory setup without distorting the thermal preference predictive probabilities. In addition, we propose a concept of smart occupant feedback request algorithm, that determines whether and when to request feedback based on the quantified value of the request. This work will lead to smarter, user-interactive comfort delivery systems that will be continuously updated through interactions with their occupants, and will provide customized indoor environments tailored to individual preferences.

1. Introduction

Many studies have shown that general thermal comfort models cannot accurately predict thermal preferences of individuals. In addition, typical HVAC systems operation based on general thermal comfort criteria often results in conservative control settings and high energy use, while it cannot achieve high levels of satisfaction for all occupants. Incorporating occupants in sensing and control frameworks (a.k.a., human-in-the-loop) has been proposed as potential solution, along with using feedback (comfort-related responses) from individual occupants to learn and update their thermal preference and provide customized indoor environments.

The most direct way to obtain comfort-related responses (data to train personalized comfort models) is requesting their responses through surveys. In [1–8], the authors requested occupants to submit responses regularly or randomly to collect data for learning thermal preferences and develop probabilistic models. However, frequent response requests are intrusive and thus not very practical in real buildings. Alternatively, occupants can voluntarily report their state of thermal comfort, through participatory user-interfaces, whenever they want [9–12]. Although the method seems practical for real world applications, occupants’ participation, a type of behavior, could be affected by multiple factors; consequently, the response distribution in collected data could be heavily biased. For example, people would actively participate and provide their feedback if they are uncomfortable believing that the
feedback could improve the thermal conditions. On the contrary, if they are satisfied with the conditions, they would not be as responsive. In that case, researchers [10,12] assumed that occupants were satisfied if there was no participation for a while. However, it is possible that occupants may not (or forget to) participate even though they are not satisfied because of their workload [12]. Also, there are other factors, e.g., the ease of interface, the expectation that the system will adapt to their response, making it difficult to develop generalizable heuristic rules for participatory data collection [11].

In this study, we present a Bayesian modeling approach which allows incorporating voluntary feedback data (comfort-related responses), collected via participatory interfaces, along with requested feedback data, into the thermal preference learning framework. The incorporation is done by explicitly considering occupants’ participation—a type of behavior—in the model. In addition, we propose a concept of smart occupant feedback request algorithm, that determines whether and when to request feedback based on the quantified value of the request.

2. Methodology

2.1. Experimental study

To collect data for the study, an experiment was conducted with five human test-subjects using three identical perimeter south-facing private offices in West Lafayette, Indiana (Fig. 1). The subjects performed their usual workload (computer-related work, reading, writing, etc.) during the experiment and wore similar clothes every time they participated. In each room, one monitor was designated for a user-interface (Fig. 2) that the subjects used to report their thermal preference by selecting one of the following three answers: “I prefer warmer” / “I prefer cooler” / “I am satisfied with current thermal conditions.”

![Figure 1. Exterior and interior view of the offices.](image1)

![Figure 2. Interface.](image2)

In the experiment, two different experimental setups were used to collect two datasets from each person: requested and participatory. In the “requested” setup, the interface was activated if the subject stayed in the room for 30 minutes and the subjects were requested to report their thermal preference. Once they submitted a response, the 30-minute timer was reset, the interface became deactivated, and the room air temperature setpoint was changed to the next value in one of the pre-determined schedules listed in Table 1. If the subject left for a short break, the timer was reset when they came back, but the setpoint was maintained. The subjects were asked to report their thermal preference six times before and after lunch. Each subject participated in this setup for four days and experienced all the schedules.

![Table 1. Setpoint schedules (°C) applied in the requested setup.](table1)

The incorporation is done by explicitly considering occupants’ participation—a type of behavior—in the model. In addition, we propose a concept of smart occupant feedback request algorithm, that determines whether and when to request feedback based on the quantified value of the request.
changes, responses submitted within 20 minutes after entering the room or after the previous response were ignored. The subjects were asked to stay in the room for three hours before and after their lunch. Different initial temperature setpoints (i.e., 25, 22.5, 20 °C) were applied at the beginning of each session to avoid bias as much as possible. Each subject participated for three days (six sessions) in this setup and experienced each initial temperature setpoint twice.

2.2. Personalized thermal preference models using requested and participatory data

Two different models were developed with the requested and participatory setups and datasets respectively (Fig. 3). For the requested preference data, we use the linear ordered probit model considering the obvious order between the thermal preference $Y$ (Model 1 in Fig. 3). The air temperature is the only input of the model, i.e., aggregated effects of all the unconsidered factors, $H_L$, are assumed to follow a Normal distribution. We assume that there is a latent quantity representing the thermal state of and mind, $L$. If this quantity is below a certain threshold, $-\tau$, one will answer “prefer warmer”; if it is above a certain threshold, $\tau$, one will answer “prefer cooler”; and in between one will answer “satisfied” (Fig. 4). However, since we do not explicitly consider all the factors affecting the thermal state of one’s body and mind, we assume $p(L|T, \Theta) = N(L|f(T), \lambda)$, where $f(t) = \beta_0 + \beta_1 t, N(\mu, \sigma^2)$ is the Probability Distribution Function (PDF) of a normal distribution with mean $\mu$ and variance $\sigma^2$, and $\Theta$ collects all the model parameters including $\beta_0$ and $\beta_1$. We set $\lambda = 1$. Subsequently, the probability of $Y$ is modeled as:

$$p(y = 1|t, \Theta) = \Phi(-\tau - f(t))$$
$$p(y = 0|t, \Theta) = \Phi(\tau - f(t)) - \Phi(-\tau - f(t))$$
$$p(y = -1|t, \Theta) = 1 - \Phi(\tau - f(t))$$

where $\Phi(\cdot)$ is the Cumulative Distribution Function (CDF) of the standard normal distribution and $y = -1, 0, 1$ represent “prefer cooler,” “satisfied,” and “prefer warmer,” respectively.

For the participatory preference data, Model 2 in Fig. 3 is used. Here, relationships between $T, L, Y$, and $H_L$ are the same as described above. $U$ is a latent quantity which represents one’s satisfaction level. We assume $U = g(L) = -|L|$, i.e., the magnitude of $L$ determines the level of satisfaction. We explicitly consider one’s participation (i.e., a type of behavior) in the model using parameters $V_1$ (participation triggered by one’s dissatisfaction/discomfort) and $V_2$ (participation due to other reasons). More specifically, if one is not satisfied with the current thermal condition, $y = -1$ or $1$, and $U + H_{V_2}$ is smaller than a threshold, $\delta$, where $H_{V_2}$ is aggregated effect of unconsidered factors on $V_2$, then the person will participate, $V_1 = 1$, and a response $y$ is received. If $U + H_{V_2}$ is higher than the threshold, although the person is not satisfied, no response is obtained. $H_{V_2}$ follows a Normal distribution of which the standard deviation is $\sigma_{V_i}$. If the person is satisfied with the current condition, $y = 0$, then $V_1 = 0$. Subsequently, the probability of $V_1$ is modeled as:

$$p(V_1 = 0|U, \Theta) = \begin{cases} 1, & y = 0 \\ 1 - \Phi(\frac{\delta - U}{\sigma_{V_2}}), & y = -1 or 1 \end{cases}$$
$$p(V_1 = 1|U, \Theta) = \begin{cases} 0, & y = 0 \\ \Phi(\frac{\delta - U}{\sigma_{V_2}}), & y = -1 or 1 \end{cases}$$

For participation triggered by other factors, $p(V_2 = 0|\Theta) = q$, $p(V_2 = 1|\Theta) = 1 - q$.

To train the personalized thermal preference models in Bayesian way, Sequential Monte Carlo (SMC) [13] was used in PyMC3 python package [14].

![Figure 3. Model structures.](image1.png)

![Figure 4. Application of the linear ordered probit model.](image2.png)
2.3. Smart feedback request algorithm: quantifying the value of information gain

With Model 2 in Section 2.2, we can learn one’s thermal preference using both requested and participatory preference responses. Since frequent response requests are intrusive, requesting regularly or randomly is not efficient nor effective. Ideally, we would mainly rely on participatory preference data and request responses only when they are truly needed. Here, the question is how to quantify the value (importance) of requesting a response at a given moment (under specific conditions). The value of information gain at any given condition also varies with our current state of knowledge regarding the person’s thermal preference. For example, if the current temperature is 30 °C, and if we are already quite certain that the person will prefer a cooler temperature, an additional requested response at that moment will not improve our state of knowledge—therefore, the request value is low and we do not need to disturb the person.

The value can be quantified with the expected improvement (reduction) in the model uncertainty as we add a datapoint. At any given moment, there are seven possible events. If we ask one’s preference, we will receive one of three responses: “prefer warmer,” “satisfied,” “prefer cooler.” If we do not request, we will receive either one of the three (participatory) responses or no response. This means that there are seven corresponding possible models as the current model is updated with the new datapoint. The expected model uncertainty after the update with/without the request can be computed as: \( \mathbb{E}(H_r) = \sum_{y=-1}^1 p(y|r) h_{r,y} \), where \( r \in \{0,1\} \) indicates whether there is a request or not and \( h_{r,y} \) estimates the model uncertainty of the corresponding updated model. Since we do not know \( p(y|r) \), we use the predictive probability from the current model, \( p(y|r, \text{Data}) \), instead. Here, \( \text{Data} \) refers to the data used to train the current model. If the expected model uncertainty of the updated model with a request is significantly lower than that without the request, \( \mathbb{E}(H_r) \ll \mathbb{E}(H_0) \), we would request a response. In this study, we use the entropy [15,16] of the posterior parameter probability distribution as the measure of the model uncertainty. To estimate the posterior parameter probability distribution with posterior samples from SMC, we use Kernel Density Estimation (KDE) function in Scipy package [17] with Gaussian Kernel and Scott’s rule for the bandwidth. To estimate the entropy with the estimated posterior distribution, we use Leave-One-Out (LOO) cross-validation method.

3. Results

3.1. Experimental Results

Fig. 5 shows data collected from the five test-subjects in the “requested” and “participatory” setups. The portion of “satisfied” responses is significantly lower in the participatory preference data. This observation corresponds to one of our hypotheses: “Occupants would rarely report their thermal preference using a participatory interface under comfortable conditions”. The data also show the differences in thermal preferences between individuals for both setups.

![Figure 5. Thermal preference responses from five subjects (request and participatory setups).](image)

3.2. Personalized Thermal Preference Results

To realize the difference between the requested and participatory datasets and their impact on personalized preference models, three personalized models were developed and compared for each subject: (i) Model 1 trained with the requested dataset; (ii) Model 1 trained with the participatory dataset; (iii) Model 2 trained with the participatory dataset. Each graph in Fig. 6 shows the predictive probability distribution with respect to air temperature computed from the personalized preference models. Solid lines represent the median values of the probabilities, and shaded areas represent the associated 95% credible intervals (i.e., 2.5- and 97.5-quantiles of the predictive probability).
Comparing graphs in the first and second rows (Model 1 trained with one’s requested and participatory datasets, respectively), we can see that the predictive probability of models developed with Model 1, with participatory data, can be highly distorted. More specifically, the probability of the person being satisfied decreases significantly due to the lack of “satisfied” responses, i.e., the green curves are shrunk for Subjects 1-4. Even in the case of Subject 5, who submitted quite a few “satisfied” responses, the probability curves moved to the right. Consequently, using this model with participatory data only is not effective for proper calculation of mostly preferred temperatures, which would have an impact on both occupant satisfaction and energy use (control operation). The graphs in the third rows of Fig. 6 show the predictive probability of models trained with Model 2 and the participatory datasets. The uncertainty areas overlap and cover the graphs of the first row, while the probability distribution curves are not distorted. This is important for avoiding wrong predictions in preferred temperatures, but also for estimating the value (importance) of information gain (a new data point) with and without a response request at a moment.

3.3. Value of Information Gain With/Without a Request
With Model 2, we can quantify the value of a response request in terms of the improvement in the model (epistemic) uncertainty (Fig. 7). The left graph shows the difference in the model uncertainty of the updated model with and without a request under different temperature conditions. Subject 1’s preference model (corresponding to one at the bottom-left corner of Fig. 6) is used as the current model. Bigger entropy difference represents higher value of the request (e.g., for 22 °C or 25 °C) meaning a request is meaningful. Then, Subject 1’s preference model is updated with seven preference data points collected between 21.5-22.5 °C in the request setup. The right graph shows the difference in the expected entropy computed for this case (e.g., significant reduction for 22 °C).

4. Conclusion
A Bayesian modeling approach was developed which allows using both participatory preference data (collected via participatory interfaces) and requested preference data for learning individual occupants’ thermal preference. This is achieved by explicitly considering occupant participation, a type of behavior,
in the model. An experiment was conducted with two different setups to collect occupants’ thermal preference responses (with both ways), which were used to develop personalized thermal preference models. The results showed that the proposed modeling approach properly maintains the uncertainty with the participatory interface as is, instead of returning distorted predictive probabilities. In addition, we propose a smart occupant feedback request algorithm to determine whether and when to request feedback, based on the quantified value of the request. The importance of the request value was computed using the expected improvement in the model uncertainty after the update. Based on our models and the smart request algorithm, we expect that smarter, user-interactive comfort delivery systems will provide customized indoor thermal environments.

Acknowledgments
This work was funded by the National Science Foundation under Grant No. 1539527 and supported by the Center of High Performance Buildings at Purdue University.

References
[1] Liu W, Lian Z and Zhao B 2007 A neural network evaluation model for individual thermal comfort Energy Build. 39 1115–22
[2] Daum D, Haldi F and Morel N 2011 A personalized measure of thermal comfort for building controls Build. Environ. 46 3–11
[3] Zhao Q, Zhao Y, Wang F, Wang J, Jiang Y and Zhang F 2014 A data-driven method to describe the personalized dynamic thermal comfort in ordinary office environment: From model to application Build. Environ. 72 309–18
[4] Chen X, Wang Q and Srebric J 2015 A data-driven state-space model of indoor thermal sensation using occupant feedback for low-energy buildings Energy Build. 91 187–98
[5] Jiang L and Yao R 2016 Modelling personal thermal sensations using C-Support Vector Classification (C-SVC) algorithm Build. Environ. 99 98–106
[6] Kim J, Zhou Y, Schiavon S, Raferty P and Brager G 2018 Personal comfort models: Predicting individuals’ thermal preference using occupant heating and cooling behavior and machine learning Build. Environ. 129 96–106
[7] Lee S, Karava P, Tzempelikos A and Bilionis I 2019 Inference of thermal preference profiles for personalized thermal environments with actual building occupants Build. Environ. 148 714–29
[8] Lee S, Joe J, Karava P, Bilionis I and Tzempelikos A 2019 Implementation of a self-tuned HVAC controller to satisfy occupant thermal preferences and optimize energy use Energy Build. 194 301–16
[9] Feldmeier M and Paradiso J A 2010 Personalized HVAC control system Internet of Things (IOT), 2010 pp 1–8
[10] Erickson V L and Cerpa A E 2012 Thermovote: participatory sensing for efficient building hvac conditioning Proceedings of the Fourth ACMWorkshop on Embedded Sensing Systems for Energy-Efficiency in Buildings (BuildSys’12) (New York, NY, USA: ACM) pp 9–16
[11] Jazizadeh F, Ghahramani A, Becerik-Gerber B, Kichaylo T and Orosz M 2014 Human-Building Interaction Framework for Personalized Thermal Comfort-Driven Systems in Office Buildings J. Comput. Civ. Eng. 28 2–16
[12] Li D, Menassa C C and Kamat V R 2017 Personalized human comfort in indoor building environments under diverse conditioning modes Build. Environ. 126 304–17
[13] Bilionis I and Koutsourelakis P S 2012 Free energy computations by minimization of Kullback–Leibler divergence: An efficient adaptive biasing potential method for sparse representations J. Comput. Phys. 231 3849–70
[14] Salvatier J, Wiecki T V. and Fonnesbeck C 2016 Probabilistic programming in Python using PyMC3 PeerJ Comput. Sci. 2 e55
[15] C. E. Shannon 1948 A Mathematical Theory of Communication Bell Syst. Tech. J.
[16] Jaynes E T 1957 Information Theory and Statistical Mechanics Phys. Rev. 106 620–30
[17] Millman K J, Aivazis M 2011 Python for Scientists & Engineers Comput. Sci. Eng. 13 9–12