OmniaScience

Journal of Industrial Engineering and Management
JIEM, 2025 – 18(3): 622-630 – Online ISSN: 2013-0953 – Print ISSN: 2013-8423
https://doi.org/10.3926/jiem.8605

Insight Indoor Adaptive Behavioural Adjustments Considering A Machine Learning-Based Method

Elena Barbadilla-Martín*, Pablo Aparicio-Ruiz, Juan Carlos Ragel-Bonilla, José Guadix

Grupo de Ingeniería de Organización, Escuela Técnica Superior de Ingeniería, Universidad de Sevilla (Spain)

**Corresponding author: ebarbadilla@us.es*
pabloaparicio@us.es, jragel@us.es, guadix@us.es

Received: November 2024
Accepted: August 2025

Abstract:

Purpose: Research focused on indoor thermal comfort has increased significantly in recent years. This is mainly due to the relationship of said variable with energy efficiency, derived from an optimal operation of conditioning systems. Related to this, it is important to highlight the relevance of the clothing insulation level, due to its connection with thermal comfort.

Design/methodology/approach: The present study proposes a machine learning-based method, by applying four Machine Learning algorithms. For this purpose, the ASHRAE Global Thermal Comfort Database II was used, and nineteen features were selected as input variables.

Findings: Among the four algorithms considered, Multi-Layer Perception more accurately predicted the clothing insulation level of the participants (R2=0,883). Moreover, as for the relevance of the input variables, those related to indoor and outdoor climatic conditions had the greatest effect on the estimation of the observed output.

Originality/value: A Machine Learning-based approach to delve into the analysis of the clothing insulation level, as opposed to other studies that rely on regression models, is proposed. Moreover, it is not based on a single technique but carries out a systematic comparative analysis of several of the most commonly in the field of thermal comfort. Additionally, it seeks to encompass a wide range of input variables to investigate their relationship with the clothing insulation level.

Keywords: machine learning, building management, building operation, adaptive behaviour

To cite this article:

Barbadilla-Martín, E., Aparicio-Ruiz, P., Ragel-Bonilla, J.C., & Guadix, J. (2025). Insight indoor adaptive behavioural adjustments considering a machine learning-based method. *Journal of Industrial Engineering and Management*, 18(3), 622-630. https://doi.org/10.3926/jiem.8605

1. Introduction

The design of buildings faces different challenges, among which, there are some related to the construction itself (the envelope or the materials used) and others related to the indoor environment. Regarding the latter, building occupants are spending more time indoors, so ensuring an adequate indoor environment in terms of thermal comfort or air quality has become a priority.

Thermal comfort could be defined as *"the condition that satisfies the thermal needs of a user according to their own sensations"* (ANSI/ASHRAE, 2020) and its analysis dates back to the 70s when (Fanger, 1970) published a model based on studies in climatic chamber to assess thermal comfort, known as the Predictive Mean Vote (PMV). Since then, numerous studies have been conducted to better understand the factors that influence thermal comfort, such as indoor temperature, relative humidity or air velocity. Subsequently, (Nicol & Humphreys, 2010) proposed another model based on field studies, known as the adaptive model, which is based on the fact that buildings' occupants can adapt to achieve thermal comfort. This approach also considers the outdoor variable as a relevant factor for indoor comfort, as well as other aspects such as psychological or behavioural adaptation.

The relevance of thermal comfort lies not only in the study itself, but also in the relationship between thermal comfort and energy efficiency. Sometimes, situations of overheating and overcooling are detected in buildings, making it essential to balance thermal comfort with energy consumption, due to the implications this has on greenhouse gas emissions, global warming and sustainable development goals.

Due to the above, the study of thermal comfort is constantly evolving and in recent years, a significant increase in the use of artificial intelligence models and machine learning (ML) algorithms can be detected in this field. ML algorithms make it possible to process large datasets and delve into complex patterns that could be difficult to identify using statistical methods. Therefore, they are being used to develop predictive thermal comfort models that take into account a wide variety of input variables for optimizing heating, ventilating and air conditioning (HVAC) systems.

For that reason, this paper addresses the optimization of thermal comfort through the study of a variable related to it, the clothing insulation level of the occupants of a building, using ML techniques. For this purpose, a literature review is carried out in section 2, the database considered, and the methods followed are presented in section 3. Section 4 shows the results obtained and discusses them and, finally, section 5 summarizes the main findings.

2. Literature Review

The clothing insulation level (clo) is defined as the intrinsic clothing ensemble insulation of a certain subject (Gagge, Burton & Bazett, 1941), existing different methods to estimate said variable exposed in standards ASHRAE 55 (ANSI/ASHRAE, 2020) or EN ISO 9920:2009 (ISO, 2009).

The analysis of clo and its prediction is relevant, not only because of its value itself, but also because the amount of thermal insulation worn by a person has a significant effect in their thermal comfort (ANSI/ASHRAE, 2020). Because of this, the value of clo is a factor in both, the calculation of the PMV index, as well as in the adaptive model. The latter, considers physiological, psychological and behavioural adaptive actions to achieve optimal thermal comfort, and includes clothing adaptation, that is, clothing changes, as part of the last group of adaptive options (De Dear & Brager, 1998).

One of the first field studies on clothing insulation was carried out by (Humphreys, 1977), where the clothing worn by children at school was analysed, as well as the variations of clothing with the change of ambient temperature throughout the day and during various days. (Humphreys, 1979) also subsequently analysed the influence of season and ambient temperature on human clothing behaviour.

Nicol and Roaf (1996) conducted a study on thermal comfort in Pakistan, in which they analysed changes in clothing as evidence of adaptive behaviour, highlighting variations in clothing insulation between summer and winter seasons.

De Dear and Brager (1998) concluded that the mean thermal insulation was significantly related to indoor operative temperature and they calculated how the mean thermal insulation would decrease while increasing the building's mean indoor temperature. (Morgan & De Dear, 2003) highlighted the relationship between the average clothing insulation level and the indoor temperature, and, especially, with the outdoor temperature. In addition, they identified that indoor clothing levels varied according to gender, company dress code and climate. With the data obtained and using linear regression, a prediction model of the clothing level was proposed.

(Haldi & Robinson, 2011) based on results from a field survey campaign conducted in Switzerland, showed that occupants' variations in clothing choices are best described by outdoor temperature variables, defining a linear regression between them. Moreover, other studies have focused on analysing clo and its relationship with different variables (Su, Wang, Zhou, Duanmu, Zhai, Lian et al., 2022; De Carli, Olesen, Zarrella & Zecchin, 2007).

The former studies, while making a significant contribution to the field of study, stand out for proposing different models that relate the level of clothing insulation with diverse variables, mainly based on regression analysis. Therefore, ML algorithms, in addition, could facilitate the ensemble analysis of a larger number of variables and deepen the relationship between them.

In this sense, in the field of thermal comfort, the application of ML algorithms for the prediction of the thermal sensation expressed by a vote on a Likert scale that represents the opinion of a subject on the indoor thermal environment, stands out (Farhan, Pattipati, Wang & Luh, 2015; Wu, Li, Peng, Cui, Liu, Li et al., 2018; Qavidel-Fard, Zomorodian & Korsavi, 2022). However, as for the study and prediction of the clothing insulation level using ML algorithms, fewer studies can be found in the literature (Ngarambe, Yun & Kim, 2019; Duhirwe, Ngarambe & Yun, 2022).

In particular (Ngarambe et al., 2019) carried out a field study and developed a deep neural network model that forecasts daily clothing insulation levels considering outdoor variables, gender, season and mode of transport as input variables. They concluded that the neural networks obtained better results compared to linear model using the same data. Also, (Duhirwe et al., 2022) developed a convolutional neural network in order to predict clothing insulation.

Therefore, having highlighted the relevance of the clothing insulation level because of its relationship with thermal comfort and energy consumption, and since it is a challenging objective to study, the present study focuses on analysing it. For that, an approach based on the comparative analysis of several machine learning algorithms is proposed, rather than focusing on a single technique, and also trying to encompass a wide range of input variables to investigate the interaction between clo with other features.

3. Sample and Methods

The sample and the methods considered in the study are described in the present section.

3.1. Sample

ASHRAE Global Thermal Comfort Database II is a public thermal comfort database which includes samples from different field studies. Its origin is in the Smart Controls and Thermal Comfort (SCATS) project, based on monitoring indoor and outdoor variables and survey campaigns regarding thermal comfort conducted in the late 90s in different European countries and, subsequently, expanded with data from new studies in other countries (McCartney & Nicol, 2002; De Dear, 1998). Nowadays, ASHRAE Global Thermal Comfort Database II includes 107.463 records from 52 field studies, comprising residential, office, classrooms and senior centre buildings, different countries, climates as well as environmental controls and demographic features of buildings' occupants (Földváry-Ličina, Cheung, Zhang, De Dear, Parkinson, Arens et al., 2018), among others. Therefore, based on the diversity of information, the data sample considered in the present study derives from the ASHRAE Global Thermal Comfort Database II.

Since the level of clothing insulation was chosen as the targeted comfort index, Table 3 presents a description of the variables considered for predicting it, taking into account the literature review in section "2. Literature review". A total of nineteen input variables and one output variable (level clothing insulation) were chosen.

Variable	Description	Symbol
Season	Spring, summer, autumn, winter	season
Climate	Cool/dry, hot/wet	climate
Age	Age of the participants	age
Sex	Male, female, undefined	sex
Height	Participants' height	ht
Weight	Participants' weight	wt
Met rate	Average metabolic rate of the participants	met
Air temperature	Indoor air temperature	ta
Operative temperature	Indoor operative temperature	top
Radiant temperature	Indoor radiant temperature	tr
Globe temperature	Indoor globe temperature	tg
Relative humidity	Indoor relative humidity	rh
Air velocity	Indoor air velocity	va
Thermal sensation	ASHRAE thermal sensation vote from -3 to 3	TSV
Thermal preference	Cooler, no change, warmer	TP
Thermal acceptability	Unacceptable, acceptable	TA
Air movement preference	Less, no change, more	vaP
Air movement acceptability	Unacceptable, acceptable	vaA
Outdoor temperature	Outdoor monthly average temperature	tout

Table 1. Variables considered in the study

3.2. Methods

Figure 1 shows the overall methods considered in which different stages, ranging from variables selection to performance evaluation, are involved.

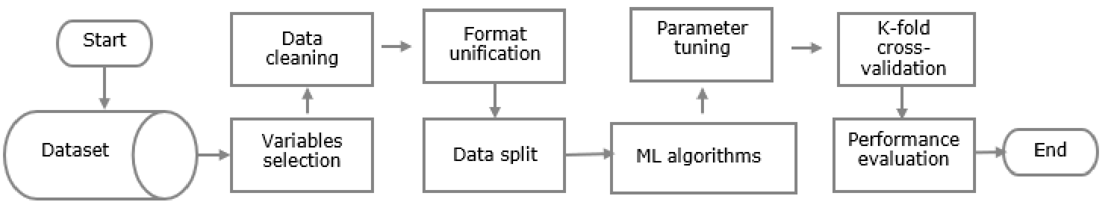

Figure 1. Methods

3.2.1. Variables Selection

The ASHRAE Global Thermal Comfort Database II comprises a total of forty-nine variables related to thermal comfort for each record. So, firstly, the raw dataset was analysed in order to identify those variables with missing values or variables related that could be redundant. Likewise, and based on the literature review (section 2), nineteen input variables (Table 1) were selected, including variables related to participants, subjective thermal comfort information, instrumental thermal or comfort measurements (Földváry-Ličina et al., 2018). The intrinsic clothing ensemble insulation of the subject participants in the field studies (clo) was defined as the output target variable.

3.2.2. Data Cleaning, Format Unification and Data Split

For the variables selected, processing was conducted to train the ML algorithms with a quality dataset. For this purpose, missing values and invalid data or outliers were identified.

Moreover, the inputs variables selected have different units and data distribution, since large numeric data variables (such as air temperature or outdoor temperature) or categorical variables (such as season or sex) could be identified. Therefore, we unified the formats of input variables before analysing their effect on the clothing insulation level.

Additionally, the 80-20 training-testing proportion was chosen because they are values commonly used in the area of knowledge (Földváry-Ličina et al., 2018).

3.2.3. ML Algorithms

Based on the literature review section, frequently ML algorithms used for thermal comfort prediction were analysed and four were considered for a comparison between them: k-nearest neighbours (K-NN), random forest (RF), gradient boosting trees (GBTR) and neural networks (Multilayer perceptron, MLP). The following is a brief description of them, highlighting their main characteristics.

k-nearest neighbours (KNN) algorithm tries to find to which class an unknown datum belongs to, by relating it to those that have similar characteristics to it (neighbors). To do this, the algorithm searches for the "k" points closest to the unknown datum by identifying the distance between them. The datum is finally classified according to the majority of the votes of the close neighbors (Müller & Guido, 2016; Mitchell, 1997).

Random forest (RF) and Gradient boosting trees (GBTR) are based on Decision Trees algorithms, which consist of a formation created by nodes, branches and leaves, and being the learning process based on decisions made through hierarchical questions. It is common in Machine Learning to use ensembles of algorithms to obtain better models, but of all these, decision tree ensembles stand out especially for classification and regression tasks. The most commonly used in this category are Random Forest and Gradient-Boosted Regression Trees (GBRT) (Müller & Guido, 2016).

Neural networks are a type of algorithm with a learning model inspired by the human brain and the operation of neurons. In particular, Multi-Layer Perception (MLP) is a type of neural network that is characterized because it has also hidden layers in which each neuron relates to the input layer and output layer and is a simple forwarding propagation method, among which exist (Luo, Xie, Yan, Ke, Yu, Wang et al., 2020).

3.2.4. Parameter Tuning And Cross-Validation

To obtain optimal hyperparameters, we used the grid searching method. Based on this method, Table 2 shows the value of the main hyperparameters for each algorithm considered.

	Parameter	Tuning range	Optimal values
KNN	n-neighbors	(1-20)	8
	leaf_size	(1-50)	22
	p	(1-2)	1
RF	n_estimators	(5-500)	470
	min_samples_split	(2-20)	7
	min_samples_leaf	(1-20)	3
GBRT	n_estimators	(5-500)	465
	min_samples_split	(2-20)	3
	min_samples_leaf	(1-20)	2
MLP	hidden layers size	(1-4)	3 (256, 128, 64)

Table 2. Key parameters for ML algorithms

In the present study, 5-fold cross-validation was chosen because they are values commonly used in the area of knowledge (Luo et al., 2020).

3.2.5. Performance Evaluation

As for metrics in order to evaluate the performance of the algorithms considered, that is, the model fit quality or the errors of the predictions of each model with respect to the test data, three rates were selected: Correlation coefficient square, mean square error and mean absolute error.

Correlation coefficient square (R2) quantifies the deviations between predicted and real system outputs (Equation 1).

$$R2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_l)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{1}$$

Where yi corresponds to each of the observed values, $\hat{y}_l$, are the values of the predictions and $\bar{y}$ is the mean of the predictions.

The mean square error (MSE) focuses on the square of the difference between observations and predictions divided by the number of cases, penalising strongly discrepancies between actual values and predictions (Equation 2).

$$MSE = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_l)^2}{n} \tag{2}$$

The mean absolute error (MAE) is calculated based on Equation 3.

$$MAE = \frac{\sum_{i=1}^{n}|y_i - \hat{y}_l|}{n} \tag{3}$$

4. Discussion

In this section the results of implementing the four algorithms considered are shown and a discussion of them is carried out. Additionally, a study on the relevance of the input variables for the prediction of the output variable is also described.

4.1. Comparison of ML Algorithms

Based on the results obtained by the Machine Learning algorithms, a comparative analysis of their performance could be conducted based on the evaluation metrics considered (R2, MSE and MAE), which are shown in Table 3.

	KNN	RF	GBTR	MLP
R2	0,679	0,782	0,653	0,883
MSE	0,020	0,019	0,029	0,009
MAE	0,102	0,089	0,123	0,067

Table 3. Performance of ML algorithms

As for the correlation coefficient square, both, Random Forest and MLP, achieved better results than the other two, with the neural network showing a significant difference in the predictions.

The mean square error for the GBTR algorithm is significantly higher, so it can be assumed that the predictions made by said algorithm have significant deviations, since this rate penalizes large errors in the predictions. As for KNN and RF, they have same the order of magnitude in term of performance. On the other hand, the neural network model obtains a lower error than the others, so it seems to achieve the best predictions.

Finally, analysing the mean absolute error, it could be concluded that MLP stands out from the rest of the models in terms of prediction accuracy. It can also be observed that GBTR continues to have a significantly higher error than other algorithms, but the difference turns out to be lower in percentage terms with respect to the previous case.

Based on the results, it could be concluded that MLP is the algorithm with the best performance in all the indices that have been considered: the value of the R2 coefficient establishes that there is a notable degree of fit to the data given. As for the mean square error and mean absolute error metrics, they indicate that there are only minor differences between the samples in the database and the predicted data.

The results are aligned with previous papers in the literature that analyse the clothing insulation level through ML algorithms, in which it can be determined that neural networks, specifically convolutional networks, seem to be the models with the best performance for clo prediction (Ngarambe et al., 2019; Duhirwe et al., 2022).

4.2. Analysis of Variables Relevance

In the present section, an analysis of the relevance of the different inputs for the prediction of the output variable is shown. For this purpose, one technique consists of training the model with all the inputs and then sequentially eliminating the inputs one by one and observing how this affects the performance of the model: if the elimination of a variable leads to a significant decrease in the performance of the algorithm, this would suggest that said variable is relevant for the prediction. The present analysis only considered the MLP algorithm, since it obtained the best results in terms of performance.

It was observed that the elimination of certain variables ("TA", "ta", "rh", "wt", "age", "met", "climate") had in many cases similar orders of magnitude (about 13% loss of accuracy in the predictions), with none of them standing out significantly. Additionally, the elimination of "TP", "tr", "tg", "TSV", "ht" variables led to similar losses in prediction accuracy of about 10%.

Therefore, the previous analysis, individualized by variables, was complemented by defining different sets of variables. Table 4 shows the set of variables considered as well as how the elimination of each set affects the accuracy of the prediction of the MLP model (in percentage terms of R2 loss).

Input variables set	Description	Loss of accuracy
Comfort variables	Includes all comfort variables	19,26%
Temperatures	Includes all variables related to temperatures	30,01%
Indoor climatic variables	Includes indoor temperature variable, air velocity and relative humidity	33,98%
Outdoor climatic variables	Includes variables related to outdoor conditions	11,85%
Climatic variables	Includes indoor climatic variables and outdoor climatic variables	37,57%
Personal factors	Includes variables related to the participants	26,13%
Season_Climate	Includes variables related to season and climate	16,49%

Table 4. Model accuracy and input variables set

In this analysis, there are clear differences between the different sets, concluding that the relevance of "Climatic variables" is especially relevant when developing a model for the prediction of clo, in particular, "Temperatures" variables. Likewise, "Personal factors" and "Comfort variables" also seem to have an important influence.

5. Conclusions

In the present study, a method for a comparative analysis of four ML algorithms for the prediction of the clothing insulation level using the ASHRAE Global Thermal Comfort Database II has been described. In addition, a study on the relevance of the input variables considered in this prediction has also been conducted. The most relevant findings are highlighted below:

- Among the four algorithms considered, the one that obtained the best results in terms of performance is the model based on neural networks.
- MLP algorithm have predicted the clothing insulation level with a correlation coefficient square of 0,883, a mean square error of 0,009 and a mean absolute error 0,067.
- As for the relevance of the input variables, although the elimination of thermal acceptability, air temperature, relative humidity, weight, age, met and climate variables showed similar loss of accuracy for the performance of MLP algorithm, none of them stood out significantly.
- Analysing the relevance of input variables, not individualized but in in terms of set, "Climatic variables" (which comprises indoor and outdoor climatic variables) had the greatest effect on the prediction of the clothing insulation level.

In future studies, it would be interesting to delve into different contexts, that is, to analyse the dataset distinguishing, for example, by countries or climates. Moreover, new variables could be introduced into the models, such as emotional factors or environmental control elements, which would lead to a more focused approach of individualized comfort. Finally, the present study performed a comparative analysis between four ML algorithms, so it could be extended by considering other models, for example, those based on deep learning and reinforced learning.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This work was granted by TED2021-130659B-I00, funded by MICIU/AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR, and 13_ US_02, funded by the Instituto Andaluz de Prevención de Riesgos Laborales (IAPRL).

References

ANSI/ASHRAE (2020). *ANSI/ASHRAE Standard 55-2020 Thermal Environmental Conditions for Human Occupancy.*

De Carli, M., Olesen, B.W., Zarrella, A., & Zecchin, R. (2007). People's clothing behaviour according to external weather and indoor environment. *Building and Environment,* 42(12), 3965-3973. https://doi.org/10.1016/j.buildenv.2006.06.038

De Dear, R., & Brager, G. (1998). Developing an adaptive model of thermal comfort and preference. *ASHRAE Transactions,* 104(1), 145-167.

De Dear, R. (1998). A Global Database of Thermal Comfort Field Experiments. *ASHRAE Transactions,* 104(1141).

Duhirwe, P.N., Ngarambe, J., & Yun, G.Y. (2022). Geographically extended occupant clothing behavior model using convolutional neural networks with hyperband technique. *Journal of Building Engineering,* 49, 104023. https://doi.org/10.1016/j.jobe.2022.104023

Fanger, P. (1970). *Thermal comfort. Analysis and applications in environmental engineering.* Copenhagen: Danish Technical Press.

Farhan, A.A., Pattipati, K., Wang, B., & Luh, P. (2015). Predicting individual thermal comfort using machine learning algorithms. In *IEEE International Conference on Automation Science and Engineering* (708-713). https://doi.org/10.1109/CoASE.2015.7294164

Földváry-Ličina, V., Cheung, T., Zhang, H., De Dear, R., Parkinson, T., Arens, E. et al. (2018). Development of the ASHRAE Global Thermal Comfort Database II. *Building and Environment,* 142, 502-512. https://doi.org/10.1016/j.buildenv.2018.06.022

Gagge, A.P., Burton, A.C., & Bazett, H.C. (1941). A practical system of units for the description of the heat exchange of man with his environment. *Science,* 94(2445), 428-430. https://doi.org/10.1126/science.94.2445.428

Haldi, F., & Robinson, D. (2011). Modelling occupants' personal characteristics for thermal comfort prediction. *International Journal of Biometeorology,* 55, 681-694. https://doi.org/10.1007/s00484-010-0383-4

Humphreys, M.A. (1977). A study of the thermal comfort of primary school children in summer. *Building and Environment,* 12(4), 231-239. https://doi.org/10.1016/0360-1323(77)90025-7.

Humphreys, M.A. (1979). The influence of season and ambient temperature on human clothing behaviour. In *Indoor Climate* (699-713).

ISO (2009). *E.C. for Standardization, ISO 9920: Ergonomics of the thermal environment - Estimation of thermal insulation and water vapour resistance of a clothing ensemble.*

Luo, M., Xie, J., Yan, Y., Ke, Z., Yu, P., Wang, Z. et al. (2020). Comparing machine learning algorithms in predicting thermal sensation using ASHRAE Comfort Database II. *Energy and Buildings,* 210, 109776. https://doi.org/10.1016/j.enbuild.2020.109776

McCartney, K.J., & Nicol, J.F. (2002). Developing an adaptive control algorithm for Europe. *Energy and Buildings,* 34, 623-635. https://doi.org/10.1016/S0378-7788(02)00013-0

Mitchell, T. (1997). *Machine Learning.* McGraw Hill.

Morgan, C., & De Dear, R. (2003). Weather, clothing and thermal adaptation to indoor climate. *Climate Research,* 24(3), 267-284. https://doi.org/10.3354/cr024267

Müller, A.C., & Guido, S. (2016). *Introduction to machine learning with Python: a guide for data scientists.* O'Reilly Media, Inc.

Ngarambe, J., Yun, G.Y., & Kim, G. (2019). Prediction of indoor clothing insulation levels: A deep learning approach. *Energy and Buildings,* 202, 109402. https://doi.org/10.1016/j.enbuild.2019.109402

Nicol, F., & Humphreys, M. (2010). Derivation of the adaptive equations for thermal comfort in free-running buildings in European standard EN15251. *Building and Environment,* 45(1), 11-17. https://doi.org/10.1016/j.buildenv.2008.12.013

Nicol, F., & Roaf, S. (1996). Pioneering new indoor temperature standards: The Pakistan project. *Energy and Buildings,* 23(3), 169-174. https://doi.org/10.1016/0378-7788(95)00941-8

Qavidel-Fard, Z., Zomorodian, Z.S., & Korsavi, S.S. (2022). Application of machine learning in thermal comfort studies: A review of methods, performance and challenges. *Energy and Buildings,* 256, 111771. https://doi.org/10.1016/j.enbuild.2021.111771

Su, X., Wang, Z., Zhou, F., Duanmu, L., Zhai, Y., Lian, Z. et al. (2022). Comfortable clothing model of occupants and thermal adaption to cold climates in China. *Building and Environment,* 207, 108499. https://doi.org/10.1016/j.buildenv.2021.108499

Wu, Z., Li, N., Peng, J., Cui, H., Liu, P., Li, H. et al. (2018). Using an ensemble machine learning methodology-Bagging to predict occupants' thermal comfort in buildings. *Energy and Buildings,* 173, 117–127. https://doi.org/10.1016/j.enbuild.2018.05.031