

COMPARISON OF THERMAL COMFORT PARAMETERS USING DIFFERENT FEATURE SELECTION ALGORITHMS

Resul ÖZLÜK^{1*}, Türkan GÖKSAL ÖZBALTA²

¹ Balıkesir University Faculty of Architecture Department of Architecture, Balıkesir,

ORCID No : <https://orcid.org/0000-0001-8309-2980>

² Ege University, Faculty of Civil Engineering, İzmir, ORCID No : <https://orcid.org/0000-0001-5195-0741>

Keywords	Abstract
<p><i>Thermal Comfort</i> <i>AI Algorithms</i> <i>Educational Building</i> <i>Environmental Parameters</i> <i>Feature Selection</i></p>	<p><i>This study aims to examine the environmental parameters affecting user thermal comfort in an amphitheater belonging to an educational building and to compare the effects of these parameters using different feature selection algorithms. Thermal comfort indicators PPD and environmental parameters were calculated using measurements from the Testo 480 device. Subjective data were obtained through surveys measuring users' thermal acceptability and thermal comfort perception. The RandomForestRegressor, SHAP, and CorrelationAttributeEval algorithms were used to compare the order of influence of environmental parameters affecting students' thermal comfort. The results showed that surface temperature and indoor air temperature are the most influential parameters on user comfort. In the SHAP and CorrelationAttributeEval algorithms, surface temperature was identified as the most influential parameter, while in the RandomForestRegressor algorithm, indoor temperature was identified as the most significant parameter. Additionally, when compared with survey results, the environmental parameter with the least effect on user comfort was found to yield the same result as the CorrelationAttributeEval algorithm. These findings provide important insights into better interpreting the factors affecting user comfort and optimizing thermal comfort in similar spaces.</i></p>

FARKLI ÖZELLİK SEÇİMİ ALGORİTMALARI KULLANILARAK ISIL KONFOR PARAMETRELERİNİN KARŞILAŞTIRILMASI

Anahtar Kelimeler	ÖZ
<p><i>Isıl konfor</i> <i>Yapay zekâ algoritmaları</i> <i>Eğitim binası</i> <i>Çevresel parametreler</i> <i>Özellik seçimi</i></p>	<p><i>Bu çalışma, bir eğitim binasına ait amfi mekânında kullanıcı ısı konforunu etkileyen çevresel parametreleri incelemek ve bu parametrelerin etkilerini farklı özellik seçimi algoritmalarıyla karşılaştırmayı amaçlamaktadır. Isıl konfor göstergelerinden PPD ve onu etkileyen çevresel parametreler, Testo 480 cihazının ölçümleriyle hesaplanmıştır. Öznel veriler, kullanıcıların ısı kabul edilebilirlik ve ısı konfor algısını ölçen anketlerle elde edilmiştir. Öğrencilerin ısı konforunu etkileyen çevresel parametrelerin etki sırasının karşılaştırılmasında Random Forest Regressor, SHAP ve Correlation Attribute Eval algoritmaları kullanılmıştır. Sonuçlar, yüzey sıcaklığı ve iç ortam sıcaklığının kullanıcı konforu üzerinde en etkili parametreler olduğunu göstermiştir. SHAP ve Correlation Attribute Eval algoritmalarında yüzey sıcaklığı, Random Forest Regressor algoritmasında ise iç ortam sıcaklığı en etkili parametre olarak belirlenmiştir. Ayrıca anket sonuçları ile karşılaştırıldığında kullanıcı konforu üzerinde en az etkili çevresel parametrenin Correlation Attribute Eval algoritması ile aynı sonucu verdiği görülmüştür. Bu bulgular, kullanıcı konforunu etkileyen faktörlerin daha iyi yorumlanmasına ve benzer mekanların ısı konfor açısından optimize edilmesine yönelik önemli sonuçlar sağlamaktadır.</i></p>

Araştırma Makalesi

Research Article

Başvuru Tarihi : 11.06.2025

Submission Date : 11.06.2025

Kabul Tarihi : 19.11.2025

Accepted Date : 19.11.2025

* Sorumlu yazar: resulozluk7@gmail.com

<https://doi.org/10.31796/ogummf.1717391>



Bu eser, Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) hükümlerine göre açık erişimli bir makaledir.

This is an open access article under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years, advanced Heating, Ventilating and Air Conditioning (HVAC) systems have become widely used in buildings thanks to technological developments. The purpose of HVAC systems is to provide users with a comfortable and clean indoor environment. Especially with the impact of the pandemic, improving indoor air quality has become even more important for user comfort (Medved et al., 2019). Indoor comfort is related to user satisfaction in terms of thermal and visual conditions within a building (Zeiler and Boxem, 2009). Therefore, examining the thermal comfort perceptions of users who spend most of their time in educational buildings is an important research area.

Thermal comfort depends on many parameters, such as human dimensions, gender, and age. However, in general, the parameters affecting thermal comfort can be classified as personal and environmental parameters. Environmental factors are categorized as air temperature, air movement, air humidity, and mean radiant temperature, while personal factors can be categorized as clothing condition (Clo) and metabolic rate (Szokolay, 2004). There are two main approaches commonly used to evaluate user thermal comfort influenced by environmental and personal parameters. The first of these approaches, the Heat Balance Model, was developed by Fanger. The Heat Balance Model approach is suitable for evaluating users in air-conditioned buildings where they have no control over the environment. In this model developed by Fanger, the comfort level of users can be measured using the Predicted Mean Vote (PMV) and Percentage of Dissatisfied People (PPD) metrics (Fanger, 1970). PMV has been the most widely used index for estimating thermal comfort (Park and Woo, 2023). The Adaptive Thermal Comfort Model can be used as an evaluation method in spaces that are naturally ventilated and interact with the environment (Liang et al., 2012). In ASHRAE 55-2017, this model defines acceptable thermal environments for naturally ventilated spaces that do not require any system. (ANSI/ASHRAE Standard 55, 2017). Since the space examined in the study has a mechanical cooling and heating system, the heat balance model was used to evaluate user comfort.

In recent years, researchers have conducted various studies on thermal comfort in educational buildings. (De Giuli et al., 2012) studied indoor air quality in educational buildings, (Aparicio-Ruiz et al., 2021) investigated the effects and comparisons of different thermal comfort models on user comfort, Heracleous and Michael (2020), examined the effects of natural ventilation on user comfort, Katafygiotou and Serghides (2014), assessed the adequacy of different environmental parameters affecting indoor thermal comfort and (Li et al., 2021) investigated the balance between a building's energy use and thermal comfort.

Thermal comfort prediction is important for building energy efficiency and user comfort. Different AI methods are commonly used to predict user comfort. (Bai et al., 2025) proposed hybrid feature selection methods based on a thermal preference prediction model. Rahmanparast et al. (2025) used different machine learning algorithms to predict PMV, one of the indicators of the Fanger method, and different feature selection methods to determine which parameters are effective on PMV. The results of the study showed that the different machine learning algorithm models used had higher accuracy than traditional models. Gao et al. (2025) used deep transfer learning to reduce the impact of insufficient data in thermal comfort modeling. Pearson correlation coefficients and deep reinforcement learning (DRL) were used to identify parameters affecting user comfort. The results indicate that accuracy exceeding 55% can be achieved with limited data in target buildings. Liu et al. (2025) applied the K-means algorithm to create a thermal comfort clustering analysis model and feature selection to improve data quality. The results show that a model with over 90% accuracy can be created using only effective factors and that the proposed model can reliably predict thermal comfort. Park and Woo (2023), applied machine learning methods to provide faster PMV calculations than traditional methods. In the study, comparisons were made between different models using Principal Component Analysis (PCA), Best Subset, and Gini Importance as the most effective features for PMV. The results show that PCA and Artificial Neural Network (ANN) have the highest accuracy. When reviewing studies conducted on this topic in general, it is clear that the importance of feature selection prior to modeling in terms of the accuracy and reliability of models has not been fully understood, and there are very few studies on this topic in the literature (Bai et al., 2025). In particular, there are limited studies comparing subjective and objective data in the studies.

The aim of this study is to compare the effects of environmental parameters affecting the thermal comfort conditions of students studying in an amphitheater space of a university building using different feature selection algorithms. The study examined the effects of parameters affecting students' thermal comfort using both objective and subjective approaches. While experimental measurements provide an objective approach based on standards, the use of questionnaires provides a subjective approach to thermal comfort.

2. Method

This study used both objective and subjective methods to evaluate the environmental parameters affecting students' thermal comfort conditions in an amphitheater belonging to a university building. First, subjective data related to thermal comfort were

collected through questionnaires administered to students. Objective data were collected using the Testo-480 measurement device. The measured environmental parameters included temperature, relative humidity (Bnem), air velocity (Va), carbon dioxide (CO₂) concentration in the environment, and surface temperatures. The collected datasets were analyzed using different feature selection algorithms (Random Forest Regressor, SHAP and Correlation Attribute Eval), and the effects of environmental parameters affecting user comfort were compared. The results were analyzed together to evaluate the consistency between subjective and objective approaches and to identify the most effective environmental parameters affecting thermal comfort.

2.1. Surveys and Measurements

Table 1 shows the devices used to measure environmental parameters and their characteristics. Measurement equipment was appropriately prepared in the working environment to continuously measure indoor air temperature, relative humidity, CO₂ concentration, air velocity, and globe temperature.

The amphitheater where measurements and surveys were conducted for this study is located on the ground floor of the education building (Figure 1). The amphitheater has a trapezoidal plan. A conventional air conditioning unit, added later, is used for heating and cooling. The seating arrangement consists of 300 seats arranged on a sloping platform. The seating is arranged in a semicircular shape, oriented toward the stage. This arrangement provides optimal viewing angles and sound distribution. Participants in the survey were university students aged 19-27. A total of 192 surveys were administered to 50 different students across multiple course sessions. The same students attending the course participated in surveys administered at different times. Measurements were taken 5 minutes after the device was placed in the room, taking into account the time it took for the device to enter the regimen. Measurements were recorded for 15 minutes during the 90-minute class period, extending until the end of the lesson. Furthermore, these measurements were taken using a device placed in the middle of the classroom and 1.1 meters above the floor, as required by thermal comfort standards. To increase the reliability of the data, measurements and surveys were conducted simultaneously during the class period. Before the survey was administered, the necessary explanations were provided to ensure that the students could easily understand the questions.

Table 1. Equipment used in a field study

Parameters	Equipment
Indoor air temperature, CO ₂ concentration and humidity	IAQ probe (Testo 480)
Outdoor air temperature	HOBO external temp/rh data logger
Air velocity	Turbulence probe (Testo 480)
Globe thermometer	Black-ball thermometer (Testo 480)

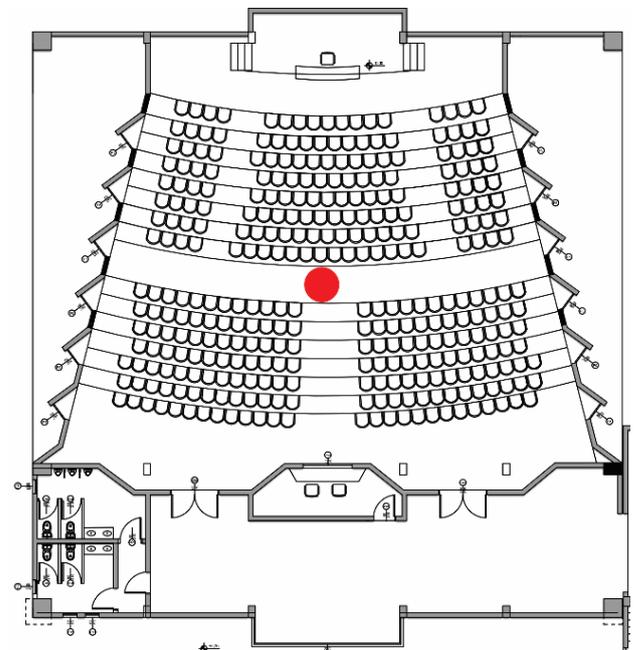


Figure 1. Study area and location of measurement equipment.

The first survey questions consisted of personal questions (gender, age). The next section asked students how they felt thermally, how they would like their environment to change thermally, and the ambient temperature, air velocity, and humidity (Appendix 1). The surveys consisted of five questions to assess occupants' perceptions of thermal comfort. The survey questions included 7- and 3-point Likert-type scales (ranging from -3 to +3, or "unacceptable" to "acceptable") to measure subjective assessments of environmental parameters such as temperature, air velocity, and relative humidity. Finally, a question asked which factor most influenced occupants' thermal comfort (Caner & Ilten, 2020). For participants'

metabolic rates and clothing conditions, values specified in ASHRAE 55 (ANSI/ASHRAE Standard 55, 2017) and ISO 7730 (Olesen and Parsons, 2002) were utilized. The ASHRAE standardized scale levels for evaluating survey

questions are listed in Table 2. This research was conducted in accordance with academic research and publication ethics rules.

Table 2. Scales used in the questionnaire survey.

Parameter	Scale						
Thermal sensation vote (TSV)	-3	-2	-1	0	+1	+2	+3
	cold	cool	A bit cool	neutral	A bit warm	warm	hot
Thermal acceptability (TA)		1		2		3	
		Acceptable		neutral		Unacceptable	
Air velocity sensation (AVS)	-3	-2	-1	0	+1	+2	+3
	Very still	Moderately still	Slightly still	Neutral	Slightly moving	Moderately moving	Much moving
Humidity sensation (HS)	-3	-2	-1	0	+1	+2	+3
	Very dry	Moderately dry	Slightly dry	Neutral	Slightly humid	Moderately humid	Very humid

Within the scope of the study, the Predicted Mean Vote (PMV) and Percentage of Dissatisfied People (PPD) models were used as thermal comfort indicators. PMV can be determined by the equation (Fanger, 1970).

$$PMV = (0.303e^{-0.036M} + 0.028)\{(M - W) - 3.96 \times 10^{-8} f_{cl} [(t_{cl} + 273)^4 - (\bar{t}_r + 273)^4] - f_{cl} h_c (t_{cl} - t_a)\} - 3.05 [5.733 - 0.007(M - W) - P_a] - 0.42 [(M - W) - 58.15] - 0.0173M(5.867 - P_a) - 0.0014M(34 - t_a) \} \tag{1}$$

$$t_{cl} = 35.7 - 0.028 \times (M - W) - I_{cl} \times \{3.96 \times 10^{-8} \times f_{cl} \times [(t_{cl} + 273)^4 - (\bar{t}_r + 273)^4] + f_{cl} \times h_c \times (t_{cl} - t_a)\} \tag{2}$$

Where, M is Metabolic rate of occupant, W is mechanical power by occupant, f_{cl} ; (surface area of body with clothes), t_{cl} ; is Clothing surface temperature (°C), T_a is temperature of air (°C), t_r ; is Mean radiant temperature (°C), P_a is partial pressure of water vapor.

The PPD equation is calculated as follows:

$$PPD = 100 - 95e^{(-0.03353PMV^4 - 0.2179PMV^2)} \tag{3}$$

2.2. Feature Selection Algorithms

Random Forest Regressor: Random Forest Regression is widely used in statistical learning and machine learning applications. It is a powerful learning algorithm based on the principles of random feature selection. This method provides high accuracy and reliability rates,

especially in classification and regression analyses. In classification problems, the model's output is determined based on the majority vote of individual trees (Z. Wang et al., 2018). In regression problems, the model's output is obtained by taking the simple arithmetic mean or weighted average of the predictions produced by all decision trees in the forest for the continuous target variable. Each tree independently produces its own prediction, and the model's final prediction is calculated as the arithmetic mean or weighted average of these predictions. This approach contributes to reducing overfitting and producing more robust predictions (Smith et al., 2013).

SHAP Algorithm: It is a method used to interpret the output of machine learning models. SHAP is based on cooperative game theory and the concept of Shapley values (Cui et al., 2024). Current machine learning models often create difficulties for researchers in interpreting model predictions and function as black box models. Unlike other methods, SHAP provides us with a detailed explanatory model of how each feature contributes to the predictions (Meddage et al., 2022). SHAP shows the importance of each feature in the predictions and helps us understand how the input features affect the predictions (Huang et al., 2020).

Correlation Attribute Eval: It evaluates the value of a trait by measuring the correlation between the trait and the class. The correlation coefficient indicates the direction and strength of the linear relationship between two random traits. This coefficient takes a value between -1 and +1 (Colgan et al., 2019).

3. Results and Discussion

3.1. Measurement and Survey Results

HOBO data loggers were used for outdoor environmental data affecting user comfort, and Testo 480 instruments were used for indoor environmental data. In this study, the relationship between the environmental parameters collected for the amphitheater and the thermal comfort indicators is presented in Figure 2.

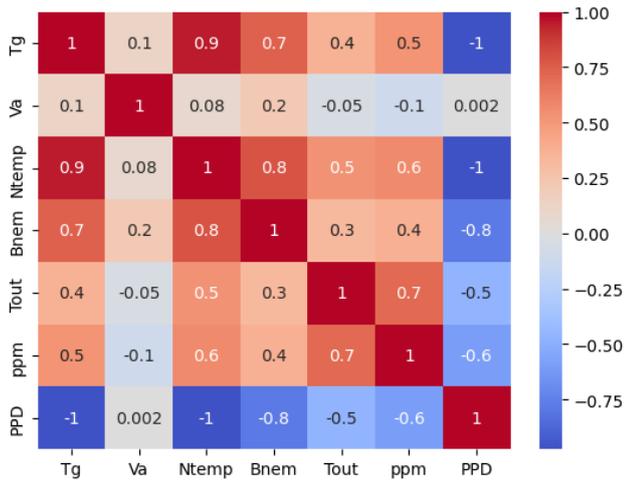


Figure 2. The relationship between PPD and other environmental parameters.

Data collected from the education building during the winter period shows the relationship between PPD (Predicted Percentage of Dissatisfied), a thermal comfort indicator, and various environmental parameters. Limit values for PPD are defined in the ISO 7730 (Olesen and Parsons, 2002) and EN 15251 (CEN, 2007) standards. According to these standards, a PPD value of 10% or less indicates that occupants are satisfied with their environment. However, this study found an average PPD value of 25%, which is insufficient for occupants' thermal comfort. However, according to the ISO 7730 standard, it corresponds approximately to Category C comfort and represents a borderline acceptable indoor environment. There is a high negative correlation between indoor globe temperature (Tg) and PPD. The high correlation indicates that an increase in surface temperature significantly reduces PPD (improves thermal comfort). An increase in surface temperature during the winter months may contribute to the environment feeling more comfortable. The correlation between Va and PPD is negligible at 0.002, indicating that airflow velocity has almost no effect on thermal comfort in this space. This can be explained by the limited air circulation during winter months (windows and doors kept closed). The high level of correlation between indoor air temperature (Ntemp) and PPD indicates that an increase in indoor

temperature improves thermal comfort. This result confirms that a higher indoor temperature during winter months reduces discomfort. The high negative correlation between Bnem and PPD indicates that as humidity increases, dissatisfaction decreases. Low humidity during winter months can make the environment feel colder; therefore, increased humidity can be considered a factor that improves comfort. These results are consistent with the findings of studies on the relationship between environmental factors and thermal comfort. The negative correlation between ppm and PPD indicates that an increase in CO₂ concentration reduces discomfort. However, high CO₂ levels typically lead to reduced air quality and decreased comfort. This result may be influenced by interactions with other environmental parameters affecting this relationship.

3.1.1. Subjective evaluation

The analysis of thermal sensation ratings (TSVs) is shown in Figure 3. TSVs were calculated by asking students, "How do you feel the temperature of your environment?" using the ASHRAE seven-point scale. Of the 192 surveys conducted, 121 (62%) were within the comfort range (between -1 and +1). Approximately 35% of users felt the environment was cold, while the remaining users felt it was warm.

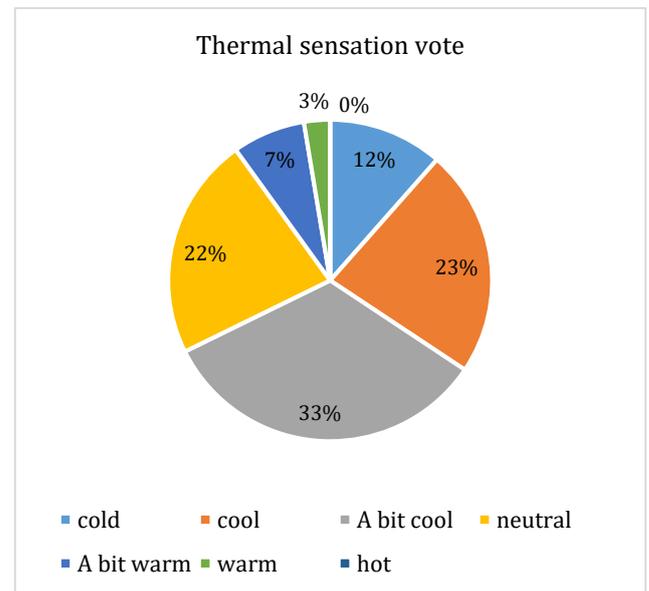


Figure 3. Distribution of thermal sensation responses during measurement.

This study focuses on the question of which factor most affects the space where users are located. Statistical information for this question is provided in Table 3. The mean value of 1.55 represents the average of students' responses regarding the factor that most impacts the environment. Temperature was the first option in this question. This value indicates that temperature is the

factor that most significantly impacts users' thermal comfort. The 95% confidence interval for the mean lies between 1.42 and 1.70. This range provides a statistically reliable estimate (Morgan et al., 2004). In addition, it is observed that the average ratings of Air velocity sensation (AVS) and Humidity sensation (HS) in

the study area are close to neutral and the students feel comfortable in terms of humidity and air exchange in the area.

Table 3. Descriptive statistics of sensation variables.

		Statistic	Std.Error	Bias	Std.Error	95% Confidence Interval	
						Lower	Upper
TSV	Mean	-1.01	.08	.00	.08	-1.17	-.84
	Std. Deviation	1.20		-.01	.05	1.08	1.30
	Variance	1.44		-.01	.13	1.18	1.70
Factor	Mean	1.55	.07	-.00	.07	1.41	1.70
	Std. Deviation	1.05		-.00	.06	.91	1.16
	Variance	1.10		-.00	.13	.84	1.36
AVS	Mean	-.09	.07	.00	.07	-.23	.05
	Std. Deviation	.99		-.01	.07	.82	1.13
	Variance	.99		-.01	.15	.68	1.29
HS	Mean	-.26	.06	.00	.06	-.38	-.13
	Std. Deviation	.87		-.00	.05	.75	.98
	Variance	.77		-.00	.10	.57	.97
Valid N (listwise)	N	192		0	0	192	192

The factor affecting thermal comfort was measured based on the answers to the question, "What is the most influential factor in your environment?". It has been found that the temperature factor has the greatest effect on thermal comfort (75%). This result shows that temperature is the most effective parameter among environmental parameters in determining thermal comfort perception (Figure 4). This result shows that people's thermal perception is more sensitive to temperature changes than to other parameters.

Additionally, factors such as humidity, air speed, and lighting are seen to be less effective compared to temperature. The low selection of air speed suggests that participants did not perceive air movement as a critical factor as temperature. These results are consistent with the low correlation between air speed and PPD observed in Figure 2. These findings suggest that more focus should be placed on the temperature factor in the design of heavily used spaces in educational buildings.

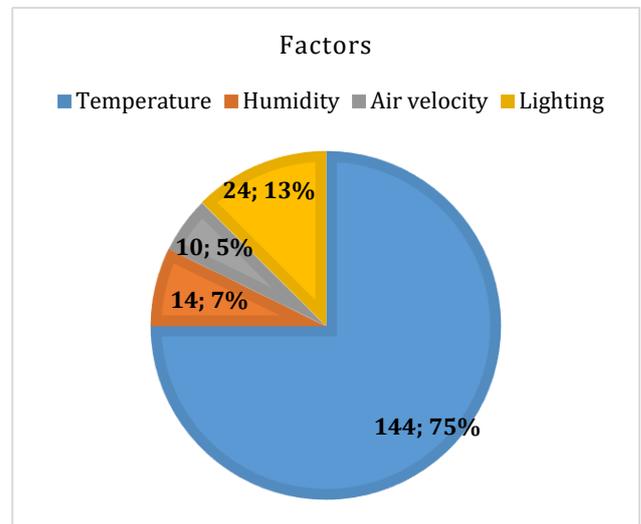


Figure 4. Distribution of factors affecting user comfort during measurements.

3.2. Feature Selection Algorithm Results and Comparison

Within the scope of the study, the effects of environmental parameters affecting user comfort were compared and analyzed using different feature selection algorithms such as Random Forest Regressor, SHAP, and Correlation Attribute Eval.

Table 4 shows the effect of each environmental parameter on each algorithm and its importance coefficient. The comparison between different feature selection algorithms is based on the relative importance rankings of the variables, not on the absolute magnitudes of the coefficients. In the SHAP and Correlation Attribute Eval algorithms, the most effective parameter was found to be surface temperature, while in the Random Forest Regressor algorithm, it was indoor temperature. In the SHAP and Random Forest

Regressor algorithms, the least effective parameter was outdoor temperature, while in the Correlation Attribute Eval algorithm, it was air velocity.

A general examination of the algorithm results reveals that the most influential environmental parameters are indoor temperature and surface temperature. The result obtained for the most influential parameter was found to be identical to the subjective (survey) results (Temperature: 75%). This result suggests that the algorithms selected for the most influential parameter performed successfully.

Furthermore, when compared with the survey results, it was observed that the environmental parameter with the least impact on user comfort (Air velocity: 5%) yielded the same ranking as the Correlation Attribute Eval algorithm.

Table 4. Comparison of results from different algorithms.

Random Forest Regressor	SHAP	Correlation Attribute Eval
Ntemp 0.22	Tg 17.59	Tg -0.97
Tg 0.21	Ntemp 1.66	Ntemp -0.95
Bnem 0.18	Va 0.92	Bnem -0.78
ppm 0.02	Bnem 0.21	ppm -0.57
Va 0.00	ppm 0.11	Tout -0.46
Tout 0.00	Tout 0.05	Va 0.01

Note: Tg: Indoor globe temperature, ppm: CO₂ concentration, Tout: Outdoor air temperature, Ntemp: Indoor air temperature, Bnem:Relative humidity, Va: Air velocity.

An examination of Figure 5 reveals that the Tg variable has the greatest impact on the model output. This demonstrates that ambient temperature is a determining factor in occupants' thermal sensation. While the Ntemp parameter has a moderate impact, the other parameters, Va, Bnem, ppm, and Tout, have a more limited impact on the model output.

Furthermore, the SHAP algorithm observed that the ranking according to the most influential parameter differed when compared to the subjective results. While air velocity (5%) was the parameter that least affected user comfort in the subjective results, the SHAP algorithm found relative humidity to be less influential than this parameter.

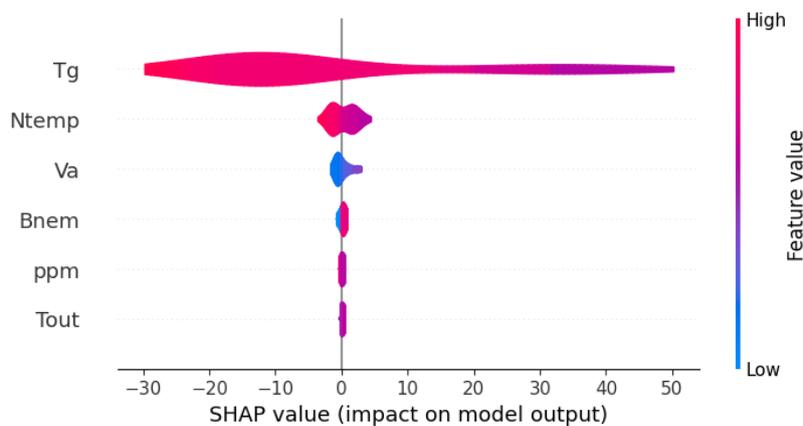


Figure 5. Effect of environmental parameters of the SHAP algorithm on user comfort.

The spread of points on the x-axis for each feature shows the range and distribution of SHAP values across different parameters (Figure 6).

The wide spread between parameter values indicates varying levels of importance for that feature across the dataset. According to Figure 6, the most significant factor is Tg and the least significant factor is Tout. In conclusion, different feature selection methods have

their advantages and limitations. Rather than evaluating the findings based on a single algorithm, this study presents comparative findings on thermal comfort parameters using different feature selection methods. SHAP is a powerful tool for explaining the contributions of features to a model.

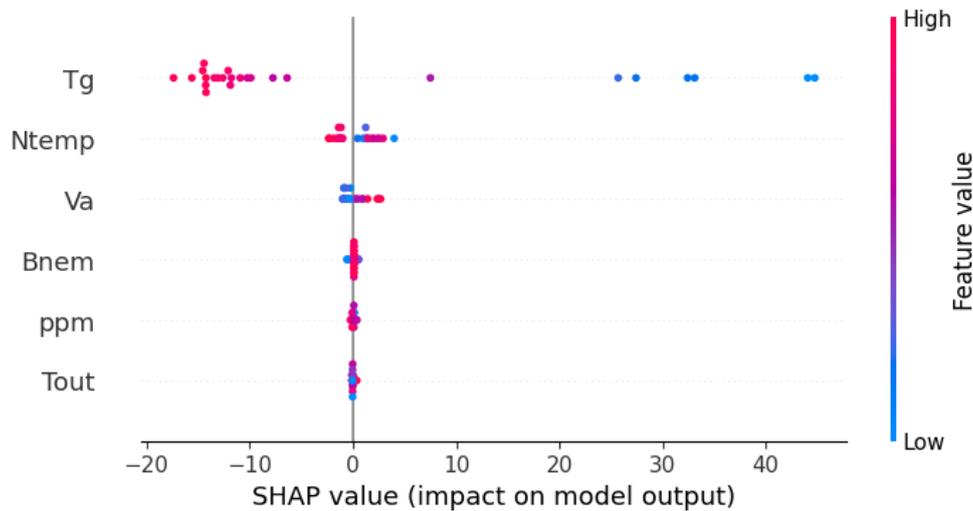


Figure 6. Effect of environmental parameters on user comfort based on SHAP algorithm results.

However, most machine learning models are difficult to interpret and understand due to their black-box (unexplainable) nature (M. Wang et al., 2023). To ensure interpretability, the SHAP algorithm, which incorporates model complexity and feature interactions, is frequently used in the literature (Meddage et al., 2022). Therefore, the SHAP algorithm was used in this study to interpret and visualize the effects of features. CorrelationAttributeEval quickly and effectively analyzed linear relationships. Random Forest Regressor is a powerful machine learning algorithm that can be used in both classification and regression problems. It combines multiple decision trees to increase the accuracy of predictions and handle missing data (Alsahaf et al., 2022). This study demonstrates that different feature selection algorithms can lead to different interpretations of the same dataset. For example, the air velocity parameter was ranked among the top three variables in the SHAP analysis, while it was ranked lowest in the CorrelationAttributeEval algorithm. This suggests that SHAP better captures nonlinear effects and that the algorithm used in thermal comfort analysis can create significant differences in interpretation. Similarly, the most influential factors differ in the importance rankings of the SHAP algorithm and the Random Forest Regressor algorithm. The differences in importance rankings across studies

provide practical insights into the impact of the feature selection method on the thermal comfort prediction model results.

4. Conclusion

This study presents findings on thermal comfort parameters using different feature selection algorithms rather than a single feature selection algorithm. The effects of environmental parameters affecting occupant comfort were analyzed in detail. The analysis compared the performance of different algorithms based on both objective and subjective data. The results obtained from all algorithms indicate that indoor temperature and surface temperature are the most important parameters affecting occupant comfort. This parallelism between subjective comfort perceptions and the importance rankings based on feature selection algorithms reinforces the consistency of the findings. These results demonstrate that the selected algorithms are effective in determining the importance rankings of environmental parameters affecting occupant comfort. However, they show that the SHAP algorithm's approach to evaluating the parameter effects offers a different perspective from other methods. While the findings of this study contribute to the studies in the literature, it also has some limitations. The study focuses on a single space of an educational building using only data from the winter season. The limited number of surveys used for data

analysis may limit generalizations for the results of different algorithms. Furthermore, this study focuses on three different feature selection algorithms, while other feature selection methods are also available in the literature. These limitations may change the order of importance of the parameters. Another limitation of this study is that the measurements were conducted at 15-minute intervals. This timeframe may not fully reflect long-term thermal changes, especially in winter

conditions. Future research could; It is recommended that studies focus on spaces with different orientations, different types of buildings (such as offices, hospitals, etc.), long-term studies involving different climatic conditions, and spaces using different air-conditioning systems. Thus, it will be important to understand how parameter importance rankings vary depending on these differences.

Appendix 1. Survey questions asked to students within the scope of the study.

Florr/Class:	Date :
Gender: Female <input type="checkbox"/> Male <input type="checkbox"/>	Age :
Please select the option that comes closest to your answer in the questions.	
1-How do you perceive the ambient temperature while you are in the environment? Very Hot (+3) <input type="checkbox"/> Hot (+2) <input type="checkbox"/> Slightly Warm (+1) <input type="checkbox"/> Neutral (0) <input type="checkbox"/> Slightly Cool (-1) <input type="checkbox"/> Cool (-2) <input type="checkbox"/> Very Cool (-3) <input type="checkbox"/>	
2- What is the temperature of your environment, as per your personal preference? Unacceptable <input type="checkbox"/> Just as I prefer <input type="checkbox"/> Acceptable <input type="checkbox"/>	
3- How do you evaluate the air speed in your environment? Much moving (+3) <input type="checkbox"/> Moderately moving (+2) <input type="checkbox"/> Slightly moving (+1) <input type="checkbox"/> Neutral (0) <input type="checkbox"/> Slightly Gentle (-1) <input type="checkbox"/> Moderately still (-2) <input type="checkbox"/> Very still (-3) <input type="checkbox"/>	
4- How would you rate the humidity of your environment? Very Humid (+3) <input type="checkbox"/> Humid (+2) <input type="checkbox"/> Slightly Humid (+1) <input type="checkbox"/> Neutral (0) <input type="checkbox"/> Slightly Dry (-1) <input type="checkbox"/> Dry (-2) <input type="checkbox"/> Very Dry (-3) <input type="checkbox"/>	
5- Which factor do you think most affects your comfort in your environment? Temperature <input type="checkbox"/> Humidity <input type="checkbox"/> Air Velocity <input type="checkbox"/> Lighting <input type="checkbox"/>	

Thanks...

References

Alsahaf, A., Petkov, N., Shenoy, V., & Azzopardi, G. (2022). A framework for feature selection through boosting. *Expert Systems with Applications*, 187. <https://doi.org/10.1016/j.eswa.2021.115895>

ANSI/ASHRAE Standard 55-2017. (2017). *Thermal Environmental Conditions for Human Occupancy*. Available at: www.ashrae.org/technology

Aparicio-Ruiz, P., Barbadilla-Martín, E., Guadix, J., & Muñuzuri, J. (2021). A field study on adaptive

thermal comfort in Spanish primary classrooms during summer season. *Building and Environment*, 203. <https://doi.org/10.1016/j.buildenv.2021.108089>

Bai, Y., Dong, Z., & Liu, L. (2025). Hybrid feature selection-based machine learning methods for thermal preference prediction in diverse seasons and building environments. *Building and Environment*, 269. <https://doi.org/10.1016/j.buildenv.2024.112450>

- Caner, I., & Ilten, N. (2020). Evaluation of occupants' thermal perception in a university hospital in Turkey. *Proceedings of the Institution of Civil Engineers: Engineering Sustainability*, 173(8), 414–428. <https://doi.org/10.1680/jensu.19.00059>
- CEN, E. 15251:2007. (2007). *Indoor Environmental Input Parameters for Design and Assessment of Energy Performance of Buildings Addressing Indoor Air Quality, Thermal Environment, Lighting and Acoustics*.
- Colgan, S., Quille, K., Mchugh, S., & Vasic, J. (2019). *Predicting Student Success. Early for a VTOS Student*. Paper presented at the International Conference on Engaging Pedagogy (ICEP), University of Limerick, Ireland.
- Cui, X., Lee, M., Koo, C., & Hong, T. (2024). Energy consumption prediction and household feature analysis for different residential building types using machine learning and SHAP: Toward energy-efficient buildings. *Energy and Buildings*, 309. <https://doi.org/10.1016/j.enbuild.2024.113997>
- De Giuli, V., Da Pos, O., & De Carli, M. (2012). Indoor environmental quality and pupil perception in Italian primary schools. *Building and Environment*, 56, 335–345. <https://doi.org/10.1016/j.buildenv.2012.03.024>
- Fanger, P. O. (1970). *Thermal Comfort: Analysis and Applications in Environmental Engineering*. Danish Technical Press.
- Gao, Y., Fu, Q., Chen, J., & Liu, K. (2025). Deep transfer learning-based hybrid modelling method for individual thermal comfort prediction. *Indoor and Built Environment*. <https://doi.org/10.1177/1420326X251317447>
- Heracleous, C., & Michael, A. (2020). Thermal comfort models and perception of users in free-running school buildings of East-Mediterranean region. *Energy and Buildings*, 215. <https://doi.org/10.1016/j.enbuild.2020.109912>
- Huang, X., Kroening, D., Ruan, W., Sharp, J., Sun, Y., Thamo, E., Wu, M. and Yi, X. (2020). A survey of safety and trustworthiness of deep neural networks: Verification, testing, adversarial attack and defence, and interpretability. In *Computer Science Review*, 37. <https://doi.org/10.1016/j.cosrev.2020.100270>
- Katafygiotou, M. C., & Serghides, D. K. (2014). Thermal comfort of a typical secondary school building in Cyprus. *Sustainable Cities and Society*, 13, 303–312. <https://doi.org/10.1016/j.scs.2014.03.004>
- Li, Q., Zhang, L., Zhang, L., & Wu, X. (2021). Optimizing energy efficiency and thermal comfort in building green retrofit. *Energy*, 237. <https://doi.org/10.1016/j.energy.2021.121509>
- Liang, H. H., Lin, T. P., & Hwang, R. L. (2012). Linking occupants' thermal perception and building thermal performance in naturally ventilated school buildings. *Applied Energy*, 94, 355–363. <https://doi.org/10.1016/j.apenergy.2012.02.004>
- Liu, Y., Li, X., Sun, C., Dong, Q., Yin, Q., & Yan, B. (2025). An indoor thermal comfort model for group thermal comfort prediction based on K-means++ algorithm. *Energy and Buildings*, 327. <https://doi.org/10.1016/j.enbuild.2024.115000>
- Meddage, P., Ekanayake, I., Perera, U. S., Azamathulla, H. M., Said, M. A. M., & Rathnayake, U. (2022). Interpretation of Machine-Learning-Based (Black-box) Wind Pressure Predictions for Low-Rise Gable-Roofed Buildings Using Shapley Additive Explanations (SHAP). *Buildings*, 12(6). <https://doi.org/10.3390/buildings12060734>
- Medved, S., Domjan, S., & Arkar, C. (2019). *Springer Tracts in Civil Engineering Sustainable Technologies for Nearly Zero Energy Buildings Design and Evaluation Methods*. Springer Nature Switzerland AG. <https://doi.org/https://doi.org/10.1007/978-3-030-02822-0>
- Morgan, G. A., Leech, N. L., Gloeckner, G. W., & Barrett, K. C. (2004). *SPSS for Introductory Statistics* (Second edition). Psychology Press. <https://doi.org/10.4324/9781410610539>
- Olesen, B. W., & Parsons, K. C. (2002). Introduction to thermal comfort standards and to the proposed new version of EN ISO 7730. *Energy and Buildings*, 34(6), 537–548. [https://doi.org/10.1016/S0378-7788\(02\)00004-X](https://doi.org/10.1016/S0378-7788(02)00004-X)
- Park, K. Y., & Woo, D. O. (2023). PMV Dimension Reduction Utilizing Feature Selection Method: Comparison Study on Machine Learning Models. *Energies*, 16(5). <https://doi.org/10.3390/en16052419>
- Rahmanparast, A., Milani, M., Camci, M., Karakoyun, Y., Acikgoz, O., & Dalkilic, A. S. (2025). A comprehensive method for exploratory data analysis and preprocessing the ASHRAE database for machine learning. *Applied Thermal Engineering*, 273. <https://doi.org/10.1016/j.applthermaleng.2025.126556>
- Smith, P. F., Ganesh, S., & Liu, P. (2013). A comparison of random forest regression and multiple linear

- regression for prediction in neuroscience. *Journal of Neuroscience Methods*, 220(1), 85–91. <https://doi.org/10.1016/j.jneumeth.2013.08.024>
- Szokolay, S. V. (2004). *Introduction to architectural science the basis of sustainable design*. Architectural Press.
- Wang, M., Li, Y., Yuan, H., Zhou, S., Wang, Y., Adnan Ikram, R. M., & Li, J. (2023). An XGBoost-SHAP approach to quantifying morphological impact on urban flooding susceptibility. *Ecological Indicators*, 156. <https://doi.org/10.1016/j.ecolind.2023.111137>
- Wang, Z., Wang, Y., Zeng, R., Srinivasan, R. S., & Ahrentzen, S. (2018). Random Forest based hourly building energy prediction. *Energy and Buildings*, 171,11–25. <https://doi.org/10.1016/j.enbuild.2018.04.008>
- Zeiler, W., & Boxem, G. (2009). Effects of thermal activated building systems in schools on thermal comfort in winter. *Building and Environment*, 44(11), 2308–2317. <https://doi.org/10.1016/j.buildenv.2009.05.005>