Personal comfort model for automatic control of personal comfort systems

Dragos-Ioan Bogatu[a], *Ongun B.* Kazanci[a], *Futa* Watanabe[b], *Yosuke* Kaneko[b], *and Bjarne W.* Olesen[a]

[a] International Centre for Indoor Environment and Energy – ICIEE, Department of Civil Engineering, Technical University of Denmark, Nils Koppels Allé, Building 402, 2800 Kgs. Lyngby, Denmark, drabo@byg.dtu.dk

[b] Mitsubishi Electric Corporation, 5-1-1 Ofuna, Kamakura, Kanagawa 247-8501, Japan.

Personal comfort models could be used for the development of automatic controls for personal environmental comfort systems (PECS). These models often use indoor environment and physiological indicators as attributes for estimating the subjective response of occupants. Traditional indoor thermal environment research and standardization recommend 7-point scales for thermal comfort or thermal sensation estimation. However, many studies apply transformations to the response, thus oversimplifying the scales and generating controversy. The aim of this study is to determine the relevance of different indicators for the development of personal comfort models while investigating the implications and resulting model accuracy when using different thermal sensation scale discretization. Two simple machine learning algorithms, namely logistic regression and Naïve Bayes, were used in a multi-class setting to predict the overall thermal sensation of individual subjects when occupying a heated or cooled chair in steady state conditions. Multiple models were generated depending on the variables included in the feature set. Additionally, two response vectors were generated based on the thermal sensation vote, a three class and a seven class one, the latter being generated by further discretizing the hot and cold spectrum of thermal sensation. Both models performed better than a random guess at identifying thermal sensation classes and reached accuracies of up to 72% when predicting the overall thermal sensation of people using PECS. Including information of the PECS operation in the model, i.e. seat temperature, increased the prediction accuracy by up to 5%. The overall accuracy was higher when using three classes for the thermal response, as implementing seven classes led to a decrease of up to 21 percentage points. Nevertheless, the latter provided a finer adjustment without affecting the model's ability to distinguish between the cold and hot spectrum, which may be an advantage for personal comfort systems that condition the microenvironment of the occupant.

Keywords. Personal environmental comfort system, automatic control, thermal sensation vote, machine learning, response scale.
DOI: https://doi.org/10.34641/clima.2022.120

1. Introduction

Personal environmental comfort systems (PECS) are a topic of interest because they could improve comfort by dealing with the inter-personal differences between people [1], [2]. Their implementation could lead to energy savings [1], [2]. However, for an optimal operation, automatic controls based on machine learning algorithms are becoming attractive as an alternative to manual interaction between the occupant and the PECS.

Monitoring and including physiological indicators such as skin temperature and heart rate in the control of the indoor environment has the potential to suplement or replace occupant feedback [3]. By including these indicators alongside environmental parameters in thermal comfort prediction models based on machine learning algorithms can achieve accuracies of up to 98% [4][5]. However, it is still unclear which algorithms, indicators, or response scale should be used [1], [2], especially for PECS. Katić et al. [6] found that mean skin temperature, hand skin temperature, and PECS control intensity influence thermal comfort and should be included as features when PECS is in use. By integrating wrist, forehead, nose, and cheek skin temperature, Aryal et al. [7] managed to obtain accuracies of up to 85% for thermal comfort prediction when using no PECS, a fan, or a heater if the models were trained using only the data for each specific system. However, although both studies registered thermal sensation using the ASHRAE 7-point scale [8], it was discretized into three classes when using it as a response vector for the thermal comfort models as in other studies [2].

The goal of the present study was to assess whether machine learning algorithms could be used for the automatic control of PECS and which indicators and thermal response scale should be used. The prediction power of two algorithms, namely logistic regression and Naïve Bayes, when used to estimate overall thermal sensation while using a PECS under steady-state conditions was investigated. Different models were built with different sets of variables to assess their relevance for thermal comfort prediction when using a PECS. By discretizing thermal sensation into three or seven classes and retraining the models, the resulting differences could be compared.

2. Methods

2.1 Dataset

The dataset consisted of 980 observations of the operative temperature (Top), the seat temperature (Tseat), the skin temperature, and the overall thermal sensation (TSV). The data were obtained from twenty human subjects (ten females and ten males) in experiments in climate chambers with a heated/cooled seat. The subjects were university students with a body mass index (BMI) between 20 and 30. Clothing was maintained constant throughout each session but differed depending on the chamber temperature, between 0.5 clo when hot and 1 clo when cold.

In all, three scenarios were emulated: hot summer (44°C), cold winter (10°C), and near comfort (20°C). The seat was designed to heat or cool depending on the chamber temperature, cold or hot respectively to maintain comfort. The seat temperature was set to 40°C when heating, 20°C when cooling, and 34°C when close to comfort (Fig. 1). The measurement of

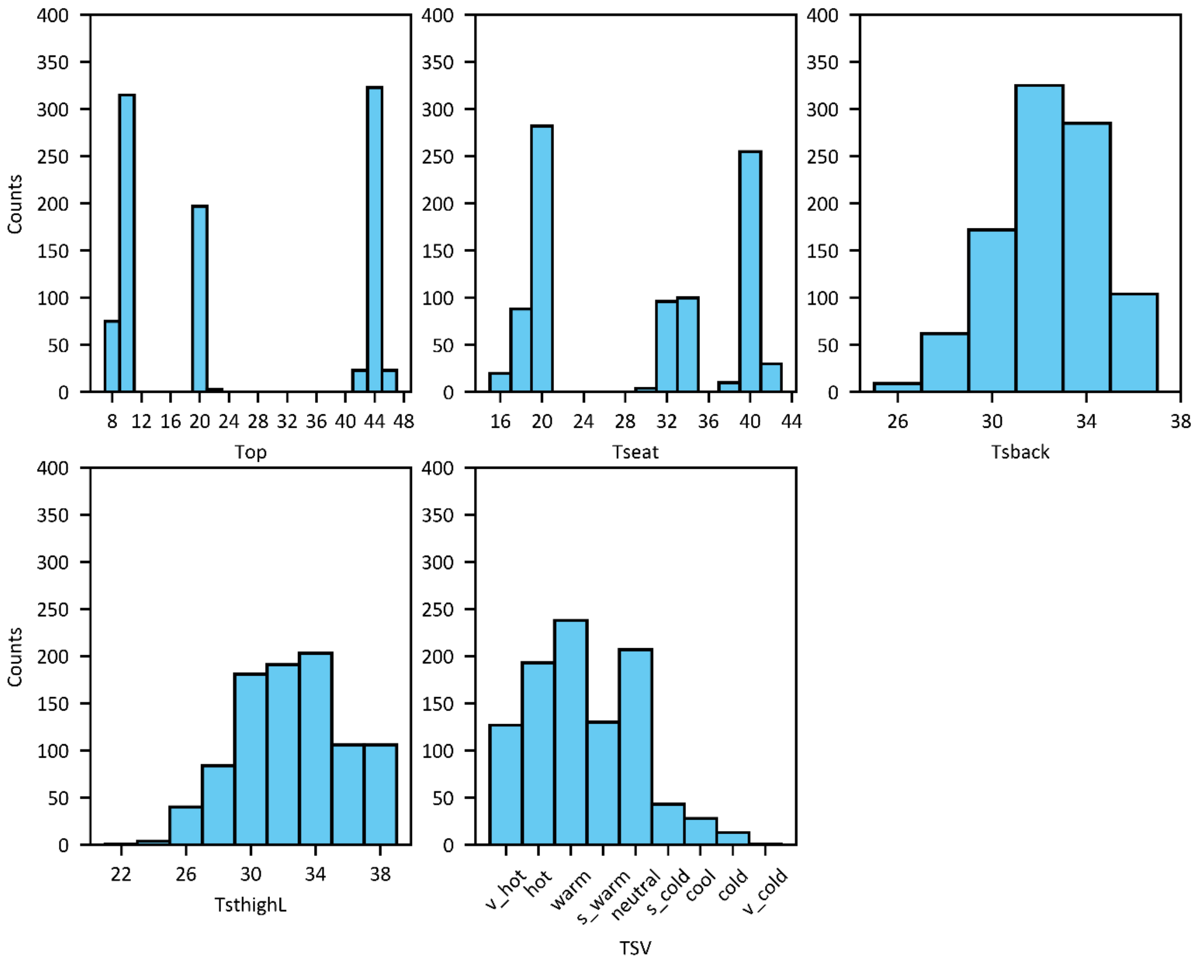

Fig. 1 - Histogram of variables in the dataset: operative temperature (Top), seat temperature (Tseat), skin temperature of the back (Tsback), left thigh (TsthighL), and left palm (TspalmL), and the overall thermal sensation of the subjects (TSV).

the operative temperature and the seat temperature had a five-minute resolution.

The skin temperature was monitored at two different points on the body, namely the back (Tsback), the left back thigh (TsthighL), both body parts in contact with the heated/cooled chair. The skin temperature had a resolution of one minute. As shown in Fig. 1, all skin temperatures were negatively skewed. The thermal sensation votes (TSV) followed a 9-point scale, namely very hot, hot, warm, slightly warm, neutral, slightly cold, cool, cold, and very cold. The TSV was positively skewed with very few “very cold” votes registered. In total ten thermal sensation votes were registered per condition from each subject.

2.2 Models

Due to the discrete and nominal nature of the multi-class predicted variable, namely the TSV, the problem represented a classification task. A logistic regression (Logit) and a Naïve Bayes (NB) algorithm were therefore selected for the study.

Six different sets of variables were used. The base set contained the operative temperature and the subject identification (sID). Three additional sets were generated by including the remaining variables, namely the seat temperature and the skin temperature of the back and thigh, in an alternating sequence. Two additional sets contained either the seat temperature and the sID or all variables except for the TSV. The sID, which indicates to which subject the variable refers, was always included since the aim was to generate an individual thermal comfort model.

For all models, the overall thermal sensation represented the response vector. However, the thermal sensation was simplified to either three (TSV3) or seven categories (TSV7). For both response vectors, the slightly hot and slightly cold classes were merged in the neutral class, as they are usually perceived as acceptable. Two other classes, hot and cold, were created in the TSV3 response by merging all the remaining hot and cold sensations, respectively. The rest of the remaining classes were kept separate for TSV7.

In total, seven different models were created for each algorithm. The first six used the TSV3 and each of the six sets of variables. For the last model, the TSV7 was used with the set containing all of the other variables in order to compare the results with the three-category scale of thermal sensation. In all instances, the performance of the two algorithms was compared to a baseline. The baseline represented a model that predicted everything as belonging to the class with the highest frequency in the response vector.

2.3 Training and test

Training and testing of the algorithms was carried out on the entire dataset in Python version 3.8.5. A two-level K-fold cross validation was used, with K_1=K_2=10 at each level. The data were shuffled before splitting them into the train and test sets. At the inner level (K_1), the respective complexity control parameter of each algorithm was optimized using only the training set of the outer level (K_2). In the regularized logistic regression algorithm (L2 penalty), the regularization parameter λ was tuned. In the NB algorithm Laplace smoothing was included, where $\alpha>0$ was chosen as a complexity controlling parameter and optimized.

At the outer level, the optimized algorithm was trained and tested using the entire dataset, thus estimating the generalization error. The optimized model was retrained at the outer level to avoid overestimating the generalization error [9]. The error rate, equal to the number of misclassified observations divided by the number of test observations, was used as a performance indicator at each level. The accuracy, precision, recall, and F1-score were computed for each model using Equations 1 to 5. The precision, recall, and F1-scores were computed using the weighted average due to the class imbalance observed in the data (Fig. 1), where the weight represented the number of true instances for each label:

$$Precision_c = \frac{TP_c}{TP_c+FP_c}, c \in [1, C] \tag{1}$$

$$Recall_c = \frac{TP_c}{TP_c+FN_c}, c \in [1, C] \tag{2}$$

$$Precision = \sum_{c=1}^{C} \frac{1}{n_c} \cdot Precision_c \tag{3}$$

$$Recall = \sum_{c=1}^{C} \frac{1}{n_c} \cdot Recall_c \tag{4}$$

$$F1 = \sum_{c=1}^{C} \frac{1}{n_c} \cdot \frac{2 \cdot Precision_c \cdot Recall_c}{Precision_c + Precision_c} \tag{5}$$

where TP is the number of true positives, FP the number of false positives, and FN the number of false negatives. C indicates the number of classes, and n represents the number of observations. In order to assess and compare the models, a pairwise statistical t-test was used [10][9].

3. Results

3.1 Correlation between the variables

The correlation (Pearson’s correlation coefficient) between the variables (Top, Tseat, Tsback, TsthighL) and the response (TSV) is shown in Fig. 2. The dark blue colour represents a strong negative correlation, while the dark red colour represents a strong positive correlation. A value close to 0 indicates no correlation.

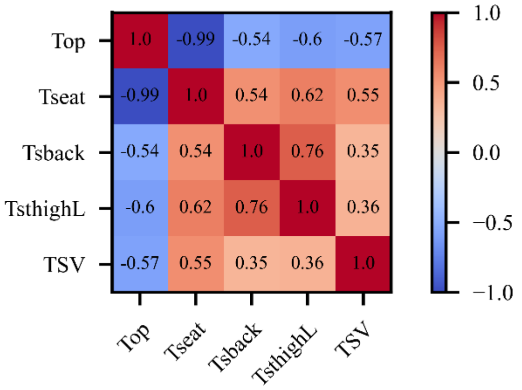

Fig. 2 - Pearson correlation between variables.

An almost perfect linear correlation was found between the Top and Tseat since when the operative temperature was low (10 °C) the seat temperature was high (40 °C) and vice-versa. All attributes were correlated with the TSV, with the Top and Tseat having a stronger correlation than the Tsback and TsthighL. The only negatively correlated attribute with the TSV was the operative temperature due to the scale's characteristics, as it spanned from very hot (0) to very cold (100). Since the back and the left thigh were in contact with the seat, the skin temperature (Tsback and TsthighL) was positively correlated with the seat temperature while negatively correlated with the operative temperature. This also meant that the two skin temperatures were linearly correlated.

3.2 Model performance

Tab. 1 shows the results of the models obtained from the two-level K-fold cross validation by algorithm and variable set. In order to quantify their ability to predict thermal sensation, the weighted precision, recall, and F1-score are given together with the overall accuracy of each model. The performance of the Baseline model was the same for each set of variables. However, it was dependent on the response vector, TSV3 or TSV7, since the most frequent thermal sensation was in the hot class for the former and in the neutral class for the latter.

Tab. 1 – Accuracy and weighted precision, recall, and F1-score by algorithm, feature set, and response used.

Algorithm	Feature set	Response	Precision	Recall	F1-score	Accuracy
Baseline		TSV3	0.32	0.57	0.41	0.57
Logit	Top, sID	TSV3	0.66	0.67	0.66	0.67
NB			0.65	0.66	0.61	0.66
Logit	Tseat, sID	TSV3	0.64	0.67	0.65	0.67
NB			0.6	0.63	0.6	0.63
Logit	Top, sID, Tseat	TSV3	0.71	0.72	0.71	0.72
NB			0.71	0.68	0.67	0.68
Logit	Top, sID, Tsback	TSV3	0.67	0.68	0.67	0.68
NB			0.73	0.69	0.69	0.69
Logit	Top, sID, TsthighL	TSV3	0.66	0.67	0.67	0.67
NB			0.71	0.67	0.68	0.67
Logit	Top, sID, Tseat, Tsback, TsthighL	TSV3	0.72	0.72	0.72	0.72
NB			0.71	0.68	0.67	0.68
Baseline		TSV7	0.15	0.39	0.22	0.39
Logit	Top, sID, Tseat, Tsback, TsthighL		0.56	0.57	0.56	0.57
NB			0.43	0.47	0.42	0.47

Both the Logit and the NB algorithms outperformed the Baseline. The accuracy of the Logit algorithm was higher by 0.1 to 0.15 when TSV3 was used as a response vector. Moreover, with the seven-point scale the difference increased by 0.18. The NB algorithm surpassed the Baseline by only 0.6 to 0.12 and 0.8 depending on the scale of the response, three or seven categories respectively.

For the same response vector, TSV3, the inclusion of both Top and Tseat in the feature set led to the highest accuracy, 72%. However, changing the operative temperature for the seat temperature as a model variable when predicting thermal sensation made little difference. Additional features, i.e. the skin temperatures, did not improve the accuracy of any of the algorithms.

The precision, recall, and F1-score were always between 60% and 72% for the Logit and NB when the response had only three classes. The model was able to identify a great degree of samples from each class while mismatching few samples [11], [12]. In general, the Logit algorithm obtained higher values. Values below 60% were observed when the response vector had seven classes. Still, with both three or seven classes, the precision and recall improved. With three classes for the TSV, the precision approximately doubled while the recall improved by up to 0.15 compared to the Baseline. When the response was changed to seven classes, the precision and recall increased by 0.41 and 0.18 for the Logit and by 0.28 and 0.8 for the NB. This led to an F1-score increase of 0.34 for the Logit and 0.2 for the NB algorithms.

Fig. 3 shows the confusion matrices for some of the Logit models. By comparing Fig. 3 a and b it was observed that the number of class matches (grey) increased while the number of critical mismatches between the cold and hot spectrum (red) decreased with the inclusion of additional variables. Additionally, although the overall mismatches increased when the response was categorized in seven classes, the number of misclassified observations between the hot and cold spectrum remained almost constant (Fig. 3 c).

3.3 Statistical evaluation

The performance of the three models relative to each other, i.e. Logit vs. Baseline, Logit vs. NB, and NB vs. Baseline, was evaluated using pairwise statistical t-tests. The 95% confidence intervals (CI), $\alpha=0.05$, and the p-values are shown in Tab. 2. The performance differs for a p-value < 0.05, i.e. the probability that the two classifiers would predict differently given the number of different predictions made.

a

		PREDICT		
		Hot	Neutral	Cold
TRUE	Hot	444	112	2
	Neutral	166	206	8
	Cold	10	26	6

b

		PREDICT		
		Hot	Neutral	Cold
TRUE	Hot	446	110	2
	Neutral	135	239	6
	Cold	5	17	20

c

		PREDICT						
		V_Hot	Hot	Warm	Neutral	Cool	Cold	V_Cold
TRUE	V_Hot	81	28	5	13	0	0	0
	Hot	32	89	36	35	0	1	0
	Warm	11	25	89	112	0	1	0
	Neutral	5	10	70	289	2	4	0
	Cool	1	0	2	12	11	2	0
	Cold	0	1	0	8	2	2	0
	V_Cold	0	0	0	0	0	1	0

Fig. 3 - Confusion matrices of different logistic regression models with features: a) Top, sID, TSV3, b) Top, sID, Tseat, Tsback, TsthighL, c) Top, sID, Tseat, Tsback, TsthighL. The correctly classified and the misclassified observations between the warm and cold spectrum are highlighted in grey and red, respectively.

Tab. 2 – The pairwise statistical t-test results between the three algorithms, Logit, NB, and Baseline.

Test	Feature set	Response	Confidence interval (CI)	p-value
Logit vs. Baseline	Top, sID	TSV3	$[0.06;\ 0.15]$	$2 \cdot 10^{-4}$
	Tseat, sID		$[0.05;\ 0.15]$	$3 \cdot 10^{-4}$
	Top, Tseat, sID		$[0.09;\ 0.19]$	$4 \cdot 10^{-6}$
	Top, Tsback, sID		$[0.07;\ 0.14]$	$1 \cdot 10^{-5}$
	Top, TsthighL, sID		$[0.06;\ 0.15]$	$4 \cdot 10^{-5}$
	Top, Tseat, Tsback, TsthighL, sID		$[0.1;\ 0.19]$	$1 \cdot 10^{-6}$
	Top, Tseat, Tsback, TsthighL, sID	TSV7	$[0.12;\ 0.25]$	$5 \cdot 10^{-6}$
Logit vs. NB	Top, sID	TSV3	$[0.01;\ 0.06]$	$8 \cdot 10^{-3}$
	Tseat, sID		$[0.02;\ 0.07]$	$5 \cdot 10^{-3}$
	Top, Tseat, sID		$[0.02;\ 0.07]$	$5 \cdot 10^{-4}$
	Top, Tsback, sID		$[0.01;\ 0.07]$	$1 \cdot 10^{-2}$
	Top, TsthighL, sID		$[0.02;\ 0.06]$	$7 \cdot 10^{-4}$
	Top, Tseat, Tsback, TsthighL, sID		$[0.02;\ 0.07]$	$3 \cdot 10^{-3}$
	Top, Tseat, Tsback, TsthighL, sID	TSV7	$[0.04;\ 0.15]$	$1 \cdot 10^{-3}$
NB vs. Baseline	Top, sID	TSV3	$[0.06;\ 0.12]$	$1.3 \cdot 10^{-6}$
	Tseat, sID		$[0.04;\ 0.1]$	$2 \cdot 10^{-4}$
	Top, Tseat, sID		$[0.05;\ 0.18]$	$2 \cdot 10^{-3}$
	Top, Tsback, sID		$[0.05;\ 0.17]$	$6 \cdot 10^{-4}$
	Top, TsthighL, sID		$[0.06;\ 0.16]$	$1 \cdot 10^{-4}$
	Top, Tseat, Tsback, TsthighL, sID		$[0.07;\ 0.17]$	$1 \cdot 10^{-6}$
	Top, Tseat, Tsback, TsthighL, sID	TSV7	$[0.05;\ 0.14]$	$2 \cdot 10^{-4}$

Tab. 2 shows that none of the 95% confidence intervals (CI's) of the pairwise t-tests contain zero. Thus, there was strong evidence that both the Logit and the NB had a higher accuracy than the Baseline. Although there was still evidence that the Logit had a higher accuracy than the NB, the effect was not as significant in this case. For all three pairwise comparisons, the p-values are significantly smaller than the chosen level, 0.05. In general, the lower and upper bounds of the confidence interval moved further away from zero as more variables were included.

4. Discussion

The focus of this study was to determine whether machine learning algorithms could be used for the control of personal comfort systems and which indicators should be used as input. Two simple algorithms, logistic regression and Naïve Bayes, were used to assess their ability to predict the overall thermal sensation based on the operative temperature, the seat temperature, and the skin temperature in contact with the seat, namely of the back and thigh, was assessed.

The Logit and NB algorithms managed to reach accuracies up to 72% for a three class TSV response, supporting the values previously reported in literature [13], [14]. Although this only meant an increase of up to 15 percentage points from the Baseline, the use of the two algorithms actually reduced the misclassification rate while increasing the models' effectiveness at identifying each class, as seen in the increase in recall and precision. Moreover, the high accuracy of the Baseline was a consequence of the skewness in thermal sensation votes, which were concentrated on the warm side. This could be due to the high heating and low cooling effectiveness of the chair as shown by the distribution of the subjects' votes. Therefore, even at low or high operative temperatures the subjects rarely reported feeling cold. The actual performance of the models when used with a more balanced dataset requires further investigation.

The highest overall increase in performance was observed when both Top and Tseat were included in the input set, confirmed by the statistical evaluation. Even though the two variables were collinear, they had an opposite effect on the TSV as the chair was supposed to counter the loss in thermal comfort caused by the change in operative temperature. Moreover, since the back side of the body was in contact with the chair while the front side was directly exposed to the ambient thermal conditions it was expected that both would be correlated almost equally although inversely with the TSV, thus providing additional information to the model.

Having the skin temperatures from the two body parts in contact with the heated/cooled chair in the set of input variables did not improve the model performance. As observed both the Tsback and TsthighL showed a weak correlation with TSV. At the same time, they were not perfectly correlated with each other, most likely due to the non-uniform seat temperature and area of contact. Moreover, since the skin temperature was measured at points in direct contact with the chair, these variables were providing the same information to the model as the seat temperature. As previous research showed, there could be other physiological and environmental indicators, e.g. relative humidity, heart rate, and the heat flux between the seat and the occupant that could improve the prediction of the TSV [15]–[18]. Additionally, in a dynamic environment the time constant of the heating or cooling element of the chair might make these aforementioned indicators more relevant.

When investigating the effect of the response scale it was determined that an increased discretization led to a decrease in performance, and this applied also to the Baseline model. Even with a higher discretization the model maintained the same distinction power between hot and cold sensations. The problem was though differentiating between the neutral and the warm or cool sensations and the increase in computational power. A higher discretization may provide additional information regarding the intensity of the thermal sensation and increased flexibility for the control but questions remain on the optimality of the scale and its discretization for thermal comfort prediction [2].

The difference in performance between the two algorithms, Logit and NB, is inherently due to the way the two operate. The logistic regression is a linear model while the relationship between environmental and physiological indicators and thermal sensation is complex [1]. The NB assumes the variables are conditionally independent [1], [9], which could also be the reason for the slightly worse performance of the NB in the present dataset. Other algorithms such as decision trees, support vector machines, and artificial neural networks should be considered since previous research has shown that they are useful for thermal comfort estimation [1], [19]. Including feature transformations such as the gradient or mean of different variables should also be considered in future model development.

5. Conclusion

Both algorithms, the Logistic Regression and the Naïve Bayes, achieved a sub-optimal performance registering an accuracy of up to 72% when predicting the thermal sensation of subjects using a PECS. However, the models performed better than a random guess.

Information on the operation of the PECS should be included in the set of input variables as the prediction accuracy increased by up to 5% when the seat temperature was added to the operative temperature. This is in line with the findings of Katić et al. [6] and Aryal et al. [7] who showed that the PECSs' settings should be used as input variables in personal comfort models as they influence individual thermal preference. However, the skin temperature of any body parts in direct contact with the PECS may not provide any useful additional information for thermal sensation prediction. These findings support the idea that control signals and measurements of the PECS output, which are easier to obtain than physiological indicators, could be used to improve the predictive accuracy of personal comfort models.

The benefit of having additional classes and hence a more accurate control came at a greater computational cost. Nevertheless, this did not affect the model's ability at distinguishing between the cold and hot spectrum.

Dataset

The datasets generated during and/or analysed during the current study are not available because of a non-disclosure agreement but the authors will make every reasonable effort to publish them in near future.

Acknowledgement

This study was financially supported by Mitsubishi Electric Corporation and by the International Centre for Indoor Environment and Energy (ICIEE), Technical University of Denmark (DTU).

References

[1] J. Kim, S. Schiavon, and G. Brager, "Personal comfort models – A new paradigm in thermal comfort for occupant-centric environmental control," *Build. Environ.*, vol. 132, no. January, pp. 114–124, 2018, doi: 10.1016/j.buildenv.2018.01.023.

[2] M. André, R. De Vecchi, and R. Lamberts, "User-centered environmental control: a review of current findings on personal conditioning systems and personal comfort models," *Energy Build.*, vol. 222,

2020, doi: 10.1016/j.enbuild.2020.110011.

[3] A. Ghahramani, G. Castro, B. Becerik-Gerber, and X. Yu, “Infrared thermography of human face for monitoring thermoregulation performance and estimating personal thermal comfort,” *Build. Environ.*, vol. 109, pp. 1–11, 2016, doi: 10.1016/j.buildenv.2016.09.005.

[4] F. Salamone *et al.*, “Integrated method for personal thermal comfort assessment and optimization through users’ feedback, IoT and machine learning: A case study,” *Sensors (Switzerland)*, vol. 18, no. 5, 2018, doi: 10.3390/s18051602.

[5] S. Lu, W. Wang, S. Wang, and E. C. Hameen, “Thermal comfort-based personalized models with non-intrusive sensing technique in office buildings,” *Appl. Sci.*, vol. 9, no. 9, 2019, doi: 10.3390/app9091768.

[6] K. Katić, R. Li, and W. Zeiler, “Machine learning algorithms applied to a prediction of personal overall thermal comfort using skin temperatures and occupants’ heating behavior,” *Appl. Ergon.*, vol. 85, no. February, 2020, doi: 10.1016/j.apergo.2020.103078.

[7] A. Aryal and B. Becerik-Gerber, “Thermal comfort modeling when personalized comfort systems are in use: Comparison of sensing and learning methods,” *Build. Environ.*, vol. 185, no. August, p. 107316, 2020, doi: 10.1016/j.buildenv.2020.107316.

[8] ASHRAE, “ASHRAE 55-2017 Thermal Environmental Conditions for Human Occupancy.” 2017.

[9] T. Herlau, M. N. Schmidt, and M. Mørup, *Introduction to Machine Learning and Data Mining*. 2021.

[10] A. Benavoli, G. Corani, J. Demšar, and M. Zaffalon, “Time for a change: A tutorial for comparing multiple classifiers through Bayesian analysis,” *J. Mach. Learn. Res.*, vol. 18, pp. 1–36, 2017.

[11] M. Sokolova and G. Lapalme, “A systematic analysis of performance measures for classification tasks,” *Inf. Process. Manag.*, vol. 45, no. 4, pp. 427–437, 2009, doi: 10.1016/j.ipm.2009.03.002.

[12] M. Grandini, E. Bagli, and G. Visani, “Metrics for Multi-Class Classification: an Overview,” pp. 1–17, 2020, [Online]. Available: http://arxiv.org/abs/2008.05756.

[13] S. Liu, S. Schiavon, H. P. Das, M. Jin, and C. J. Spanos, “Personal thermal comfort models with wearable sensors,” *Build. Environ.*, vol. 162, no. July, p. 106281, 2019, doi: 10.1016/j.buildenv.2019.106281.

[14] D. Li, C. C. Menassa, V. R. Kamat, and E. Byon, “HEAT - Human Embodied Autonomous Thermostat,” *Build. Environ.*, vol. 178, no. March, 2020, doi: 10.1016/j.buildenv.2020.106879.

[15] T. Chaudhuri, D. Zhai, Y. C. Soh, H. Li, and L. Xie, “Random forest based thermal comfort prediction from gender-specific physiological parameters using wearable sensing technology,” *Energy Build.*, vol. 166, pp. 391–406, 2018, doi: 10.1016/j.enbuild.2018.02.035.

[16] Z. Deng and Q. Chen, “Development and validation of a smart HVAC control system for multi-occupant offices by using occupants’ physiological signals from wristband,” *Energy Build.*, vol. 214, p. 109872, 2020, doi: 10.1016/j.enbuild.2020.109872.

[17] I. Pigliautile, S. Casaccia, N. Morresi, M. Arnesano, A. L. Pisello, and G. M. Revel, “Assessing occupants’ personal attributes in relation to human perception of environmental comfort: Measurement procedure and data analysis,” *Build. Environ.*, vol. 177, no. March, p. 106901, 2020, doi: 10.1016/j.buildenv.2020.106901.

[18] Y. F. Zhang, D. P. Wyon, L. Fang, and A. K. Melikov, “The influence of heated or cooled seats on the acceptable ambient temperature range,” *Ergonomics*, vol. 50, no. 4, pp. 586–600, 2007, doi: 10.1080/00140130601154921.

[19] N. Ma, D. Aviv, H. Guo, and W. W. Braham, “Measuring the right factors: A review of variables and models for thermal comfort and indoor air quality,” *Renew. Sustain. Energy Rev.*, vol. 135, no. August 2020, p. 110436, 2021, doi: 10.1016/j.rser.2020.110436.