Performance of Machine Learning Algorithms considering Spatial Effects Assessment for Indoor Personal Thermal Comfort in Air-Conditioned Workplace

Syafiq Asyraff Ahmad[1], *Sheikh Ahmad* Zaki [1] , *Azizul* Azizan[2] and *Noor Syazwanee* Md Taib[1]

[1]Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia.
[2]Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia.

Abstract. Personal comfort models were developed to circumvent most of the constraints imposed by the Predicted Mean Vote (*PMV*) and present adaptive models, which consider the average response of a large population. Although there has been a lot of research into new input features for personal comfort models, the spatial data of the building, such as windows, doors, furniture, walls, fans, and heating, ventilation, and air conditioning (*HVAC*) systems, (the location of its occupants with those elements), have not been thoroughly examined. This paper investigates the impact of the spatial parameter in predicting personal indoor thermal comfort using various machine learning approaches in air-conditioning offices under hot and humid climates. The Decision Tree, Random Forest, Support Vector Machine, K-Nearest Neighbour, and Neural Network were trained using a field study dataset that was done in nineteen office spaces yielding 628 samples from 42 occupants. The dataset is divided randomly into training and testing datasets, with a ratio of 80% and 20%. This study examines how well machine learning predicts personal thermal comfort with spatial data compared to without spatial data; where the spatial parameters have shown a significant influence on model prediction accuracies, Mean Absolute Error (*MAE*), and Root Mean Squared Error (*RMSE*). The result shows the average MAE is decreased by 10.6% with the Random Forest (RF) getting the most MAE reduction by 23.8%. Meanwhile, the average RMSE is reduced by 11.8% with the RF giving the most RMSE cutback by 30.6%. Consequently, the spatial effect analysis also determines which area of the room has cold or heat clusters area that affects thermal comfort that contributes to the design of sustainable buildings.

1 Introduction

Over the past few decades, the world has seen a sharp increase in the amount of electricity used for air conditioners in buildings, especially in the Association of Southeast Asian Nations (*ASEAN*) reaching over 80 TWh in 2020 [1]. The International Energy Agency (*IEA*) forecasts that the *ASEAN* will see increased air conditioner ownership due to the region's ongoing economic expansion and population growth. The transformation of tropical lands also contributes to the increase in indoor time when surface temperatures rise. Building design and operation are primarily driven by two factors, energy efficiency and thermal comfort [2]. Setting better efficiency cooling criteria by governments is a crucial and relatively easy step. However, efforts to conserve energy should not sacrifice the occupant's comfort.

Thermal comfort is a crucial factor in occupant productivity, health, and well-being [3,4]. It has been demonstrated that a rise in indoor temperature can make occupants feel less attentive, which lowers cognitive function [5]. Based on the American Society of Heating, Refrigerating, and Air-Conditioning Engineers (*ASHRAE*) [6] thermal comfort is a mental state that expresses contentment with the thermal environment and is quantified using subjective evaluation. Because of this, occupants must be contented with their temperature surroundings if they want to work comfortably.

The most common model that predicts thermal comfort is the predicted mean vote (*PMV*) and adaptive thermal comfort model. The *PMV* model developed by P. O. Fänger [7] is the most widely used indicator in the field of thermal comfort. It is based on experimental data collected in a steady-state climatic chamber with heat balance ideas. Fanger's concept creates a group thermal model but not a model for individual thermal comfort. A comfortable *PMV* index is one with 95% happy responders. The thermal sensation vote is an integer with an *ASHRAE* scale from −3 to +3. model. *PMV* model relies on personal factors such as clothing insulation, and metabolic rate, and environmental factors such as indoor temperature, humidity, radiant temperature, and air velocity. Another common model that was used in thermal comfort is the adaptive model. The adaptive model was created because of several studies showing that people actively interact with their surroundings. The model takes into consideration numerous psychological, physiological, and behavioural components as individuals regulate their temperature environment using various tactics including moveable windows, clothing, heaters, fans, and so forth.

Adaptive models are frequently applied in environments with natural ventilation.

The ability of the *PMV*s to give accurate results based on diverse factors and analyse individual variations in occupants' thermal preferences is one of its weaknesses. As a result, the machine learning (*ML*) approach has been used to investigate the prediction of personal thermal comfort due to its significant capacity for self-study, rapid computation, and sophisticated problem-solving. To predict personal thermal comfort more accurately, several features are examined, to unmask and quantify the differences between different individuals in an environment. This enables a better understanding of specific comfort needs and requirements.

Although there has been a lot of research into new input features for personal comfort models, the spatial data of the building, such as its windows, doors, furniture, walls, fans, and *HVAC* systems, as well as the location of its occupants with those elements, have not been thoroughly examined.

2 Literature Review

An essential component of building thermal management is the prediction of thermal comfort and subsequent environmental adjustment. Over the years, several thermal comfort prediction models have been created, some of which include machine learning in the prediction process. The support vector machine (*SVM*) approach, for instance, was applied by the researchers [8] to the RP-884 thermal comfort database to develop a unique model with self-learning and self-correction capabilities. The properties of the *SVM* algorithm and the sample distribution characteristics of the RP-884 were used in the study to estimate the applicability range. The model might considerably reduce the previous models' errors. Comparing the new model to the *PMV* model, the sum of squares for residuals (*SSE*) was reduced by 96.4%, while the fitting degree (Rnew) increased by 83.7%.

[9] adopted four common classification algorithms including logistic regression, k-nearest neighbour, support vector machine, and random forest. Given the data collected in the case study, the random forest classifier produces the highest classification accuracy. [10] described an artificial neural network (*ANN*) with a momentum function technique to efficiently tackle the problem of indoor thermal comfort prediction. To enhance classification performance and decrease mean square error (*MSE*), the innovative swarm algorithm (*CSO*) was used, which automatically produces the most effective architectural model of ANN.

However, current machine learning-based personal comfort models, like conventional thermal comfort research, concentrated on occupants' physiological and psychological differences without considering the spatial impact. Some research examined spatial differences in the machine learning based personal prediction model, such as incorporating desk-specific air velocity measurements to get the district ambient environment [11] and spatial impact in summer and winter settings [12]. The other differences produced by different spatial locations and spatial arrangement designs were not considered. Herein, significant spatial parameters should be addressed in the development of machine learning-based personal comfort models, particularly in hot and humid climates such as *ASEAN* nations, where the influence on prediction accuracy considering these aspects should be thoroughly examined. This study aimed to investigate the spatial effects on indoor personal thermal comfort by employing machine learning algorithms in an air-conditioned workplace under a hot and humid climate.

3 Methodology

The framework and processes that were implemented are depicted in Fig. 1; to create a thermal comfort model based using machine learning.

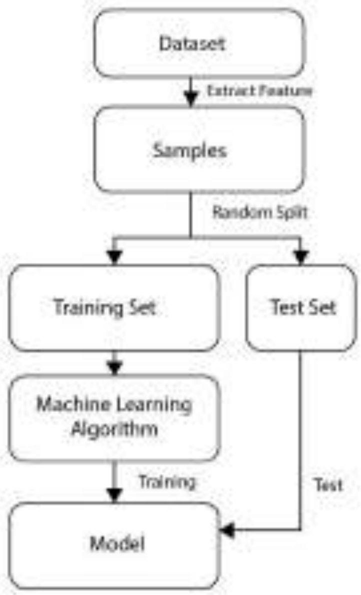

Fig. 1. The framework and processes of machine learning based, building thermal comfort model.

3.1 Dataset

The machine learning model is built using a dataset from [13]. This dataset has collected 628 samples from field thermal comfort studies in 5 buildings. It was gathered from public institutions, Universiti Teknologi Malaysia in Kuala Lumpur (UTMKL) and Universiti Teknologi MARA in Shah Alam (UiTM) which are built around densely populated urban zones in Malaysia. The dataset includes air temperature (T_a), globe temperature (T_g), relative humidity (*RH*), and air velocity (v_a), as well as subjective questionnaire answers. This experiment used *TSV* as a prediction output.

3.2 Data Input

The data that was selected for the input of this experiment can be grouped into 2 models.

3.2.1 Model 1

Model 1 consists of the 6 inputs of Fanger's model [7] consist of air temperature (T_a), relative humidity (*RH*), air velocity (v_a), mean radiant temperature (*MRT*), clothing insulation (*clo*), and metabolic rate (*Met*).

3.2.2 Model 2

Model 2 consists of the 6 inputs of Fanger's model with the distance of air-conditioned (*ACD*) and windows (*WD*) relative to the occupant in meters. Fig. 2 show the distance of ACD and WD that was calculated relative to the occupant.

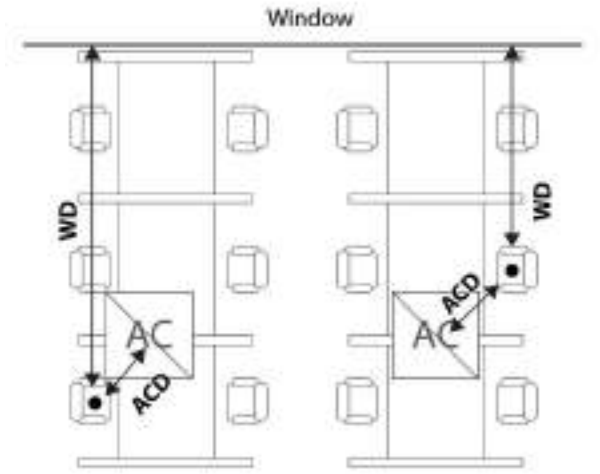

Fig. 2. The distance of ACD and WD is calculated relative to the occupant.

3.3 Machine Learning Algorithm

The machine learning algorithm used is decision tree (*DT*), random forest (*RF*), support vector machine (*SVM*), K-nearest neighbour (*KNN*), and neural network (*NN*). The setting that was used in all these algorithms is as follows. *DT* is set with the Gini criterion, best splitter with the 2 minimum samples split, and 1 minimum samples leaf. *SVM* is set with a linear kernel. *KNN* on the other hand, uses 3 neighbours with uniform weight and 30 leaf size. *NN* is set with 100 hidden layers, with rectified linear unit function for activation, and Adam optimizer. *NN* also uses 500 max iterations. Finally, *RF* is set with 500 trees, with entropy criterion.

3.4 Performance Metric

Accuracy, mean absolute error (*MAE*), and root mean square error (*RMSE*) were chosen as measures to evaluate the different ML models. A *ML* model's accuracy shows the algorithm's performance in recognising correlations and trends between variables in a dataset based on the input, or training data. *MAE* is the difference in size between an observation's predicted value and its actual value. Meanwhile, *RMSE* measures the model's absolute fit to the data, or how close the observed data points are to the model's anticipated values.

4 Results and Discussion

In total, 628 data are used to establish the personal thermal comfort prediction model in two models, which are model 1 and model 2. Table 1 shows the results of the *ML* prediction performance comparison between model 1, model 2, and *PMV*.

Based on Table 1, the spatial parameters give a significant contribution towards increasing the accuracy and reducing the *MAE* and *RMSE*. This is because spatial parameter takes to account the distance between air-conditioned and the distance between windows with indoor occupants. The average accuracy is increased by 4% when using the spatial parameter, with the *RF* giving the highest accuracy, an increase of 11.9%.

Table 1. The results of the *ML* prediction performance comparison between model 1, model 2, and *PMV*.

Model	Algorithm	Accuracy (%)	MAE	RMSE
Model 1	*DT*	46.8	0.762	1.094
	SVM	49.2	0.69	1.035
	NN	47.6	0.825	1.098
	RF	48.4	0.706	1.087
	KNN	41.2	0.865	1.165
Model 2	*DT*	47.6	0.716	1.027
	SVM	50	0.634	0.959
	NN	50.8	0.634	0.967
	RF	60.3	0.468	0.781
	KNN	44.5	0.817	1.085
PMV	-	-	0.946	1.208

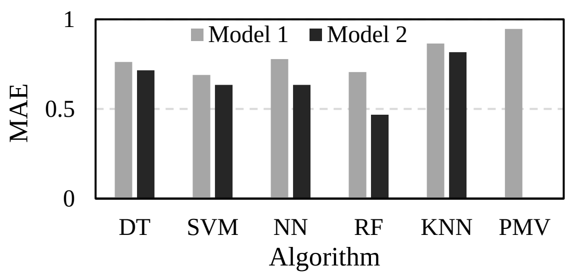

Fig. 3. The results of the *MAE* comparison between model 1, model 2, and *PMV*.

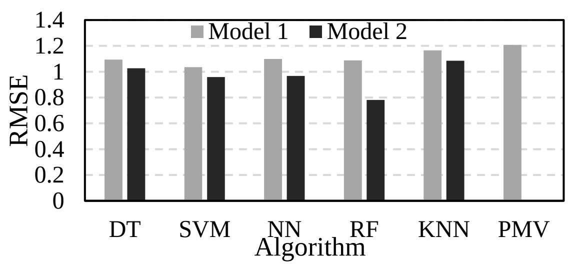

Fig. 4. The results of the *RMSE* comparison between model 1, model 2, and *PMV*.

Fig. 3 and Fig. 4 show the comparison of the *MAE* and *RMSE* of *PMV* with five *ML* algorithms that use model 1 and model 2 respectively. Based on this figure, when comparing model 1 and model 2, the average *MAE* is decreased by 10.64% with the *RF* giving the most *MAE* reduction by 23.8%. Furthermore, the average *RMSE* is reduced by 11.76% with the *RF* producing the most *RMSE* cutback by 30.6%.

When *RF* is using model 2 which is using spatial parameters, it is *MAE* performance compared with *PMV*, shows a significant reduction of 47.8%. The RMSE is also reduced by 42.7%.

Overall, model 2 produced good results when used with *ML* algorithms, reducing *MAE* and *RSME* error compared to *ML* methods that solely used 6 inputs of Fanger's model. The random forest using the model 2, performs best, obtaining 60% accuracy, *MAE* of 0.46, and *RMSE* of 0.78. Based on the result obtained, this observation is consistent with [12], where the researchers noted that the spatial features have a considerable impact on model prediction performance.

5 Conclusion and Future Works

This paper has assessed spatial effects on indoor personal thermal comfort by employing machine learning algorithms in the air-conditioned workplace. The spatial parameters revealed have a significant influence on model prediction accuracies, *MAE*, and *RMSE*. Although the accuracy increased slightly, the error performance metrics such as *MAE* and *RMSE* showed a considerable reduction when compared to *ML* performance using model 1. Compared with the *PMV*, it shows significant results of 47.8% for the *MAE* and 42.7% for the *RMSE*.

There are a few limitations of this research work mainly the dataset's limited size, the variations of people, variations of experimental locations and this paper only focuses on the spatial parameter of the distance of window and AC. Hence, for future work, the dataset from additional people in multiple experimental rooms in varied weather and season conditions needs to be acquired. Other than that, the spatial parameter can be expand further. This spatial effect analysis also needs to identify which section of the indoor has a cold or heat cluster area that affects people's thermal comfort. This study can help to contribute to the design of sustainable buildings.

This work was supported by Universiti Teknologi Malaysia Fundamental Research [Vot 22H26] and The Japanese Chamber of Trade & Industry, Malaysia (JACTIM) under Universiti Teknologi Malaysia [Vot 4B813].

References

1. IEA, World Energy Outlook 2022, IEA, Paris https://www.iea.org/reports/world-energy-outlook-2022, License: CC BY 4.0 (report); CC BY NC SA 4.0 (Annex A) (2022)
2. J. Guenther, S. Oliver, Feature Selection for Thermal Comfort Modeling based on Constrained LASSO Regression, IFAC-PapersOnLine, Volume 52, Issue 15, 2019, Pages 400-405 (2019)
3. L. Lan, P. Wargocki, & D. P. Wyon, Z, Lian. Effects of thermal discomfort in an office on perceived air quality, SBS symptoms, physiological responses, and human performance. Indoor air. 21. 376-90. (2011)
4. D. P. Wyon, I. B. Andersen, G.R. Lundqvist, The effects of moderate heat stress on mental performance. Scandinavian J. Work, Environment & Health, 5, 352-361. (1979)
5. C. Ana, N. Sandro, S. Petar, P. Toni, A. Aleksandar, C. Velimir, Investigation of personal thermal comfort in office building by implementation of smart bracelet: A case study, Energy, Volume 260, 2022, 124973. (2022)
6. ASHRAE, ANSI/ASHRAE Standard 55-2017 Thermal Environmental Conditions for Human Occupancy, In. Atlanta: ASHRAE. (2017)
7. P.O.Fänger, Fundamentals of Thermal Comfort, Advances In Solar Energy Technology. W. H. Bloss and F. Pfisterer, Oxford, 3056-3061. (1988)
8. X. Zhou, L. Xu, J. Zhang, B. Niu, M. Luo, G. Zhou, X. Zhang, Data-driven thermal comfort model via support vector machine algorithms: Insights from ASHRAE RP-884 database. Energy and Buildings, 211, 109795. (2020)
9. D. Li, C. C. Menassa, V. R. Kamat, Personalized human comfort in indoor building environments under diverse conditioning modes. Building and Environment, 126, 304-317. (2017)
10. F. Mousa, H. Hamdy, Real-Time Intelligent Thermal Comfort Prediction Model. International Journal of Advanced Computer Science and Applications, 12. (2021)
11. J. Guenther, S. Oliver, Feature Selection for Thermal Comfort Modeling based on Constrained LASSO Regression. IFAC-PapersOnLine, Volume 52, Issue 15, 2019, Pages 400-405. ISSN 2405-8963. (2019)
12. G. Puyue, C. Yuanzhi, Z. Zihan, Z. Cheng, C. Bing, S. Stephen, Investigating spatial impact on indoor personal thermal comfort. Journal of Building Engineering, Volume 45, 2022,103536. (2022)
13. S. M. Taib, S. A. Zaki, H. B. Rijal, A. A. Razak, A. Hagishima, W. Khalid, M. S. M. Ali, Associating thermal comfort and preference in Malaysian universities' air-conditioned office rooms under various set-point temperatures. Journal of Building Engineering 54, 104575. (2022)