A data-driven approach for predicting occupant thermal comfort in offices

Fatma Deghim[1], Farzan Banihashemi[1], Sebastian Koth[2] and Werner Lang[1]

[1]Institute of Energy Efficient and Sustainable Design and Building, TU Munich, Munich, Germany

[2]Chair of Building Technology and Climate Responsive Design, TU Munich, Munich, Germany

E-mail(s): [fatma.deghim, farzan.banihashemi, sebastian.koth, w.lang]@tum.de

Abstract: Since people spend most of their time in indoor environments, working on thermal comfort is of increasing interest to researchers. Indoor thermal comfort is currently modeled using knowledge-based methods such as the predicted mean vote (PMV) model. The PMV model was developed by averaging the feedback of large groups of people on thermal comfort in laboratory studies. Hence the model is limited when predicting the thermal comfort of individuals. Researchers investigated new approaches to model individual-specific thermal comfort responses that showed high potential such as the Personal Comfort Models (PCM). In this paper, we evaluated different machine learning (ML) classifiers predicting personal thermal comfort in offices. Our final model performed 30% better than PMV in predicting thermal sensation votes (TSV) on the ASHRAE Global Thermal Comfort Database II. We performed a case study to evaluate the trained models, where we obtained comparable F1-scores between the PMV model and the trained models. This result on the case study is explained by the mild environment in the conducted experiment and the imbalanced data with high presence of the neutral thermal sensation vote. This work exemplifies the potential and the limitations of using classifiers trained on the ASHRAE dataset to predict TSV on a real use case.

Keywords: Thermal sensation, group-based personal comfort modeling, data-driven, PMV, ASHRAE Global Thermal Comfort Database II

1 Introduction

Generally Heating, Ventilation and Air Conditioning (HVAC) systems are used to ensure thermal comfort in indoor environments based on standards such as ASHRAE 55 [1]. With predefined setpoints, current HVAC systems provide a "one-size fits all" environment, thereby ignoring that comfort is subjective and can differ from one person to another depending on several environmental and personal factors. Apart from the high energy consumption to provide this defined comfortable environment, studies showed that people do not necessarily feel comfortable at the end [2]. Thus, with the emergence of the Internet of Things (IoT), the importance of developing intelligent HVAC systems that correctly predict the thermal sensation of individuals increased. These systems will learn from users real-time feedback and adapt the environmental parameters of the room to make it more comfortable accordingly. A better understanding of occupants' behavior in indoor environments will help fill the performance gaps of the PMV and adaptive models and provide better individual indoor conditions that are not only comfortable but also healthy [3], [4].

Several researchers have used data-driven methods to predict individuals' feedback on thermal comfort. This approach is called personal comfort modeling (PCM) and was discussed and reviewed in detail by Kim [5] and Martins [6]. The models use real-time sensor data combined with personal information and individuals' feedback obtained from surveys. The objective and subjective data are used to train ML models in predicting individual' thermal sensations and preferences. The sensor data typically include indoor and outdoor environmental parameters: indoor and outdoor air temperature, indoor mean radiant temperature, relative humidity, and air velocity [5], [6]. Certain studies like the done by Lee (2021) [7] used physiological parameters, including skin temperature from different body parts and heart rate using thermal cameras or wearable devices. Other studies started using data from Personal Comfort Systems (PCS) to learn individual thermal preferences[8]. Additional to sensor data, personal information such as age, sex, clothing, and metabolic rate were used to build the models using ML.

Various ML algorithms have been evaluated with the goal of improving thermal sensation predictions. Prominent examples include: Random Forest [9], [10], Classification Trees [9], Gaussian process classification [7], Gradient Boosting Method [11], Support Vector Machine [10]–[12], K-Nearest Neighbors [10], [11], and Artificial Neural Networks [10]. Although previous studies showed the high potential of the PCMs in predicting thermal comfort of individuals especially compared to the PMV or adaptive method, these Blackbox models lack transparency when it comes to evaluating them on real use cases or reproducing them. In fact, in previous research, mainly data from own experiments was used to develop PCMs using ML [7], [9], [10]. Others used the ASHRAE database to create thermal comfort models but did not test them on real use cases for example [12]. Any change in the data, be it only the combination of input parameters or using a completely different training data set implies adaptations to the model properties and therefore its capabilities. This research paper aims to answer the following research question: Can data-driven models trained on an open-source dataset accurately predict the thermal sensation of new individuals compared to conventional methods?

2 Methodology

The proposed approach for predicting individuals' TSV includes three steps: (1) data preprocessing of the ASHRAE dataset, (2) model development using ML algorithms, and (3) model evaluation on collected data from experiments at the Technical University of Munich, as seen in Figure 1.

2.1 Data preprocessing

The open-source database ASHRAE Global Thermal Comfort Database II was launched in 2014. This database comprises 81,846 rows of raw data divided into measured environmental parameters

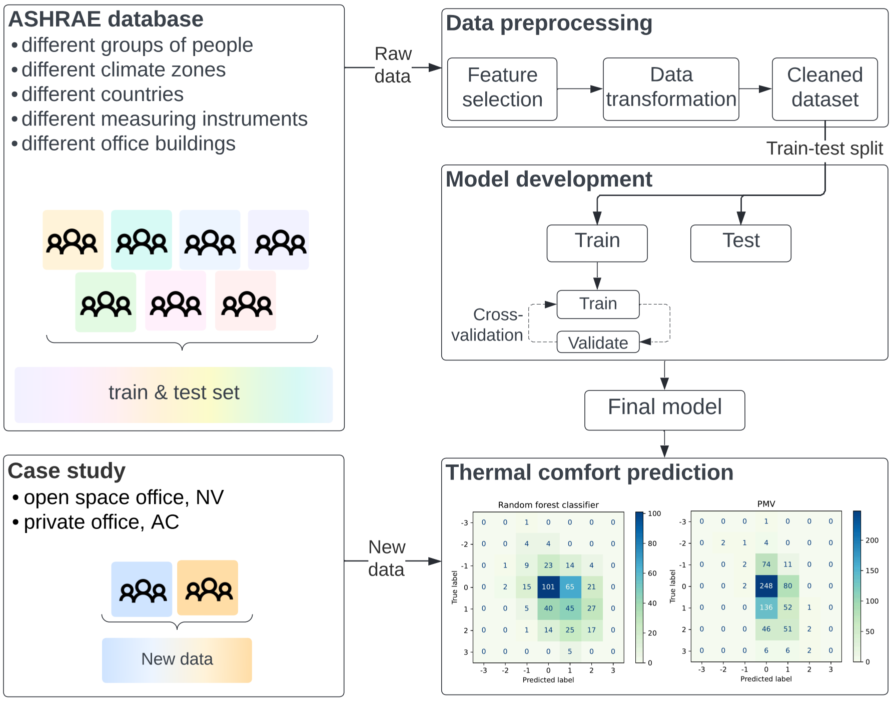

Figure 1: Proposed methodology

and subjective user feedback collected between 1995 and 2016 on different continents according to predefined data collection requirements. The database furthermore comprises several "building categories" such as classrooms, multi-family houses, offices and "building types" that concern cooling strategies [13]. Within the framework of this paper, the data was filtered to only include data from offices which results in a raw dataset with 61,691 rows.

Data preparation: First, the features vector x and the target vector y are selected. As stated before commonly three types of parameters were used for PCM development: (1) indoor environmental features, (2) outdoor environmental features, and (3) personal features. The ASHRAE dataset only provides the monthly outdoor temperature. Hence, it was excluded from model training. In this study, we use additional features such as the country, the climate zone, the season, and the building type. The target value is the individual's Thermal Sensation Votes (TSV), which is an occupant's feedback on thermal sensation at the moment of the survey and is given on a 7-point scale corresponding to the categories: -3 (cold), -2 (cool), -1 (slightly cool), 0 (neutral), 1 (slightly warm), 2 (hot), 3 (warm). The rows with null values and errors in the target column are removed. Moreover, the climate zones are grouped to the five main groups according to Köppen climate classification: A tropical, B dry, C temperate and D continental. The E polar group was absent in the dataset. Figure 2 depicts the TSV distribution in the cleaned data set. **Encoding**: The selected features are a mixture of numerical and categorical data. Since most of ML algorithms can not handle categorical values, these must be converted to numerical values. The preprocessing method from python scikit-learn "One hot encoding" is used for this purpose. **Data scaling**: Feature values in the dataset show wide ranges, for example, indoor air temperature ranges between 13.4 and 37.6 ℃ and air velocity ranges between 0.01 and 0.5 m/s, which can lower the model performance and its learning speed. Feature scaling is a good practice in this case. The most common techniques of feature scaling are Normalization and Standardization [14]. In the realm of this work standardization was used. **Data resampling:** Figure 2 shows the class imbalance in the dataset where the classes -1 and 1 represent the majority of the TSVs. To deal with the problem of data imbalance, researchers combined classes together and worked with a 3 or 5-point scale TSV [9]. In this study, the 7-point scale TSV is kept to consider the nuances in an individual's TSV. The “RepeatedStratifiedKFold” cross-validation with 10 folds and 3 repetitions as well as other algorithm-specific hyperparameters are used to overcome the problem of imbalanced data. After data preparation, the obtained dataset had a total of 9,070 valid rows further summarized in Table 1.

Table 1: Features used for model training (ASHRAE) and evaluation (case study)

Feature type	Feature name	Range of values (ASHRAE)	Range of values (Case study)
Indoor environment	Indoor air temperature (℃)	[13.4, 37.6]	[19.51, 30.8]
	Air velocity (m/s)	[0.01, 0.51]	[0, 0.39]
	Relative humidity (%)	[15.5, 88.8]	[23.84, 74.47]
Personal information	Age	[10, 75]	[25, 39]
	Sex	Female, male	Female, male
	Clothing insulation (clo)	[0.3, 2.2]	[0.23, 1.75]
	Metabolic equivalent of task (Met)	[0.8, 6.8]	[1.1, 5.3]
Additional features	Climate zones	A, B, C	C
	Country	Australia, Brazil, Iran, India, Germany	Germany
	Season	Summer, autumn, winter, spring	Winter, spring
	Building type	Naturally ventilated (NV), air-conditioned (AC), mixed mode (MM)	NV, AC

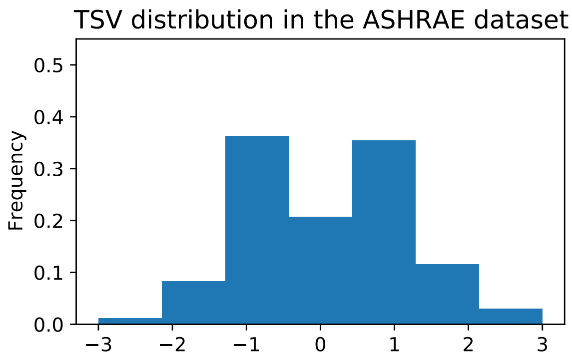

Figure 2: Label distribution (ASHRAE dataset)

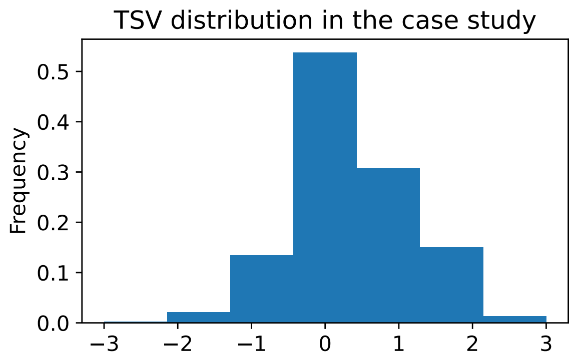

Figure 3: Label distribution (case study)

2.2 Model development

The dataset is split in an 80% train and a 20% test set using stratified train-test split for imbalanced data. Four ML classifiers were trained: Support Vector Machine (SVM), Random Forest (RF), Gradient Boosting (GB) and K-Nearest Neighbors (KNN). The hyperparameters were tuned using gridSearchCV from python scikit-learn and the BayesianOptimazationOracle Tuner in Keras.Tuner [15] and showed comparable results with the Keras.Tuner being computationally more effective. The best parameter search was conducted using the Bayesian optimization with a Gaussian Process model with defining the maximum number of trials. The RepeatedStratifiedKFold was used to evaluate the algorithms during hyperparameter tuning with 10 folds and 3 repetitions. The hyperparameter tuning was done with the goal of maximizing the mean f1-score across all folds and all runs. This cross-validation method has, compared to simple cross-validation, the advantage of keeping the same percentage of each class in every iteration by stratification and the cross-validation being repeated, allows a more accurate evaluation of the algorithm. After hyperparameter tuning, the following models were developed using the ASHRAE dataset. The RF was trained with 300 trees and a max_depth of 16 using a balanced subsample. The GB was trained with 200 boosting stages, a max_depth of 6 and a learning rate of 0.01. The KNN was trained with 26 nighbors and leaf_size 25 and using the distance weights methods for points. The SVM was trained using C=10, an 'rbf' kernel and a gamma of 0.1.

2.3 Thermal comfort prediction - case study

The case study consists of two field studies conducted in two types of offices: one naturally ventilated open space office and one air-conditioned single office. For the first experiment, a survey was conducted at the Institute of Energy Efficient and Sustainable Design and Building during the winter and spring 2022. For three weeks in February and in May, 12 healthy individuals, 8 females and 4 males, participated in the survey. The participants answered questions about their (1) clothing, (2) activity level in the last 30 minutes before taking the survey, (3) current thermal sensation and (4) current thermal preference. The participants were asked to take the survey two to three times per week and three times a day: in the morning, before lunch break, and in the afternoon. For the second experiment, the survey was conducted in the Senselab from the Chair of Building Technology and Climate Responsive Design, Prof. Dipl-Ing. Thomas Auer, TUM [16] over six weeks in October and November 2021. 10 healthy participants, 6 females and 4 males, participated in the survey and answered the same questions as in the first experiment several times a day. Each participant did the experiment for two days in the survey period. In addition to the survey data, the case recorded the office environmental parameters with measurement boxes situated near to the participants. Indoor air temperature, relative humidity, and air velocity were recorded at the time of the survey. The clothing insulation and the metabolic rate of each participant were calculated according to the ASHRAE 55 [1]. The metabolic rate was calculated by averaging the activity levels in the last 30 minutes before the survey.

3 Results and discussion

In this study we developed ML models to predict the TSV of a group of individuals based on the ASHRAE database. These models were then evaluated on previously unseen data in a case study. Additionally, the PMV index is computed using pythermalcomfort package [17] according to the ASHRAE standard [1] to compare the ML models and the conventional comfort model. All six parameters for PMV calculation (indoor air temperature, mean radiant temperature, relative humidity, air velocity, metabolic rate, and clothing insulation) are known. The calculated PMV indices are converted to integers between -3 and 3 to allow a precise comparison to the personal TSVs [12]. Table 2 summarizes the performance of the proposed classifiers as well as the PMV on the ASHRAE test set and the case study. In general, the four classifiers performed similarly in predicting the TSV on the test set and showed better results than the PMV with F1-scores of 0.49 and 0.47 compared to 0.17. The models were further evaluated on the unseen data of the case study. For this specific case, where the indoor environment was relatively mild and people tend to feel neutral, the PMV showed comparable F1-score and accuracy to the ML models. The RF, GB and the PMV were the top three best models with F1-scores of 0.37 and 0.39. The two-dimentional confusion matrices of these three models in Figures 4 and 5 show the actual class (rows) and prediction value (column). Due to data imbalance and despite the different techniques used to overcome this problem, we can see that the labels -1, 0 and 1 are the most predicted classes for the trained models followed by the class 2. These classes are in fact the most present classes in the training set 2, whereas the PMV model predicted mainly the classes 0 and 1 with higher correct prediction for the class 0 which represented more than half of the TSVs in the case study (Figure 3). 80% of the values of the class 0 were correctly predicted by the PMV but only for this class the PMV showed this high true positive prediction. Cheung [2] stated in his study as well that the PMV showed high accuracies in predicting the neutal TSV 0. The Mean Absolute Errors (MAE) between prediction and actual values range between 0.66 and 0.92 for all the models. The MAE were less than 1 which means that the models predict in average one level higher or lower than the person really feels.

Table 2: Classification model performances on test set (ASHRAE) and on new data (case study)

Classification models	Test set (ASHRAE)				New data (Case study)			
	F1-score	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall
RF	0.49	0.5	0.5	0.5	0.39	0.39	0.39	0.39
KNN	0.47	0.5	0.48	0.49	0.37	0.36	0.39	0.39
GB	0.47	0.49	0.47	0.49	0.37	0.38	0.39	0.38
SVM	0.47	0.49	0.47	0.49	0.21	0.27	0.46	0.27
PMV	0.17	0.22	0.33	0.22	0.37	0.45	0.42	0.45

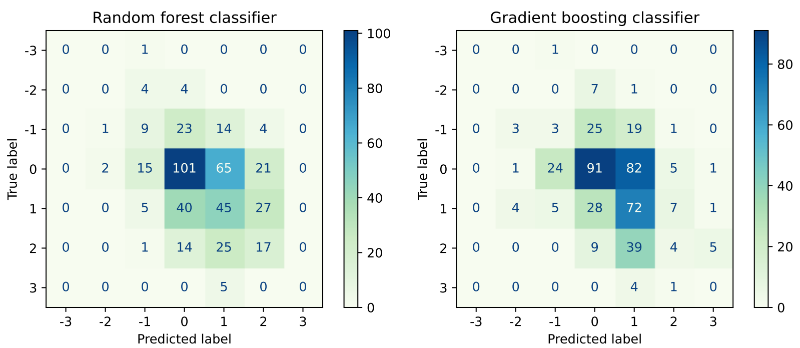

Figure 4: Random forest and gradient boosting classifier confusion matrix (case study)

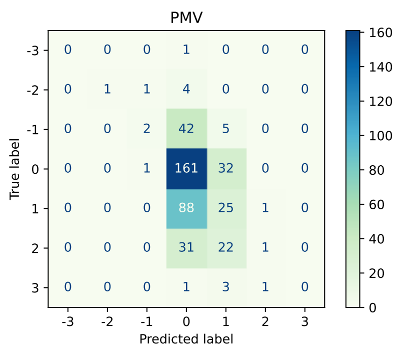

Figure 5: PMV confusion matrix (case study)

4 Conclusions

This study used the ASHRAE dataset to develop thermal comfort models that learn from different groups of individuals to predict individual thermal sensation of a specific group of people. The ASHRAE dataset is the only database of this kind that aggregates a significant number of measured and reported information from individuals from different continents. Nevertheless, in this considerable amount of data coming from different studies, the quality of the measured data should be questioned. The method performed sufficiently well and the model predicted new individuals' TSVs correctly almost 40% of the time. Due to the small number of participants in the experiments and the mild indoor environment of the case study, the collected data was less rich and diverse than the ASHRAE dataset which can be assumed to have led to the poor prediction of some of the TSV classes. This context lacks extreme indoor conditions which also allowed the PMV to perform as good as the data-driven models in this specific case but not on the ASHRAE test set where the PMV failed in predicting 80% of the votes. This method exemplifies limitations directly related to the data quality, its diversity, and its distribution in the training set and in the evaluation set. For a better understanding of people's thermal sensation, physiological data and behavioral data recorded on long time periods are essential. This time-dependent high amount of data will not only allow training models that predict personal thermal sensation with higher accuracies, but also allow understanding the underlying physiological properties of the different thermal sensations over time. Additionally, this data type is user independent and will therefore reduce uncertainties and inconsistencies of surveys.

Acknowledgements

This work was conducted as part of the research project “NuData Campus” funded by the Federal Ministry of Economics Affairs and Climate Action (BMWK). The authors are grateful to the BMWK for its support. The authors are grateful to the Senselab team (Building Technology and Climate Responsive Design, Prof. Dipl-Ing. Thomas Auer, TUM) and all the participants in the case study.

References

[1] ASHRAE, “Standard 55-2013 - thermal environmental conditions for human occupancy”, 2013.

[2] T. Cheung, S. Schiavon, T. Parkinson, P. Li, and G. Brager, “Analysis of the accuracy on pmv–ppd model using the ashrae global thermal comfort database ii”, *Build Environ*, 2019.

[3] A. Warthmann, D. Wölki, H. Metzmacher, and C. Van Treeck, “Personal climatization systems—a review on existing and upcoming concepts”, *Applied Sciences*, 2018.

[4] W. van Marken Lichtenbelt, M. Hanssen, H. Pallubinsky, B. Kingma, and L. Schellen, “Healthy excursions outside the thermal comfort zone”, *Build Res Inf*, 2017.

[5] J. Kim, S. Schiavon, and G. Brager, “Personal comfort models – a new paradigm in thermal comfort for occupant-centric environmental control”, *Build Environ*, 2018.

[6] L. A. Martins, V. Soebarto, and T. Williamson, “A systematic review of personal thermal comfort models”, *Build Environ*, 2022.

[7] J. Lee and Y. Ham, “Physiological sensing-driven personal thermal comfort modelling in consideration of human activity variations”, *Build Res Inf*, 2021.

[8] A. Aryal, B. Becerik-Gerber, G. M. Lucas, and S. C. Roll, “Intelligent agents to improve thermal satisfaction by controlling personal comfort systems under different levels of automation”, *IEEE Internet Things J*, 2020.

[9] A. Aryal and B. Becerik-Gerber, “Thermal comfort modeling when personalized comfort systems are in use: Comparison of sensing and learning methods”, *Build Environ*, 2020.

[10] S. Liu, S. Schiavon, H. P. Das, M. Jin, and C. J. Spanos, “Personal thermal comfort models with wearable sensors”, *Build Environ*, 2019.

[11] W. Jung, F. Jazizadeh, and T. E. Diller, “Heat flux sensing for machine-learning-based personal thermal comfort modeling”, *J Sens*, 2019.

[12] X. Zhou, L. Xu, J. Zhang, *et al.*, “Data-driven thermal comfort model via support vector machine algorithms: Insights from ashrae rp-884 database”, *Energy Build*, 2020.

[13] V. F. Ličina, T. Cheung, H. Zhang, *et al.*, “Development of the ashrae global thermal comfort database ii”, *Build Environ*, 2018.

[14] A. Burkov, *The hundred-page machine learning book*. A. Burkov Quebec City, Canada, 2019.

[15] T. O’Malley, E. Bursztein, J. Long, F. Chollet, H. Jin, L. Invernizzi, *et al.*, *Kerastuner*, https://github.com/keras-team/keras-tuner, 2019.

[16] B. Kobas, S. C. Koth, K. Nkurikiyeyezu, G. Giannakakis, and T. Auer, “Effect of exposure time on thermal behaviour: A psychophysiological approach”, *Signals*, 2021.

[17] F. Tartarini and S. Schiavon, “Pythermalcomfort: A python package for thermal comfort research”, *SoftwareX*, 2020.