



## Transfer learning for thermal comfort prediction in multiple cities

Nan Gao<sup>a</sup>, Wei Shao<sup>a</sup>, Mohammad Saiedur Rahaman<sup>a</sup>, Jun Zhai<sup>a</sup>, Klaus David<sup>b</sup>,  
Flora D. Salim<sup>a,\*</sup>

<sup>a</sup> School of Computing Technologies, RMIT University, Melbourne, Victoria, 3000, Australia

<sup>b</sup> University of Kassel, Kassel, 34131, Germany

### ARTICLE INFO

#### Keywords:

Human–building interaction  
Thermal comfort  
Transfer learning  
HVAC automation  
Smart building

### ABSTRACT

The HVAC (Heating, Ventilation and Air Conditioning) system is an important part of a building, which constitutes up to 40% of building energy usage. The main purpose of HVAC, maintaining appropriate thermal comfort, is crucial for the best energy usage. Additionally, thermal comfort is also important for well-being, health, and work productivity. Recently, data-driven thermal comfort models have achieved better performance than traditional knowledge-based methods (e.g. the predicted mean vote model). An accurate thermal comfort model requires a large amount of self-reported thermal comfort data from indoor occupants which undoubtedly remains a challenge for researchers. In this research, we aim to address this data-shortage problem and boost the performance of thermal comfort prediction. We utilize sensor data from multiple cities in the same climate zone to learn thermal comfort patterns. We present a transfer learning-based multilayer perceptron model from the same climate zone (TL-MLP-C\*) for accurate thermal comfort prediction. Extensive experimental results on the ASHRAE RP-884, Scales Project and Medium US Office datasets show that the performance of the proposed TL-MLP-C\* exceeds the performance of state-of-the-art methods in accuracy and F1-score.

### 1. Introduction

Recently, Internet of Things (IoT) devices have been widely used in urban environments. In addition, sensors have become the backbones of smart cities that enable spatial and situational awareness of real-time dynamic phenomena, e.g., pedestrian movement [1], parking events [2], and energy consumption [3,4]. As one of the most important parts of cities, buildings account for approximately 40% of the global energy usage and 60% of the worldwide electricity usage [5]. Large proportions of these usages are contributed by buildings' HVAC systems [6]. The main goal of the HVAC system is to maintain the indoor occupant comfort at minimal energy usage. To achieve overall satisfaction with an indoor environment, thermal comfort is considered to be the most influential factor compared with visual and acoustic comfort [7].

*Thermal comfort* is the state of mind that expresses satisfaction with the thermal environment [8]. Thermal discomfort not only affects occupant productivity, work performance and engagement [9,10], but it also has a negative influence on lifelong health. Hence, it is important to maintain a thermally comfortable environment for the well-being of occupants while minimizing buildings' energy usage. A crucial step towards this goal is to create an accurate model for thermal comfort.

The *Predicted Mean Vote* (PMV) model proposed by Fanger et al. [11] developed with principles of human heat balance and adopted by the ASHRAE Standard 55, is one of the most prevalent models. It relates the thermal comfort scale with six different factors (see Fig. 1).

However, some researchers revealed the discrepancy between the thermal sensation vote reported by occupants and the predicted mean vote [12]. This discrepancy is likely because a variety of parameters such as time factors (e.g., hour, day, and season) [13,14], personal information (e.g., heart rate, age, and gender) [15], environmental factors (e.g., colour, light, and outdoor climates) [16], culture (e.g., dress code and economic status) [17], short- and long-term thermal exposure [14], etc. may affect thermal comfort. Therefore, a data-driven method is a better choice than the traditional PMV model since more parameters could be utilized to improve the performance of thermal comfort prediction.

Some researchers have applied data-driven machine learning techniques for thermal comfort prediction for a specified group of people. However, it is usually difficult to obtain sufficient labelled data, which limits the performance of data-driven models. Recently, various thermal comfort studies have been conducted worldwide; and several databases, including databases covering multiple cities and climate

\* Corresponding author.

E-mail addresses: [nan.gao@rmit.edu.au](mailto:nan.gao@rmit.edu.au) (N. Gao), [wei.shao@rmit.edu.au](mailto:wei.shao@rmit.edu.au) (W. Shao), [saiedur.rahaman@rmit.edu.au](mailto:saiedur.rahaman@rmit.edu.au) (M.S. Rahaman), [jun.zhai@rmit.edu.au](mailto:jun.zhai@rmit.edu.au) (J. Zhai), [david@uni-kassel.de](mailto:david@uni-kassel.de) (K. David), [flora.salim@rmit.edu.au](mailto:flora.salim@rmit.edu.au) (F.D. Salim).

<https://doi.org/10.1016/j.buildenv.2021.107725>

Received 17 November 2020; Received in revised form 14 January 2021; Accepted 15 February 2021

Available online 23 February 2021

0360-1323/© 2021 Published by Elsevier Ltd.

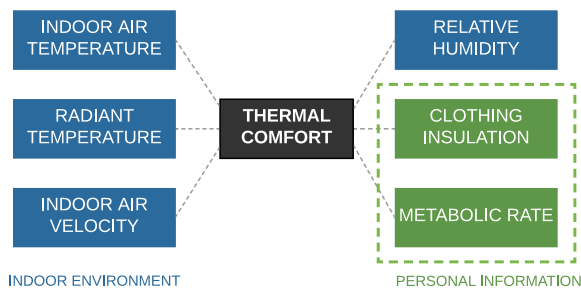


Fig. 1. Six factors affecting thermal comfort (PMV model).

zones, are currently available online (see Section 3). Since the sensor data inferred from different cities may have very divergent patterns caused by building materials, construction requirements and climate changes, previous research has mainly focused on investigating how people living in specific cities react to their thermal environment (e.g., the hot-arid climate in Kalgoorlie-Boulder, Australia [18] and the humid subtropical climate in Brisbane Australia [19]).

We aim to explore whether we can utilize sensor data from multiple cities to benefit a target building. We hypothesize that the performance of thermal comfort modelling can be boosted by conducting transfer learning using data from multiple cities. Therefore, we wish to answer the following research questions: *Can we predict occupants' thermal comfort accurately by learning from multiple buildings in the same climate zone when we do not have enough data? If so, which features contribute the most to effective thermal comfort transfer learning?*

In this research, we present the *transfer learning-based multilayer perceptron* (TL-MLP) model and *transfer learning-based multilayer perceptron from the same climate zone* (TL-MLP-C\*) model for predicting occupants' thermal sensation with insufficient labelled data. ASHRAE RP-884 [11] and the Scales Project [20] are chosen as the source datasets, and the Medium US Office [21] is used as the target dataset. Extensive experiments on these three public databases show that the proposed thermal comfort models outperform the popular knowledge-driven and data-driven models. To summarize, the contributions are as follows:

- To the best of our knowledge, we are the first to transfer the knowledge from similar thermal environments (climate zones) to a target building for effective thermal comfort modelling. We propose the TL-MLP and TL-MLP-C\* thermal comfort models and confirm that the thermal comfort sensor data from multiple cities in the same climate zone can benefit the small thermal comfort dataset of a target building in another city, with insufficient training data.
- Extensive experimental results show that the proposed TL-MLP and TL-MLP-C\* models outperform the popular knowledge-driven and data-driven models for thermal comfort prediction and can be implemented in buildings without adequate thermal comfort labelled data.
- We identify the significant feature sets for effective thermal comfort transfer learning. We also find that the combination of age, gender, outdoor environmental features and the six factors from the PMV model can lead to the best prediction performance for transfer learning-based thermal comfort modelling.

## 2. Related work

First, we list the previous literature for traditional thermal comfort modelling methods and transfer learning applications. Then, we discuss the current gaps and identify the advantages of this work.

### 2.1. Traditional thermal comfort modelling methods

The PMV model developed by Fanger et al. [11] and adaptive model developed by De Dear et al. [19] are the most famous knowledge-driven thermal comfort models. The adaptive model is based on the idea that occupant can adapt to different temperatures at different times and that outdoor weather affects indoor comfort. Occupants can achieve their comfort through personal adjustments such as clothing changes or window adjustments [22]. Clear et al. [23] explored how adaptive thermal comfort could be supported by new ubiquitous computing technologies. They noted that IoT sensing technologies can help build a more sustainable environment where people are more active in maintaining and pursuing their thermal comfort, which is less energy-intensive and less tightly controlled.

In recent years, data-driven thermal comfort modelling has become increasingly more popular and huge efforts have been made to apply machine learning to thermal comfort modelling [15,24–29]. Ran et al. [24] used rotation forests to predict occupants' thermal comfort using thermographic imaging information. Similarly, Ghahramani et al. [25] used a hidden Markov model (HMM) based method to predict thermal comfort using the infrared thermography of faces. Chaudhuri et al. [15] established a random forest-based model for different genders using physiological signals (e.g., skin conductance and blood pressure). However, all the thermal comfort models mentioned above require the installation of additional devices (individual thermal cameras, smart eyeglasses, and physiological sensors) and may lead to privacy concerns.

The performance of traditional machine learning algorithms on thermal comfort prediction has been discussed in [26]. Researchers compared nine widely used machine learning algorithms for thermal sensation prediction using the ASHRAE Comfort Database II. They found that ML-based thermal sensation prediction models generally have higher accuracy than traditional PMV models and that the random forest has the best performance compared to other ML algorithms.

As the non-traditional machine learning algorithms, artificial neural networks have been increasingly used in thermal comfort modelling. Ferreira et al. [27] controlled an HVAC system to achieve the desired thermal comfort level and energy savings. They applied several neural network models to calculate the PMV index for model-based thermal comfort prediction. Hu et al. [28] implemented a black-box MLP neural network for thermal comfort modelling, which obtained better prediction performance than the PMV model and traditional white-box machine learning models. Compared to most previous research using a coarse-grained neural network architecture (link input attributes and thermal comfort score directly), Zhang et al. [29] used the MLP neural network to model the relationship between controlling building operations and thermal comfort factors. Their proposed fine-grained DNN approach for thermal comfort modelling outperforms the coarse-grained modelling and other popular machine learning algorithms.

### 2.2. Transfer learning applications

Although great contributions have been made to improve the prediction accuracy of thermal comfort through various machine learning techniques, there is still a main bottleneck for data-driven thermal comfort modelling — the accessibility of sufficient thermal comfort data. Transfer learning allows researchers to learn an accurate model using only a tiny amount of new data and a large amount of data from a previous task [30].

Transfer learning has been applied to many real-world applications involving image/video classification, natural language processing (NLP), recommendation systems, etc. For instance, transfer learning has been used for children's Automatic Speech Recognition (ASR) task [31]. Researchers learn from adult models to child models through

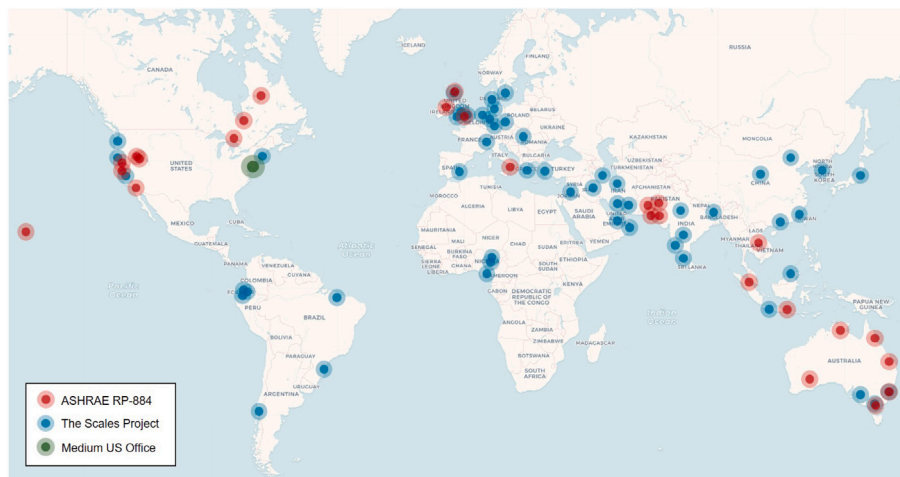


Fig. 2. Locations of different studies in ASHRAE RP-884 database, The scales project database and Medium US Office dataset. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

a Deep Neural Network (DNN) framework. They investigated the transfer learning techniques between adult and child ASR systems in acoustic variability (layers near the input) and pronunciation variability (layers near the output), updated both the top-most and bottom-most layers and kept the rest of the layers fixed.

Some existing work has focused on transfer learning using sensor data. Wang et al. [32] proposed a transfer learning based-framework for cross-domain activity recognition. First, they used the majority voting technique to obtain the pseudo label of the target domain. Intra-class knowledge transfer was interactively performed to convert two domains into the same feature subspace. Then, the labels of the target domain can be ignored by the second annotation. Ye et al. [33] learned human activity labels by leveraging annotations across multiple datasets with the same feature space, even though the datasets may have different sensing deployments, sensing technologies and different users.

Recently, a transfer active learning framework was proposed to predict thermal comfort [34]. They considered thermal comfort prediction as inductive transfer learning where labelled data are available in both source and target domains but users do not have access to all labelled data in the target domain. They used the parameters transferred from the source domain to the target domain. The biggest disadvantage of their method is that they assume the feature spaces in both domains must be the same, which is not applicable in daily life as there may be unique useful features in the target dataset.

Similarly, Hu et al. [35] adopted transfer learning for thermal comfort modelling and assumed that the feature space of the source domain is a subset of that of the target domain. They connected the classifiers from the source domain and target domain and then built a new classifier to obtain knowledge from the source domain; however, they did not explain why the network structure works well. Besides, they trained the thermal comfort model for a lab study and learned knowledge from the data from buildings all over the world in the ASHRAE RP-884 dataset, but they did not consider the differences in the thermal environments in different climate zones.

Overall, there are several advantages of our work: (1) We are the first to transfer the knowledge from similar thermal environments (climate zones) to the target building for effective thermal comfort modelling. Most previous research has focused on building a thermal comfort model for one target building [15,24–29]. Although a few researches [35] have started to use transfer learning for building thermal comfort models, their target datasets are collected from laboratory studies and do not consider the influences of different climate zones. (2) Unlike some research that uses data collected from laboratory studies [15,24,25,35], we build thermal comfort models using data

from field studies in both the target and source domains, which is much more meaningful in real-world scenarios. (3) Compared with some research utilizing additional devices (e.g., thermal cameras in [24], eyeglasses in [25], and wristbands in [35]), our research is easier and cheaper to conduct, and better protects the privacy of occupants.

### 3. Data sets introduction

#### 3.1. Overview

ASHRAE RP-884 Database [8] is one of the most popular public databases for human thermal comfort study, which has been used in numerous previous research [36–39]. ASHRAE RP-884 dataset was initially collected to develop De Dear's adaptive model, involving more than 25,000 observations collected from 52 studies and 26 cities over different climate zones all over the world. We adopt this public database as one of the source datasets in our research.

The Scales Project Dataset [20] is published in 2019 which contains thermal comfort responses from 57 cities in 30 countries for 8225 participants. This dataset aims at exploring participants' thermal comfort, thermal sensation, thermal acceptances and to investigate the validity of assumptions regarding the interpretation of responses from the survey. This public dataset is used as one of the source datasets in the research.

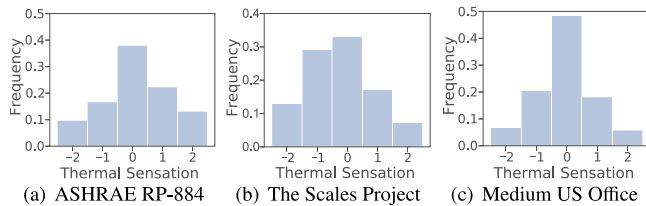
Medium US Office Dataset [21] developed by Langevin et al. [21] is a popular dataset used by recent thermal comfort studies [29,40]. It collected data from 24 participants (16 females and 8 males) in the Friends Center Office building in Philadelphia city, USA. Longitudinal thermal comfort surveys are distributed online 3 times daily (morning, mid-day and afternoon) for a continuous 2-week period in each of the four project seasons between July 2012 and August 2013. Data types vary from daily surveys to sensor data including but not limited to the indoor air temperature, air velocity, relative humidity, CO<sub>2</sub> concentration and illuminance. This public dataset is used as the target dataset in the research.

The locations of all cities with data used in the study are displayed in Fig. 2. The red points represent the 26 cities in the ASHRAE RP-884 database, the blue points indicate the 57 cities in the Scales Project dataset, and the green point indicates Philadelphia in the Medium US Office dataset. In this research, we aim to learn the knowledge from data in cities indicated by red points and blue points to benefit one building in Philadelphia (green point).

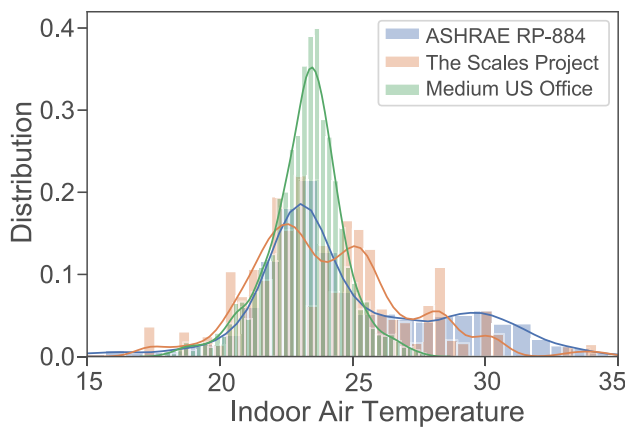
Table 1 shows the basic information for the three datasets. The first two datasets have different building types (HVAC, naturally ventilated and mixed ventilated) while there is only one HVAC building in the

**Table 1**  
Information for source dataset and target dataset.

Dataset	ASHRAE RP-884	The scales project	Medium US office
Instances	25,623	8225	2497
Participants	Unknown (48% M, 52% F)	8225 (53% M, 46% F)	24 (33% M, 67% F)
Indoor AT Range (C)	6.2 – 42.7	13.2 – 34.2	17.9 – 27.8
Indoor RH Range (%)	2.0 – 97.8	18.0 – 82.4	15.7 – 72.4
Indoor AV Range (m/s)	0.01 – 1.71	0.00 – 0.70	0.02 – 0.19
MR Range (Met)	0.64 – 6.82	N/A	1.00 – 6.80
CL Range (Clo)	0.04 – 2.29	N/A	0.21 – 1.73



**Fig. 3.** Distribution of thermal sensation over different datasets.



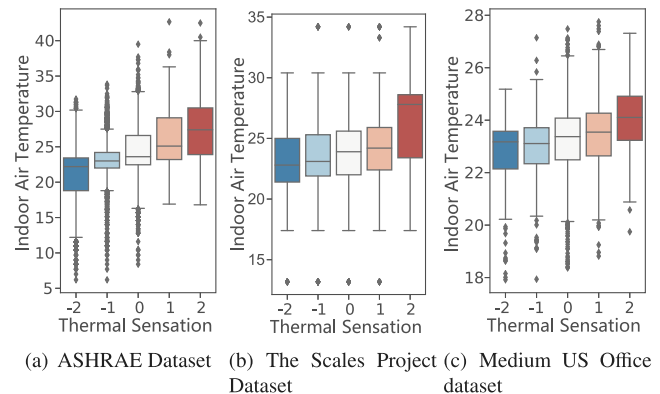
**Fig. 4.** Distribution of the indoor air temperature over different domains.

Medium US Office dataset. Since the ASHRAE and Scales datasets include different climate zones all over the world, they have wider indoor air temperature ranges than the Friends Center building in the Medium US Office (17.9 °C–27.8 °C). Different from the first two datasets, the Medium US Office dataset has much smaller groups of participants. Besides, the ranges of the indoor relative humidity (Indoor RH), indoor air velocity (Indoor AV), metabolic rate (MR), and clothing level (Clo) in the Medium US Office dataset are smaller than those in the ASHRAE dataset.

### 3.2. Preliminary analytics

Fig. 3 shows the distribution of thermal sensation for the ASHRAE RP-884 dataset, the Scales Project and the Medium US Office dataset. Since the numbers of instances of the sensation scale for +3 (Hot) and –3 (cold) are far less than those of the other instances in both datasets, we merged +3 (hot) and +2 (warm) into one class, and –3 (cold) and –2 (cool) into one class. In the office environment, indoor environmental factors such as temperature are generally maintained at a relatively comfortable level (17.9 °C–27.8 °C in the Medium US dataset), and people can also choose to adjust their clothing level and behaviour (e.g., open the heater vents and have hot drinks) if they are too cold or too hot.

Although the regression model is effective in many time-series problems [2,41], the classification method still dominates the thermal comfort area. Therefore, in this paper, we choose classifiers rather than



**Fig. 5.** Boxplots of thermal sensation and the indoor temperature.

regressors for effective thermal comfort prediction. Besides, based on the previous discussion, thermal sensation scales are classified into 5 categories (i.e., cold or cool, slightly cool, neutral, slightly warm, hot or warm).

The above three datasets have similar thermal sensation distributions, and occupants feel neutral towards the thermal environment most of the time. We can observe that there are more responses for feeling slightly warm or cool than feeling warm/cool or hot/cold, which accords with our thermal comfort feelings in daily life. Meanwhile, the thermal sensation distributions in the ASHRAE dataset and the Scales Project dataset are more uniform than the distribution of the Medium US Office dataset. This is because the ASHRAE dataset and the Scales Project dataset consist of a variety of data from different climate zones all over the world while the Medium US Office dataset includes data from only one building.

Indoor air temperature is one of the most significant factors affecting occupants’ thermal feelings. Fig. 4 shows the distribution of the indoor air temperature for the three datasets. Most temperature values range from 22 °C–24 °C. However, there are also some differences between these three distributions. The ASHRAE and the Scales Project datasets have higher indoor air temperatures because some thermal sensation responses are from hot climate areas. In contrast, in the Medium US Office dataset, the indoor temperature distribution seems to be centred at approximately 20 °C to 27 °C.

From Fig. 5, we can see the relationship between the indoor air temperature and thermal sensation scale. Usually, a higher indoor air temperature indicates a higher thermal sensation scale for all three datasets. Interestingly, in the Medium US Office dataset, the average indoor air temperature for feeling cold or cool is slightly higher than that for feeling slightly cool. This phenomenon may be due to there being too few subjects (24 participants in total) in the Medium US Office dataset. Additionally, the other factors such as the relative humidity, age, gender, and outdoor weather will affect the thermal sensation. This is the reason why we use as many features as possible to build a more accurate and robust thermal comfort prediction model.

From the above analysis, there are observable differences between the ASHRAE, Scales Project and Medium US Office datasets. One of the

reasons is that buildings in these three datasets are located in various climate zones, where climate variability can lead to a different working environment, occupant cognition and behaviour, therefore affecting occupants' thermal sensation in different buildings. Considering that the three datasets share many similarities in occupant thermal comfort and that the number of instances in the target dataset is very limited, we explore occupants' thermal comfort by learning from multiple buildings in the same climate zone with similar climate conditions. We will then introduce the proposed thermal comfort modelling in Section 4.4.

## 4. Methodology

### 4.1. Problem definition

To learn sensor data from multiple datasets for thermal comfort modelling, some notations need to be defined in this paper. Firstly, we give the definition of a 'task' and a 'domain'. A domain  $D$  can be represented as  $D = \{\mathcal{X}, P(X)\}$ , which contains two parts: the feature space  $\mathcal{X}$  and the marginal probability distribution  $P(X)$ , where  $X = \{x_1, x_2, \dots, x_n\} \in \mathcal{X}$ . The task  $\mathcal{T}$  can be represented as  $\mathcal{T} = \{y, f(\cdot)\}$ , which contains two components: the label space  $y$  and a target prediction function  $f(\cdot)$ .  $f(\cdot)$  cannot be observed but can be learnt from the training data, which could also be considered as a conditional function  $P(y|x)$ .

In the context of traditional machine learning, the common assumption is that the training and test data share exactly the same feature space and data distribution [42]. However, once the new task  $\mathcal{T}$  arrives and its data distribution  $P(X)$  is different from the previous task, the new model must be rebuilt from the beginning using the current data. Such method requires extra effort and is very expensive in most cases. Compared with traditional machine learning methods, transfer learning can tolerate differences in data distribution and utilize knowledge from other sources to target tasks.

In this research, we transfer the knowledge from the source domain (RP-884 and the Scales Project datasets) to benefit thermal comfort prediction in the target domain (Medium US Office dataset). Although both domains have different features, they share several common features such as the indoor air temperature, indoor relative humidity, indoor air velocity, indoor mean radiant temperature, clothing level, metabolic rate, and occupants' age and gender. Therefore, predicting thermal comfort falls under *transductive transfer learning* [43], which can be formally defined as follows: given a source domain  $D_s$  and the corresponding learning task  $\mathcal{T}_s$ , a target domain  $D_t$  and the corresponding learning task  $\mathcal{T}_t$ , we aim to improve the performance of the prediction function  $f(\cdot)$ , in  $\mathcal{T}_t$  by discovering the knowledge from  $D_s$  and  $\mathcal{T}_s$ , where  $D_s \neq D_t$  and  $\mathcal{T}_s = \mathcal{T}_t$ .

Fig. 6 shows the thermal comfort transfer learning system in which we could use the transfer learning method to learn knowledge from the source datasets and benefit the target dataset in a specified city.

### 4.2. Feature selection

Human thermal sensation is influenced by a variety of factors such as time factors [13], personal information [15], environmental changes [16], and culture [17]. In this research, several features are chosen for thermal comfort transfer learning based on the following criteria: (1) the features were commonly studied in previous thermal comfort research and (2) the features are easy calculate or collect by using passive sensing or self-reported responses. In summation, we divide the features into three broad categories: indoor environmental features, outdoor environmental features and personal features. Table 2 displays the selected features in the Medium US Office dataset.

**Indoor Environmental Features.** The indoor environment affects occupants' thermal comfort directly, and we adopt the following basic indoor environmental features derived from Fanger's PMV model [11] for thermal comfort prediction: the air temperature, mean radiant

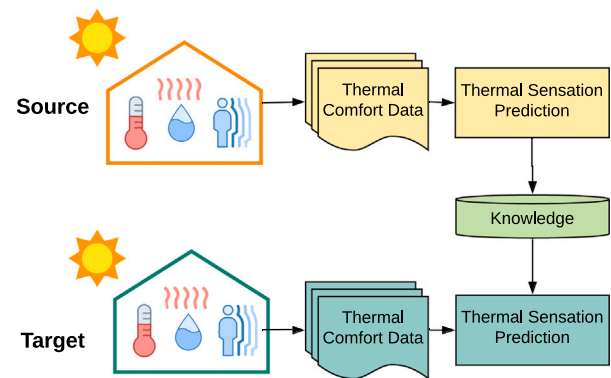


Fig. 6. Thermal comfort transfer learning system.

temperature, air velocity and relative humidity. The air temperature is the average temperature of the air surrounding the occupant at a location and time. The radiant temperature indicates the radiant heat transferred from a surface, and the mean radiant temperature is affected by the emissivity and temperature of the surrounding surfaces, viewing angles, etc. The air velocity is the average speed of air with respect to the direction and time. The relative humidity is the ratio of the amount of water vapour in the air to the amount of water vapour that the air can hold at a specified pressure and temperature.

**Outdoor Environmental Features.** Outdoor weather conditions can have physiological effects on individuals thermal perception and clothing preference in different seasons [12,44]. For instance, in summer people tend to choose lightweight clothing, which will influence their indoor thermal comfort. The most popular measurements of the outdoor environment include the outdoor air temperature and outdoor humidity, which will also be adopted in this research.

**Personal Features.** Studying personal features is crucial for effective thermal comfort modelling because thermal sensation is a subjective measurement and different individuals perceive the same environment differently. In this research, we selected the following personal features: clothing insulation, metabolic rate, age and gender. Clothing insulation has a major impact on the thermal comfort level because it affects heat loss and thus the heat balance. Previous research shows the relationship between age and thermal sensation [17,45]. Besides, Sami et al. [46] found a significant gender difference in thermal comfort: females tend to prefer a higher room temperature than males and feel both uncomfortably hot and uncomfortably cold more often than males. Hence, gender and age are considered to be the features for thermal comfort modelling.

The features in a source domain can be considered as a subset in the target domain. The ASHRAE dataset shares eight features with the Medium US Office dataset while the Scales Project dataset only shares six features with the target dataset. Although there are various other features in these three datasets such as occupant behaviour data (e.g., adjusting heaters/curtains/ thermostats) and background survey data (e.g., acceptable temperature), we simplify the thermal comfort prediction and therefore do not show the other features.

### 4.3. Imbalance class distribution

As the thermal sensation scale has 5-point values, we regard thermal comfort prediction as a classification task. Fig. 3 shows the distributions of the ASHRAE RP-884, Scales Project and Medium US Office datasets. It is clear that the three distributions are imbalanced, and the number of thermal sensation instances for  $-1$  (cool) to  $1$  (warm) far exceeds the number of other instances. To train a fair classifier, we must address this class imbalance issue in thermal comfort data. Take the binary classification as an example. If class  $M$  is 95% and class  $N$  is 5% in the

**Table 2**  
Selected features in the Medium US Office dataset.

Category	Data source	Feature name	Description	Units
Indoor	HOBO Datalogger (15 min)	Indoor_AT	Indoor temperature	°C
		Indoor_RH	Indoor relative humidity	%
		Indoor_AV	Indoor air velocity	m/s
		Indoor_AMRT	Indoor radiant temperature	°C
Outdoor	Weather Analytics (15 min)	Outdoor_AT	Outdoor temperature	°C
		Outdoor_RH	Outdoor humidity	%
Personal	Daily survey (3 times/day)	CL	Clothing insulation	clo
		MR	Metabolic rate	Met
	Background survey (once)	Age	Participants' age	Years
		Gender	Participants' gender	

dataset, we can simply reach an accuracy of 95% by predicting class  $M$  each time, which provides a useless classifier for our purpose. In this research, we assume that the survey responses are 'correct'. Although there may be some biases (e.g., rating bias, anchoring bias, and social desirability bias) in self-reported data, we will not discuss them in this paper.

To address an imbalanced dataset, oversampling and undersampling are efficient techniques to adjust the class distribution of the dataset. Under-sampling (e.g., *clustering*, *edited nearest neighbours* [47] and *Tomek links* [48]) can balance the dataset by reducing the size of the majority class. However, undersampling methods are usually used when we have sufficient data. Oversampling (e.g., the *synthetic minority oversampling technique* [49] and *adaptive synthetic sampling* [50]) aims to balance the dataset by increasing the number of minority classes, which can be applied when the data are insufficient.

*Generative Adversarial Networks* (GANs) have been successfully applied in various fields to learn the probability distribution of a dataset and synthesize samples from the distribution [51,52]. A GAN uses a generator  $G$  to capture the underlying data distribution of a dataset and a discriminator  $D$  to estimate the probability that a given sample comes from the original dataset rather than being created by  $G$ . Some techniques such as the *TableGAN* [53] and *TabularGAN* [54] have been proposed to handle the imbalance of tabular data. In particular, Quintana et al. [55] used the *TabularGAN* to synthesize a small thermal comfort dataset. They found that when the amount of synthesized data is no larger than the amount of real data, the thermal comfort dataset can achieve similar performance to the real samples.

In the thermal comfort classification problem, labelled thermal comfort responses are usually few. Therefore, in this research, we synthesize survey responses to handle the imbalance of thermal sensation classes. The *TabularGAN*<sup>1</sup> is used in this research to generate tabular data based on the generative adversarial network. It can learn each column's marginal distribution by minimizing the KL divergence, which is more suitable for thermal comfort classification problems compared with other methods such as the *TableGAN*, *edited nearest neighbours* [47], *SMOTE* [49], etc. The reason why we did not adopt the *TableGAN* is that it optimizes the prediction accuracy on synthetic data by minimizing the cross entropy loss while *TabularGAN* focuses more on the marginal distribution. The *TabularGAN* learns each column's marginal distribution by minimizing the KL divergence, which is more suitable for the thermal comfort classification problem.

#### 4.4. Thermal comfort modelling

Traditional algorithms for thermal comfort modelling is isolated and occurs purely based on specific buildings in the same climate zone. No thermal comfort knowledge is retained that can be transferred from one thermal comfort model to another. Recently, the transfer learning technique has been intensively studied in different applications [31,

34]. It aims to leverage knowledge from source tasks and then apply them to the target task. There are various transfer learning techniques that can be roughly grouped into three categories: *inductive transfer learning*, *unsupervised transfer learning* and *transductive transfer learning* [56]. Inductive transfer learning [57] aims to improve performance on the current task after having learned a different but related skill or concept on a previous task. Unsupervised transfer learning [58] focuses on solving unsupervised learning tasks in the target domain such as dimensionality reduction, clustering, and density. Transductive transfer learning aims to utilize the knowledge from the source domain to improve the performance of the prediction task in the target domain.

Transductive transfer learning can exploit the different levels of information captured from different layers in the neural network. Generally, layers close to the input data capture specific characteristics in the dataset while deeper layers capture information more relevant to the tasks (e.g., object types in image recognition and thermal sensation labels in thermal comfort prediction). The Medium US Office dataset, as described in Section 3.1, differs in cities and climate zones from the ASHRAE dataset and the Scales Project dataset. In different climate zones, there are various factors possibly contributing to thermal comfort, e.g., climate characteristics and occupants' recognition and endurance. This motivates us to investigate transfer learning between the ASHRAE/Scales Project datasets and Medium US office dataset in climate variability, which is close to the layers near the input.

We assume that climate variability affects the lower-level neural network only. Therefore, these layers need to be adapted to better represent the Friends Center office building in the target dataset. This can be regarded as retaining the knowledge of higher-level mappings from the source dataset. Hence, we retain the last hidden layer of the models on the ASHRAE and Scales Project datasets as shown in Fig. 7. Then, the thermal comfort neural network will be retrained with the Medium US Office dataset until convergence to find the optimal parameters for the lower hidden layers.

## 5. Experiment

In this section, we conduct experiments on the proposed thermal comfort transfer learning models and compare the performance with the state-of-the-art techniques and different configurations. We address the two research questions: *Can we predict occupants' thermal comfort accurately by learning from multiple buildings in the same climate zone when we do not have enough data? If so, which features contribute the most to effective thermal comfort transfer learning?* Specifically, we explore how the numbers of hidden layers and sample size of the training set in the target building affect thermal comfort transfer learning performance.

### 5.1. Experimental Setup

In our research, the source domain (ASHRAE RP-884 dataset and the Scales Project dataset) and the target domain (Medium US Office dataset) share some common features, which include four indoor environmental variables (air temperature, indoor relative humidity,

<sup>1</sup> Python package for TabularGAN: <https://pypi.org/project/tgan/>.

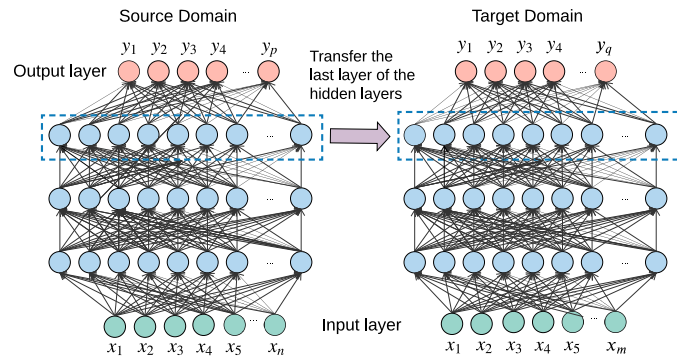


Fig. 7. The architecture for thermal comfort transfer learning.

mean radiant temperature, and indoor air velocity), two environmental variables (air temperature and humidity) and two personal variables (age and gender). In addition, the ASHRAE RP-884 and Medium US Office datasets share two other personal variables (clothing insulation and metabolic rate). The shared features make it possible to transfer knowledge to the target domain from the source domain.

**Preprocessing.** As discussed in Section 3.2, we first merge the minority classes and reclassify the thermal sensation into five categories. Then, we standardize the features by scaling them to unity variance for better classification performance. Considering that the thermal sensation classes are extremely imbalanced, in order to train a meaningful classifier, the *TabularGAN* [54] technique is applied for synthesizing the samples in all the classes except the majority class in the training set. Here, 50% of the samples in each class were synthesized while ensuring that the number of samples per category did not exceed the number of samples in the majority class.

Taking the Medium US Office dataset as an example, there are 2497 instances in the original dataset. After removing the null values and categorizing the thermal sensation responses, there were 1090 ‘neutral’ responses, 462 ‘slightly cool’ responses, 408 ‘slightly warm’ responses, 154 ‘cool or cold’ responses and 131 ‘warm or hot’ responses. After synthesizing the data using the *TabularGAN*, there were 981 ‘neutral’ responses, 624 ‘slightly cool’ responses, 551 ‘slightly warm’ responses, 208 ‘cool or cold’ responses and 177 ‘warm or hot’ responses in the training set (90% of the dataset).

**Architecture.** In this research, we choose the multilayer perceptron (MLP) neural network as the classifier for the source domain and target domain. Each neural network consists of two hidden layers with 64 neurons in each layer. The Relu function is used as the activation function in hidden layers. Then, the softmax function is applied to the output layer as the activation function. We train the classifier with the categorical cross-entropy loss function and the Adam optimizer with learning rate = 0.001. The batch size is set to 200 and the max epoch has been set to 500. Besides, the fixed random seed is chosen for dataset shuffling and training.

**Evaluation.** Similar to previous thermal comfort studies [15,24,25, 28], the accuracy and weighted F1-score are chosen as the performance metrics. Accuracy reflects the overall performance of the thermal comfort model. Since our priority goal is to correctly predict the thermal sensation for as many occupants as possible to achieve overall thermal comfort/energy savings in the building, accuracy is used as the main evaluation metric in this problem. We also adopt the weighted F1-score as the best metric to assess the accuracy of capturing performance across imbalanced classes. The F1-score considers both false positives and false negatives to strike a balance between the precision and recall. The ‘weighted-average’ calculates the metrics for each class and finds their average weighted by the number of true instances for each class. Compared with the ‘macro-average’ method, the ‘weighted average’ considers class imbalances. The weighted F1-score is helpful for evaluating thermal sensation classifiers as it considers all imbalanced

classes. That is, it evaluates the classifiers for different user groups with different thermal sensation levels instead of all occupants globally.

**Baselines.** For the baseline, three different categories of baselines are selected for comparison with our proposed method: random guessing, the PMV model and multiple traditional machine learning models. Random guessing generates the sample from the distribution of thermal comfort and regards it as a predicted value. Similar random baselines have been widely used in previous thermal comfort studies such as [35,59]. The PMV model is the most prevalent thermal comfort model worldwide. In the experiment, we will only use the four indoor environmental variables, the metabolic rate and clothing insulation to calculate the PMV score  $p_s$  according to the formula in [60] for the target dataset. Then, the thermal sensation class  $C(p_s)$  is calculated using Eq. (1).

$$C(p_s) = \begin{cases} -2, & \text{if } p_s \leq -1.5 \\ -1, & \text{if } -1.5 < p_s \leq -0.5 \\ 0, & \text{if } -0.5 < p_s \leq 0.5 \\ 1, & \text{if } 0.5 < p_s \leq 1.5 \\ 2, & \text{if } p_s \geq 1.5 \end{cases} \quad (1)$$

For the multiple traditional machine learning models, we choose K-nearest Neighbours [61], Naive Bayes [62], Support Vector Machine (with Linear, RBF and Polynomial kernel) [63], Decision Tree [64], Random Forest [65], AdaBoost [66] as baselines. Naive Bayes [62] from Bayes family methods is chosen due to its fast speed and working well with high dimensions. Support Vector Machine [63] technique is efficient for handling high dimensional spaces. Different from algorithms like SVM, AdaBoost [66] is fast, simple and easy to use with less need for tuning parameters. K-nearest Neighbours [61] is a simple method storing all available instances and classifying data instances according to a similarity measure, which has been widely used in the pattern recognition and statistical prediction area. Random Forest [65] is an ensemble learning method for classification operated by building multiple decision trees. It can cope with high-dimensional features and judge the feature importance.

Compared to the PMV model using six factors for thermal comfort prediction, the multiple machine learning algorithms use ten features as input features (see Table 2). Besides, all three of the above baselines build a thermal comfort classification model using the Medium US dataset.

**Cross-validation.** We apply the *k-fold cross-validation* [69] ( $k = 10$ ) method for effective thermal comfort classification. The advantage of 10-fold cross-validation is that it estimates the unbiased generalization performance of the thermal comfort prediction model. In the experiment, the data from the target domain (US Medium Office dataset) are randomly partitioned into 10 folds, each fold serves as the testing data iteratively, and the remaining 9 folds are used as the training data. The cross-validation process is repeated 10 times, and the prediction results (accuracy and weighted F1-score) are averaged to produce a single estimation.

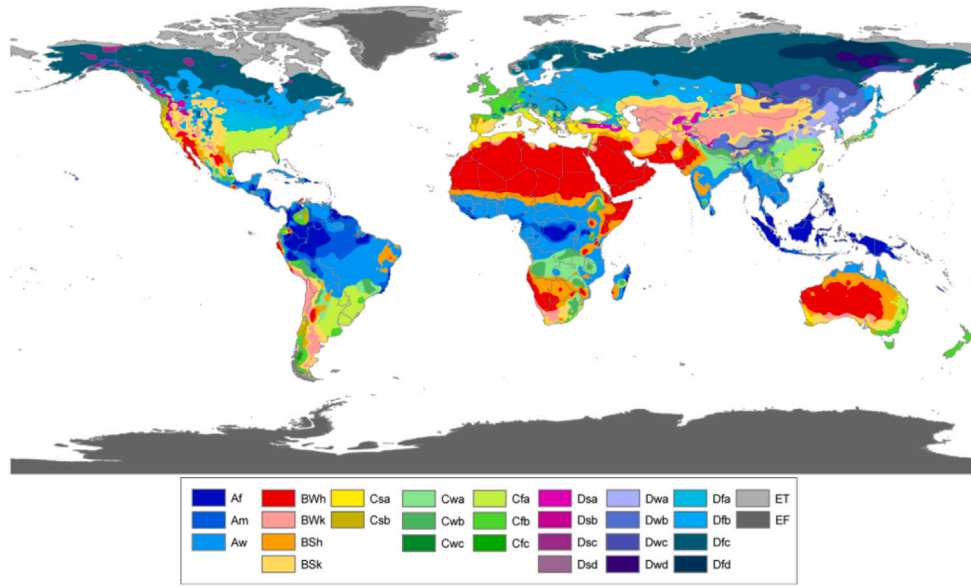


Fig. 8. 'Köppen World Map High Resolution' by Peel, M. C. et al. [67], licenced under Creative Commons Attribution-Share Alike 3.0 Unported [68], Desaturated from original.

**Table 3**  
Classification of the ASHRAE RP-884 database for HVAC buildings according to climates.

Climate	Number of cities	Instances
Tropical	5 (Townsville, Jakarta, Darwin, Bangkok, and Singapore)	3826
Dry	6 (Honolulu, Kalgoorlie-Boulder, Karachi, Quettar, Multan, and Peshawar)	3290
Temperate	12 (Brisbane, Melbourne, Athens, South Wales, Sydney, San Francisco, Merseyside, San Ramon, Antioch, Auburn, Oxford, and Saidu)	3512
Continental	3 (Ottawa, Montreal, and Grand Rapids)	2808
All	26	13 436

**Climate zone divisions.** We adopt the Köppen climate classification updated by Peel et al. [67], which is one of the most widely used climate classification systems in the world. As shown in Fig. 8, the Köppen climate classification divides climates into five main climate zones: A (tropical), B (dry), C (temperate), D (continental), and E (polar). Each large climate zone is then divided into several small subzones based on temperature patterns and seasonal precipitation. All specific climates are assigned a main group of climate zones (the first letter).

In our study, the target domain (Philadelphia in the US) belongs to the 'temperate' climate zone. In the source domain, the Scales Project dataset includes 8225 instances from 57 cities in total, and 5411 instances from 32 cities (e.g., Yokohama, Sydney, and Cambridge) were located in the 'temperate' climate zone. The ASHRAE RP-884 database consists of 25 623 thermal comfort responses from 26 cities in total, where 12 cities (e.g., Berkeley, Athens, and Chester) [70] are situated in the same climate zone as Philadelphia.

We run the proposed TL-MLP model and TL-MLP-C\* models with the ASHRAE database and the Scales Project database as the source domain and the Medium US Office dataset as the target domain. In particular, for both proposed models, we only use the data from buildings with HVAC systems in all datasets. For the TL-MLPC\* model, we use the data from the buildings with HVAC systems in the same climate zone as the source domain and the Friends Center building as the target domain.

Besides, we classify the HVAC buildings in the ASHRAE RP-884 database into different climates (see Table 3). The table shows that in the ASHRAE RP-884 database, there are 13 436 observations from buildings with HVAC systems in total and 3512 such observations in

**Table 4**  
Prediction performance for different algorithms on the target dataset.

Algorithm	Accuracy (%)	F1-score (%)
PMV	33.35 (2.40)	32.45 (2.35)
Random	27.23 (1.30)	29.30 (1.40)
KNN	41.43 (2.95)	41.93 (2.85)
SVM (Linear)	29.44 (5.19)	30.92 (4.84)
SVM (RBF)	37.93 (3.86)	40.91 (4.04)
SVM (Poly)	34.02 (4.59)	37.66 (5.15)
Decision tree	43.33 (4.94)	43.34 (4.87)
Random forest	51.41 (3.03)	52.93 (3.69)
Naive Bayes	40.43 (4.10)	39.40 (3.97)
AdaBoost	42.94 (3.22)	42.41 (3.94)
MLP	50.35 (3.81)	50.67 (4.51)
TL-MLP	50.76 (4.31)	53.60 (4.43)
TL-MLP-C*	54.50 (4.16)	55.12 (4.14)

the 'temperate' climate zone. Since the Scales Project dataset recorded the Köppen climate and HVAC status information during the data collection, after calculation, there were 4621 observations from buildings with HVAC systems in total and 3245 observations collected from buildings with HVAC systems located in the 'temperate' climate zone.

5.2. Overall prediction result

Table 4 shows the performance of different thermal comfort modelling algorithms. We use all ten features described in Section 4.2 on most algorithms except for the PMV model. From Table 4, we can see



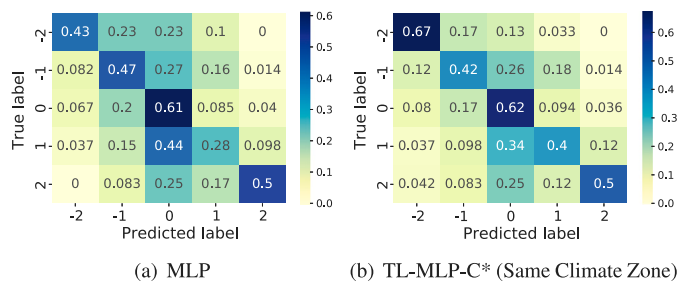


Fig. 9. Confusion matrix on the target domain.

that the PMV model performs better than only the random baseline and SVM classifiers (kernel = ‘Linear’) in accuracy. The F1-score of the linear SVM is still higher than that of the PMV model. This may be because we use more features in machine learning classifiers while the PMV model only has six factors. We will discuss the prediction performance with different feature sets later in Section 5.3.

Table 4 shows that the random forest algorithm performs the best on all metrics compared with the PMV model, random baseline and other data-driven models including eight traditional machine learning classifiers. This may be because the random forest is usually regarded as the best classification algorithm for small datasets [35] and has been proven to have the highest prediction accuracy for thermal sensation [26].

Most importantly, we find that the TL-MLP has a higher F1-score for thermal comfort classification than other machine learning methods without using transfer learning. Although the TL-MLP has better prediction performance than the MLP on all metrics, the prediction accuracy of the TL-MLP is slightly lower than that of the random forest. The potential reason is that the TL-MLP transfers knowledge from all HVAC buildings in the world regardless of the different climate zones, leading to lower prediction accuracy than that of the random forest. Excitingly, the TL-MLP-C\* model works better than all of the state-of-the-art algorithms on both metrics (accuracy and F1-score), indicating the effectiveness of the proposed approach.

To further investigate how the proposed TL-MLP-C\* improves the prediction performance compared to the MLP, we show the confusion matrices for the MLP and TL-MLP-C\* in Fig. 9. The figure shows that the MLP model can predict label 0 (neutral) with the highest probability of 0.61, which is similar to the 0.62 of the TL-MLP-C\*. However, it still has high chances to misclassify labels 1 (slightly warm) to 0 (neutral). Instead, the transfer learning-based thermal comfort model TL-MLP-C\* can predict labels more accurately than the traditional MLP model, especially for the minority classes (-2, -1, 1). It can predict 67% of the label -2 (cool or cold) instances and 40% of the label 1 (slightly warm) instances correctly and achieves an average accuracy of 54.50% for all classes from -2 to 2.

In summary, our proposed transfer learning-based models (TL-MLP and TL-MLP-C\*) achieve remarkable performance for thermal comfort prediction compared with the random baseline, traditional PMV model and data-driven algorithms without transfer learning. In particular, the TL-MLP-C\* model outperforms the state-of-the-art algorithms on both metrics (accuracy and F1-score). Furthermore, the improved prediction performance of the TL-MLP-C\* is significant compared to that of the standard MLP model.

### 5.3. Impact of different feature combinations

We will now explore how accurately the proposed TL-MLP and TL-MLP-C\* models work when only a set of features is available. Usually, indoor sensors are inexpensive and unobtrusive and have been installed in many buildings with HVAC systems. However, some features may be unavailable due to factors such as privacy, costs, etc. For instance,

Table 5

Prediction performance for different feature sets on the target dataset.

Sets	Algorithm	Accuracy (%)	F1-score (%)
$\mathcal{X}_a$	PMV	33.35	32.45
	Random Forest	<b>34.77</b>	34.92
	MLP	33.18	34.06
	TL-MLP	33.53	35.90
	<b>TL-MLP-C*</b>	33.98	<b>39.32</b>
$\mathcal{X}_b$	Random Forest	43.43	43.18
	MLP	42.96	45.31
	TL-MLP	44.10	45.88
	<b>TL-MLP-C*</b>	<b>47.10</b>	<b>51.15</b>
$\mathcal{X}_c$	Random Forest	51.41	52.93
	MLP	50.35	50.67
	TL-MLP	50.76	53.60
	<b>TL-MLP-C*</b>	<b>54.50</b>	<b>55.12</b>

occupants may not be willing to report their age, which reflects their metabolism level and influence their thermal comfort feelings. Besides, it is somewhat inconvenient to install outdoor weather stations outside a building to capture outdoor environmental changes (e.g., outdoor air temperature and humidity) more accurately than the official weather stations used for local weather forecasting.

Hence, in the experiment, we will divide our features into 3 different sets  $\mathcal{X}_a$ ,  $\mathcal{X}_b$ , and  $\mathcal{X}_c$  based on PMV factors, personal factors and outdoor environmental factors, respectively; and then compare the different sets and explore which features contribute the most to effective thermal comfort transfer learning. The feature sets are as follows:

- $\mathcal{X}_a$ : Six basic factors introduced in the PMV model: indoor air temperature, indoor air velocity, indoor relative humidity, indoor radiant temperature, clothing insulation and metabolic rate. This is the most common feature set for thermal comfort modelling used in previous studies [35].
- $\mathcal{X}_b$ : Six factors from  $\mathcal{X}_a$  and two personal factors: age and gender. Personal factors such as gender and age can be easily collected through background surveys.
- $\mathcal{X}_c$ : Eight factors from  $\mathcal{X}_b$  and two outdoor environmental factors including the outdoor air temperature and outdoor relative humidity. The above two outdoor environmental features need to be accessed from the outdoor weather station near the target building.

For different feature sets, we use the same oversampling methods and fixed random seeds in neural network training. Table 5 shows the prediction performance for different feature sets on the target dataset. The random forest and MLP algorithms are chosen for comparison with the TL-MLP and TL-MLP-C\* algorithms due to their relatively high performance, as shown in Table 4. For the  $\mathcal{X}_a$ ,  $\mathcal{X}_b$ , and  $\mathcal{X}_c$  feature sets, we can observe that the performance of the TL-MLP and TL-MLP-C\* models increases as the number of features increases. In addition, the TL-MLP-C\* model has the highest accuracy and F1-score in each feature set.

For feature set  $\mathcal{X}_a$ , the PMV model works slightly better than the MLP model in accuracy but worse in F1-score. The random forest algorithm achieves the best performance in accuracy while TL-MLP-C\* achieves the highest F1-score. With transfer learning from source datasets, the TL-MLP and TL-MLP-C\* have similar prediction accuracies to the traditional PMV model. This shows that the advantages of the proposed TL-MLP and TL-MLP-C\* models cannot be fully utilized when the number of features is limited.

In feature set  $\mathcal{X}_b$ , all data-driven models achieve better prediction performance than using only feature set  $\mathcal{X}_a$ . This shows that personal information (age and gender) could effectively improve thermal comfort prediction. Moreover, the TL-MLP-C\* model has the best prediction performance compared with the other methods in both metrics when considering personal factors.

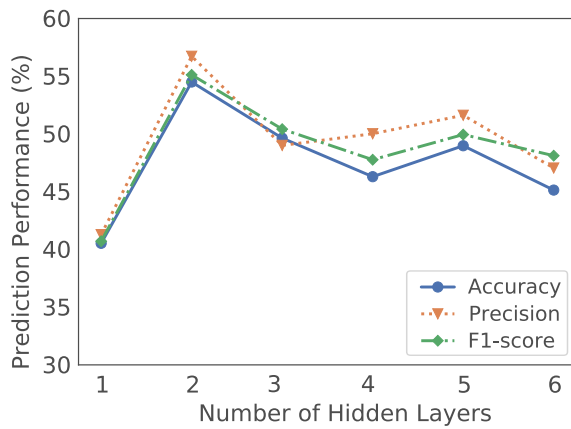


Fig. 10. Prediction performance with different number of hidden layers.

In comparison to feature sets  $\mathcal{X}_a$  and  $\mathcal{X}_b$ , the random forest, MLP, TL-MLP and TL-MLP-C\* work best among all metrics on the feature set  $\mathcal{X}_c$ . This proves that outdoor environmental changes can affect occupants' thermal sensation in HVAC buildings and shows the necessity to consider outdoor features for effective thermal comfort modelling.

#### 5.4. Impact of the number of hidden layers

We also conduct adaption experiments by using different numbers of hidden layers in the TL-MLP-C\* model. Fig. 10 shows the prediction accuracy and F1-score for TL-MLP-C\* with different numbers of hidden layers. We can observe that the prediction performance is worst in all metrics with only one hidden layer. Since our proposed method transfers the last layer of the hidden layer, if we set only one hidden layer, the target dataset will have little contribution to the prediction model. When the number of hidden layers is set to 2, the proposed TL-MLP-C\* model has the highest prediction performance in accuracy and F1-score. As the number of hidden layers continues to increase, the prediction performance tends to decrease, which may be due to the model being overfitted with more trainable parameters.

Finally, although our proposed TL-MLP-C\* model has better thermal comfort prediction performance than the state-of-the-art methods, the achieved accuracy (54.50%) is still not remarkably high. There are several potential reasons: (1) We adopt the *TabularGAN* to resample the minority classes for meaningful classification. Fifty percent of the instances in each class were synthesized while ensuring that the number of samples per category did not exceed the number of samples in the majority class. Although some previous works achieve slightly higher accuracy for thermal comfort prediction (e.g., 63.09% in [35] and 62% in [26]), they only assigned slightly higher weights to the instances in the minority classes, which cannot handle the class imbalance problem as well as our method. (2) Predicting thermal comfort is challenging since many factors affect occupants' thermal sensation (as discussed in Section 1). There may also be many response biases during the survey. Therefore, the classification accuracy in most previous research is also not good and rarely higher than 60%, even for personal thermal comfort modelling. (3) It could be better to regard the thermal comfort prediction as a regression problem instead of a classification problem. For example, classifying '-2' (cool) to '-1' (slightly cool) should be more acceptable than classifying '-2' (cool) to '+2' (warm). We will study the thermal comfort regression in future work.

## 6. Conclusion

A huge amount of sensor data has been generated in cities worldwide. Recently, utilizing such data from multiple cities to benefit a target city has become a critical issue. In this research, we applied

the idea of transfer learning to the thermal comfort area and proposed two transfer learning-based thermal comfort prediction models: TL-MLP and TL-MLP-C\*. For the first time, we transferred the knowledge from similar thermal environments to a target building for effective thermal comfort modelling. Furthermore, we improved the prediction performance and built meaningful classifiers by using a GAN-based re-sampling method (i.e., *TabularGAN*) to imbalance the class distribution of occupants' thermal sensation.

By retaining the last hidden layer of the neural network from the source domain (ASHARE RP-884 and Scales Project datasets), we trained the thermal comfort model for the Friends Center building from the Medium US Office dataset and found the optimal parameter settings for lower hidden layers. Extensive experimental results showed that the proposed TL-MLP and TL-MLP-C\* models outperform the state-of-the-art algorithms for thermal comfort prediction. Interestingly, the most significant feature sets are identified for effective thermal comfort transfer learning.

This research provides the possibility of building thermal comfort models with limited data. The publicly available thermal comfort data from similar climate zones can be used to benefit the thermal comfort modelling in the target building. However, the current studies have some limitations that needed to be addressed in future research: (1) First, we only used the Friends Center Office as the target building which is located in the 'temperate' climate zone. The performance of transfer learning on more target buildings in the same or different climate zones should be explored in the future. (2) Although the proposed method can benefit the target building with a small amount of labelled data, the prediction model will achieve the best performance when at least six factors are provided. In real-world scenarios, if there are only several factors (lower than six) in the target building, our method can still work by setting the values of missing factors to the same distribution as the source domain, but the prediction performance for the target building will be affected. (3) We only investigated the MLP models but more advanced transfer learning architectures can be explored to find transferable representations between the source domain and target domain in future studies.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgements

This research was supported by the Australian Government through the Australian Research Council's Linkage Projects funding scheme (project LP150100246). This paper is also a contribution to the IEA EBC Annex 79.

#### References

- [1] X. Wang, J. Liono, W. Mcintosh, F.D. Salim, Predicting the city foot traffic with pedestrian sensor data, in: Proceedings of the 14th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services, ACM, 2017, pp. 1–10.
- [2] W. Shao, Y. Zhang, B. Guo, K. Qin, J. Chan, F.D. Salim, Parking availability prediction with long short term memory model, in: International Conference on Green, Pervasive, and Cloud Computing, Springer, 2018, pp. 124–137.
- [3] H. Song, A. Qin, F.D. Salim, Evolutionary model construction for electricity consumption prediction, *Neural Comput. Appl.* (2019) 1–18.
- [4] S. Zhan, A. Chong, Building occupancy and energy consumption: Case studies across building types, *Energy Built Environ.* (2020).
- [5] W.H. Sadid, S.A. Abobakr, G. Zhu, Discrete-event systems-based power admission control of thermal appliances in smart buildings, *IEEE Trans. Smart Grid* 8 (6) (2017) 2665–2674.
- [6] M. Rahman, M. Rasul, M.M.K. Khan, Energy conservation measures in an institutional building in sub-tropical climate in australia, *Appl. Energy* 87 (10) (2010) 2994–3004.

- [7] M. Frontczak, P. Wargocki, Literature survey on how different factors influence human comfort in indoor environments, *Buil. Environ.* 46 (4) (2011) 922–937.
- [8] Refrigerating, A.-C. Engineers, A.N.S. Institute A.S. of Heating, Thermal environmental conditions for human occupancy, in: American Society of Heating, Refrigerating and Air-Conditioning Engineers, Vol. 55, 2004.
- [9] N. Gao, W. Shao, M.S. Rahaman, F.D. Salim, n-Gage: Predicting in-class emotional, behavioural and cognitive engagement in the wild, *Proc. ACM Interact. Mobile Wearable Ubiquit. Technol.* 4 (3) (2020) 1–26.
- [10] M.S. Rahaman, J. Liono, Y. Ren, J. Chan, S. Kudo, T. Rawling, F.D. Salim, An ambient-physical system to infer concentration in open-plan workplace, *IEEE Internet Things J.* (2020).
- [11] P.O. Fanger, et al., Thermal comfort. analysis and applications in environmental engineering. Thermal comfort, *Anal. Appl. Environ. Eng.* (1970).
- [12] R. Becker, M. Paciuk, Thermal comfort in residential buildings—failure to predict by standard model, *Buill. Environ.* 44 (5) (2009) 948–960.
- [13] F. Auffenberg, S. Stein, A. Rogers, A personalised thermal comfort model using a bayesian network, in: Twenty-Fourth International Joint Conference on Artificial Intelligence, 2015.
- [14] C. Chun, A. Kwok, T. Mitamura, N. Miwa, A. Tamura, Thermal diary: Connecting temperature history to indoor comfort, *Buill. Environ.* 43 (5) (2008) 877–885.
- [15] T. Chaudhuri, D. Zhai, Y.C. Soh, H. Li, L. Xie, Random forest based thermal comfort prediction from gender-specific physiological parameters using wearable sensing technology, *Energy Build.* 166 (2018) 391–406.
- [16] O. Seppänen, W. Fisk, M. Mendell, Association of ventilation rates and co2 concentrations with health and other responses in commercial and institutional buildings, *Indoor air* 9 (4) (1999) 226–252.
- [17] M. Indraganti, K.D. Rao, Effect of age gender economic group and tenure on thermal comfort: a field study in residential buildings in hot and dry climate with seasonal variations, *Energy Build.* 42 (3) (2010) 273–281.
- [18] K. Cena, R.J. de Dear, Field study of occupant comfort and office thermal environments in a hot, arid climate, *ASHRAE Trans.* 105 (1999) 204.
- [19] R. De Dear, G.S. Brager, Developing an adaptive model of thermal comfort and preference, 1998.
- [20] M. Schweiker, A. Abdul-Zahra, M. André, F. Al-Attrash, H. Al-Khatiri, R.R. Alprianti, H. Alsaad, R. Amin, E. Ampatzi, A.Y. Arsano, et al., The scales project, a cross-national dataset on the interpretation of thermal perception scales, *Sci. data* 6 (1) (2019) 1–10.
- [21] J. Langevin, P.L. Gurian, J. Wen, Tracking the human-building interaction: A longitudinal field study of occupant behavior in air-conditioned offices, *J. Environ. Psychol.* 42 (2015) 94–115.
- [22] L. Bourikas, E. Costanza, S. Gauthier, P. James, J. Kittley-Davies, C. Ornaghi, A. Rogers, E. Saadatian, Y. Huang, Camera-based window-opening estimation in a naturally ventilated office, *Buill. Res. Inf.* 46 (2) (2018) 148–163.
- [23] A.K. Clear, J. Morley, M. Hazas, A. Friday, O. Bates, Understanding adaptive thermal comfort: new directions for ubiComp, in: Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing, ACM, 2013, pp. 113–122.
- [24] J. Ranjan, J. Scott, Thermalsense: determining dynamic thermal comfort preferences using thermographic imaging, in: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, ACM, 2016, pp. 1212–1222.
- [25] A. Ghahramani, G. Castro, S.A. Karvigh, B. Becerik-Gerber, Towards unsupervised learning of thermal comfort using infrared thermography, *Appl. Energy* 211 (2018) 41–49.
- [26] M. Luo, J. Xie, Y. Yan, Z. Ke, P. Yu, Z. Wang, J. Zhang, Comparing machine learning algorithms in predicting thermal sensation with ashrae comfort database ii, *Energy Build.* (2020) 109776.
- [27] P. Ferreira, A. Ruano, S. Silva, E. Conceicao, Neural networks based predictive control for thermal comfort and energy savings in public buildings, *Energy Build.* 55 (2012) 238–251.
- [28] W. Hu, Y. Wen, K. Guan, G. Jin, K.J. Tseng, Itcm: Toward learning-based thermal comfort modeling via pervasive sensing for smart buildings, *IEEE Internet Things J.* 5 (5) (2018) 4164–4177.
- [29] W. Zhang, W. Hu, Y. Wen, Thermal comfort modeling for smart buildings: A fine-grained deep learning approach, *IEEE Internet Things J.* (2018).
- [30] W. Dai, O. Jin, G.-R. Xue, Q. Yang, Y. Yu, Eigentransfer: a unified framework for transfer learning, in: Proceedings of the 26th Annual International Conference on Machine Learning, ACM, 2009, pp. 193–200.
- [31] P.G. Shivakumar, P. Georgiou, Transfer learning from adult to children for speech recognition: Evaluation, analysis and recommendations, 2018, arXiv preprint arXiv:1805.03322.
- [32] J. Wang, Y. Chen, L. Hu, X. Peng, S.Y. Philip, Stratified transfer learning for cross-domain activity recognition, in: 2018 IEEE International Conference on Pervasive Computing and Communications (PerCom), IEEE, 2018, pp. 1–10.
- [33] J. Ye, Slearn: Shared learning human activity labels across multiple datasets, in: 2018 IEEE International Conference on Pervasive Computing and Communications (PerCom), IEEE, 2018, pp. 1–10.
- [34] E.L.A. Natarajan, A transfer active learning framework to predict thermal comfort, *Int. J. Prognostics Health Manage.* 10 (2019) 13.
- [35] W. Hu, Y. Luo, Z. Lu, Y. Wen, Heterogeneous transfer learning for thermal comfort modeling, in: Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, 2019, pp. 61–70.
- [36] D.H.C. Toe, T. Kubota, Development of an adaptive thermal comfort equation for naturally ventilated buildings in hot-humid climates using ashrae rp-884 database, *Front. Archit. Res.* 2 (3) (2013) 278–291.
- [37] J. Langevin, J. Wen, P.L. Gurian, Relating occupant perceived control and thermal comfort: Statistical analysis on the ashrae rp-884 database, *HVAC R Res.* 18 (1–2) (2012) 179–194.
- [38] S. Lu, W. Wang, C. Lin, E.C. Hameen, Data-driven simulation of a thermal comfort-based temperature set-point control with ashrae rp884, *Buill. Environ.* 156 (2019) 137–146.
- [39] J. von Grabe, S. Winter, The correlation between pmv and dissatisfaction on the basis of the ashrae and the mcintyre scale—towards an improved concept of dissatisfaction, *Indoor Built Environ.* 17 (2) (2008) 103–121.
- [40] M. Schweiker, A. Wagner, The effect of occupancy on perceived control, neutral temperature, and behavioral patterns, *Energy Build.* 117 (2016) 246–259.
- [41] N. Gao, W. Shao, F.D. Salim, Predicting personality traits from physical activity intensity, *IEEE Comput. Mag.* (2019).
- [42] M. Kabloi, A review of transfer learning algorithms, 2017.
- [43] A. Arnold, R. Nallapati, W.W. Cohen, A comparative study of methods for transductive transfer learning, in: ICDM Workshops, 2007, pp. 77–82.
- [44] L. Zhang, D. Wei, Y. Hou, J. Du, Z. Liu, G. Zhang, L. Shi, et al., Outdoor thermal comfort of urban park—a case study, *Sustainability* 12 (5) (2020) 1961.
- [45] M. Indraganti, R. Ooka, H.B. Rijal, Thermal comfort in offices in india: behavioral adaptation and the effect of age and gender, *Energy Build.* 103 (2015) 284–295.
- [46] S. Karjalainen, Gender differences in thermal comfort and use of thermostats in everyday thermal environments, *Buill. Environ.* 42 (4) (2007) 1594–1603.
- [47] D.L. Wilson, Asymptotic properties of nearest neighbor rules using edited data, *IEEE Trans. Syst. Man Cybern.* (3) (1972) 408–421.
- [48] I. Tomek, Two modifications of cnn, *IEEE Trans. Syst. Man and Cybern.* 6 (1976) 769–772.
- [49] N.V. Chawla, K.W. Bowyer, L.O. Hall, W.P. Kegelmeyer, Smote: synthetic minority over-sampling technique, *J. Artif. Intell. Res.* 16 (2002) 321–357.
- [50] H. He, Y. Bai, E.A. Garcia, S. Li, Adasyn: Adaptive synthetic sampling approach for imbalanced learning, in: 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), IEEE, 2008, pp. 1322–1328.
- [51] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, *Adv. Neural Inform. Process. Syst.* 27 (2014) 2672–2680.
- [52] N. Gao, H. Xue, W. Shao, S. Zhao, K.K. Qin, A. Prabowo, M.S. Rahaman, F.D. Salim, Generative adversarial networks for spatio-temporal data: A survey, 2020, arXiv preprint arXiv:2008.08903.
- [53] N. Park, M. Mohammadi, K. Gorde, S. Jajodia, H. Park, Y. Kim, Data synthesis based on generative adversarial networks, *Proc. VLDB Endow.* 11 (10) (2018) 1071–1083.
- [54] L. Xu, K. Veeramachaneni, Synthesizing tabular data using generative adversarial networks, 2018, arXiv preprint arXiv:1811.11264.
- [55] M. Quintana, C. Miller, Towards class-balancing human comfort datasets with gans, in: Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, 2019, pp. 391–392.
- [56] S.J. Pan, Q. Yang, A survey on transfer learning, *IEEE Trans. Knowl. Data Eng.* 22 (10) (2009) 1345–1359.
- [57] I.I.I. H. Daume, D. Marcu, Domain adaptation for statistical classifiers, *J. Artif. Intell. Res.* 26 (2006) 101–126.
- [58] W. Dai, Q. Yang, G.-R. Xue, Y. Yu, Self-taught clustering, in: Proceedings of the 25th International Conference on Machine Learning, ACM, 2008, pp. 200–207.
- [59] M. Abouelenien, M. Burzo, R. Mihalcea, K. Rusinek, D. Van Alstine, Detecting human thermal discomfort via physiological signals, in: Proceedings of the 10th International Conference on Pervasive Technologies Related to Assistive Environments, 2017, pp. 146–149.
- [60] M.C.G. da Silva, J. Pires, A. Loureiro, L. Pereira, P. Neto, A. Gaspar, D. Viegas, N. Soares, M. Oliveira, J. Costa, Spreadsheets for the calculation of thermal comfort indices pmv and ppd, 2014, URL: <http://www.researchgate.net>.
- [61] K. Fukunaga, P.M. Narendra, A branch and bound algorithm for computing k-nearest neighbors, *IEEE Trans. Comput.* 100 (7) (1975) 750–753.
- [62] I. Rish, et al., An empirical study of the naive bayes classifier, in: IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence, Vol.3, 2001, pp. 41–46.
- [63] J.A. Suykens, J. Vandewalle, Least squares support vector machine classifiers, *Neural Process. Lett.* 9 (3) (1999) 293–300.
- [64] S.R. Safavian, D. Landgrebe, A survey of decision tree classifier methodology, *IEEE Trans. Syst. Man Cybern.* 21 (3) (1991) 660–674.
- [65] A. Liaw, M. Wiener, et al., Classification and regression by randomforest, *R news* 2 (3) (2002) 18–22.

- [66] X.-z. Z.H.A.O, X. W.A.N.G, X.-f. Z.H.U., Research of samples selection in eye detection based on adaboost algorithm, *Comput. Technol. Dev.* 2 (2010).
- [67] M.C. Peel, B.L. Finlayson, T.A. McMahon, Updated world map of the köppen-geiger climate classification, 2007.
- [68] Creative commons license deed, URL <https://creativecommons.org/licenses/by-sa/3.0/deed.en>.
- [69] Y. Bengio, Y. Grandvalet, No unbiased estimator of the variance of k-fold cross-validation, *J. Mach. Learn. Res.* 5 (Sep) (2004) 1089–1105.
- [70] T.L. Yong, H. Djamil, Exploring köppen-geiger climate classification of the ashrae rp-884 database, *Int. J. Recent Technol. Eng.* 7 (6) (2019) 854–860.