The colours of comfort: From thermal sensation to person-centric thermal zones for adaptive building strategies

Authors

Stephanie Gauthier[1*], Leonidas Bourikas[2], Farah Al-Atrash[3], Chihye Bae[4], Chungyoon Chun[4], Richard de Dear[5], Runa T. Hellwig[6], Jungsoo Kim[5], Suhyun Kwon[4], Rodrigo Mora[7], Himani Pandya[8], Rajan Rawal[8], Federico Tartarini[9], Rohit Upadhyay[7], Andreas Wagner[10]

Affiliations

[1] Energy & Climate Change Group, Faculty of Engineering and Physical Sciences, University of Southampton, Burgess Road, SO16 7QF, Southampton, UK, s.gauthier@soton.ac.uk

[2] Lancaster Institute for the Contemporary Arts, Lancaster University, LA1 4YW, Lancaster, UK

[3] Architecture and Interior Architecture, German Jordanian University, Mu'ath bin Jabal Street, 35247, Amman, Jordan

[4] Interior Architecture and Built Environment, Yonsei University, 50 Yonsei-Ro Seodaemun-Ku, 3722, Seoul, Republic of Korea

[5] School of Architecture, Design and Planning, The University of Sydney, 148 City Road, 2006, Sydney, Australia

[6] Department of Architecture, Design and Media Technology CREATE, Aalborg University, Rendsburggade 14, 9000, Aalborg, Denmark

[7] Building Science, British Columbia Institute of Technology (BCIT), 3700 Willingdon Avenue, V5G 3H2, Vancouver, Canada

[8] Centre for Advanced Research in Building Science and Energy, CEPT University, K.L.Campus, Navarangpura, 380009, Ahmedabad, India

[9] Sustainable Buildings Research Centre, University of Wollongong, Squires Way, Innovation Campus, 2519, Wollongong, Australia

[10] Building Science Group (Fachgebiert Bauphysik & Technischer Ausbau), Karlsruhe Institute of Technology (KIT), Englerstr. 7, D, 76131, Karlsruhe, Germany

Keywords

Adaptive thermal comfort; Clustering; Individual comfort profiles; Person-centric thermal zones; Adaptive building strategies

Abstract

Thermal comfort research has been traditionally based on cross-sectional studies and spatial aggregation of individual surveys at building level. This research design is susceptible to compositional effects and may lead to error in identifying predictors to thermal comfort indices, in particular in relation to adaptive mechanisms. A relationship between comfort and different predictors can be true at an individual level but not evident at the building level. In addition, cross-sectional studies overlook temporal changes in individual thermal perception due to contextual factors. To address these limitations, this study applied a longitudinal research design over 8 to 21 months in eight buildings located in six countries around the world. The dataset comprises of 5,567 individual thermal comfort surveys from 258 participants. The analysis aggregated survey responses at participant level and clustered participants according to their thermal sensation votes (TSV). Four TSV clusters were introduced, representing four different thermal sensation traits. Further analysis reviewed the probability of cluster membership in relation to demographic characteristics and behavioural adaptation. Finally, the analysis at individual level enabled the introduction of a new metric, the thermal zone (Zt), which in this study ranges from 21.5ºC to 26.6ºC. The thermal sensation traits and person-centric thermal zone (Zt) are a first step into the development of new metrics incorporating individual perceived comfort into dynamic building controls for adaptive buildings.

Highlights

- Review of longitudinal comfort surveys (N= 5,576) from 258 participants and 6 countries.
- Clustering of individual thermal sensation votes to identify four thermal sensation traits.
- Introduction of the person-centric thermal zone (Zt) to inform adaptive building strategies.

1. Introduction

1.1 Research background

Thermal comfort is widely recognised as one of the key parameters in relation to the design and operation of buildings [1]. The quality of the indoor environment has been associated with staff productivity [2] [3] and health [4] [5] [6]. Occupants' dissatisfaction with the indoor environment has been identified as one of the main drivers of interaction with the building controls and systems [1] [7]. Modern office buildings are likely to have a Building Management System (BMS) with multi-zonal controls and input by environmental sensors [1] [8]. Despite the feedback from these sensors, building managers often have to override the BMS controls to address occupants' comfort related complaints. The building managers' response is to manually adjust the heating, ventilation and air conditioning (HVAC) system settings resulting in unnecessary energy use without usually achieving an increase of occupants' satisfaction. In shared open plan office spaces in particular, the cause of complaints is likely to be related with zonal micro-climatic conditions. These issues can be due to the design or commissioning of the systems, the change of use of the room, spatiotemporal variations such as direct sunlight and occupants' variations in perceived thermal comfort, thermal expectations and thermal preferences.

Thermal comfort research, current guidelines and standards are based on the thermal votes of a large group of people. This entails cross-sectional studies and spatial aggregation of individual surveys at room or building level. A cross-sectional research design does not address temporal variations within a space or within participants (i.e. variation in individual thermal sensitivity and preference). The indoor environment of office buildings is mainly designed to comply with standard indoor air quality requirements that focus on ventilation rates. With regard to thermal comfort, employers need to provide a "reasonable temperature" in the context of health and safety in the workspace. Standards [9] [10] [11] and guidance [12] [13] have been developed for the interpretation of "reasonable" temperature into recommendations for thermal comfort in office buildings. Common practices in industry assess the occupants' comfort with empirical models that rely heavily on assumptions of occupancy and activity. These assumptions are made at group level rather than individual level. These may be one of the reasons for the observed gap between experienced/perceived and expected/preferred indoor environmental conditions [12], as compositional effects may occur. A relationship between comfort and different predictors can be true at an individual level but not evident at the building level.

The adaptive thermal comfort model was introduced as an alternative to the predictive model. This model accounts for occupants' adaptive behaviour to outdoor weather forcing [14]. The fundamental adaptive principle is that people will always react to thermal changes in order to restore their comfort [7]. In the adaptive model the comfort temperature is a function of the outdoor ambient temperature. In particular the comfort temperature for free-running buildings is based on the estimation of the exponential-weighted outdoor temperature running mean (T_{rm}) [15]. The transferability and validity of the equation however rely on the selection of a representative value for the constant 'α' in the T_{rm} equation and the selection of the "half-life", the time step interval for the calculation of the running mean that might not be uniformly applicable to all climates [16]. In addition, modern lifestyles often prolong the time people spend in conditioned environments every day. As a result, people might adapt to the conditioned environment and loose the thermal connection with the outdoor environmental conditions [17] [18]. Individual thermal sensation and preferences also imply diversity in the individual adaptive mechanisms and different adaptation time [19]. To explore the variation within individuals' thermal sensation and preference, a longitudinal research design would be more appropriate instead of the typical cross-sectional survey design.

Recently, thermal comfort research has expanded on the development and use of personalized comfort models [20] [21] [22] [23]. However, it still remains unclear as to when and how these personal comfort models could be applied in working spaces. Although some buildings will opt for entirely passive design with natural ventilation, passive cooling and heating, most of office buildings have mechanical systems to comply with standards and supplement any natural ventilation or passive cooling strategies. These systems include mechanical ventilation, comfort cooling, comfort heating and air conditioning. To successfully combine both aspects, many new buildings will opt for mixed mode ventilation strategies (MM), where natural and mechanical approaches are combined. The system may operate concurrently or change-over on seasonal or daily basis. This combined, hybrid approach can save energy, increase the usable space, reduce the operation and maintenance costs and increase occupants' satisfaction [24].

As mentioned above, individually controlled spaces may not be possible in shared office. Yet, if individuals' thermal preference was grouped in a generic thermal zone, then few settings may be implemented within a shared space. There are two main approaches to group individual profiles. Firstly, through a deductive process and supervised techniques, where regression or classification analysis are applied. Secondly, through an inductive process and unsupervised techniques, where clustering or association analysis are applied. The strength of the second approach is that no prior knowledge or hypothesis is imposed on the data. Such techniques have been applied in many fields of research, from game theory [26] to profiling energy customers [27] [28]. As people's physical, psychological and contextual characteristics vary greatly; individual thermal sensation, preference and comfort are likely to vary between people but also within a person itself. Clustering analysis could identify groups of individuals with similar thermal comfort variations; leading to personalised thermal environments. Current research has assessed individual thermal sensitivity to integrate individual comfort profiles in multi-occupancy spaces' HVAC control systems [25]. Low energy, adaptive buildings would benefit from personalised controls with online training of new adaptive algorithms that dynamically adjust the building systems' operation and proactively adapt the environmental conditions according to historical data, automated occupancy detection and the thermal preferences of the people currently in each zone [25].

Adaptive comfort regression models have been developed to account for the active response of people to environmental changes in order to achieve comfort [13]. Linear regression models are used to examine the relationship between thermal sensation and indoor operative temperature. The neutral temperature (T_n) (i.e. average thermal sensation is neutral) is calculated by solving the regression between thermal sensation (TSV) and operative temperature (T_{op}) [12]. PMV at this neutral temperature is expected to have a value within [-0.5 to 0.5]. T_n is sensitive to three factors, the regression coefficient, the residual standard deviation of thermal sensation and the sample size [16]. Several datasets have been used for the estimation of T_n for different outdoor temperatures [16]. There are two prevalent methodologies for the estimation of T_n in buildings. de Dear and Brager [12] introduced a method that assesses T_n by the significance of the regressions between thermal sensation on indoor operative temperatures ($\alpha = 0.05$). However, in the case of fully air-conditioned buildings, this method is likely to inflate the means of the regression gradients because a low gradient will have a negative effect on the statistical significance [16]. Humphreys, Nicol and Roaf [16] proposed an alternative method that requires only knowledge of the regression coefficient and does not require detailed thermal sensation and operative temperature data for the regression. This approach allows the use of mean values of thermal sensation and operative temperature for the whole dataset. The regression coefficient was empirically derived from the ASHRAE RP-884 and the SCATs databases [12] [16]. The validity of the results depends on the number of observations and the errors introduced from the use of the mean temperature as a predictor variable [16]. This paper applies the method introduced by de Dear and Brager [12], as the dataset was collected from mixed-mode (MM) buildings. Furthermore, the thermal sensation and indoor temperature observations in the dataset have large standard deviations and a wide scatter.

This study is looking to contribute to research on individual thermal comfort in the context of adaptive buildings. Personal comfort and its relation to indoor, outdoor and personal factors were assessed. Individual thermal sensations were clustered to establish personal comfort traits. A new metric, the person-centric thermal zones, is introduced. This metric could be used with automated occupancy detection and individual comfort profiling to optimise HVAC zonal settings and add to the BMS predictive capabilities that would dynamically adapt and transition the indoor environment to the forecasted three-dimensional occupancy (i.e. who, where and when) and external weather forcing.

1.2 Choice of variables

Thermal comfort studies traditionally use indices and scales to quantify subjective questionnaire survey responses such as thermal sensation, comfort and preference. In this study, thermal sensation (feeling) is described through the Thermal Sensation Vote (TSV) using a 7-level scale (graded from -3: cold to +3: hot with 0 being the neutral). Thermal Comfort Vote (TCV) is assessed through a binary question (1: comfortable, 2: uncomfortable). Thermal Preference Vote (TPV) uses a 3-level scale with the central point being “want no change” (-1: want cooler to +1: want warmer). One of the most common indices of warmth is the Standard Effective Temperature (SET). The SET method provides a simplified parameterisation of the principle of equal skin heat losses between an idealised environment (i.e. 50% RH, <0.1 m/s air speed and average radiant temperature equal to the air temperature and an occupant with 1.0 met activity level and 0.6 clo clothing level) and an actual person in the real environment [11]. SET was calculated using the ASHRAE Database II [29] validated comfort calculator. In contrast, the operative temperature (T_{op}) is an environmental variable that varies as a function of the indoor air temperature (T_a), black globe temperature (T_r) and indoor air velocity (V_a). Thirteen variables were selected for the analysis; a summary of the main descriptive statistics is shown in Table 3.

This study evaluated the clustering of individual comfort indices as a method to generalise common person centric adaptive comfort attributes. The analysis is structured around the definition of three distinct thermal zones. Each of these zones is developed with the generalisation of a comfort measure but instead of averaging the individual responses, this study assessed their variability and used it to enhance the role of individual perception and preferences. First, participants' responses were clustered according to the mean ($\bar{x}$) and the standard variation ($_{SD}$) as a measure of spread around the mean (see Section 3.2). Three thermal zones (neutral Z_n, comfort Z_c, preferred Z_p) were defined at the overlapping temperature range as calculated with individual regression models for four distinct clusters of responses. These zones (Z_n, Z_c, Z_p) of intra-cluster temperature cross-section are a generalisation of three measures of comfort; the neutral temperature (T_n), the comfort temperature (T_c) and the preferred temperature (T_p) respectively. The final step synthesized the three thermal comfort indices and the corresponding zones into a single thermal zone, Z_t. The final thermal zone deviates from the commonly used averaging approach to introduce a range of temperature that represents the individual variability as a result of personal adaptation and individual perception attributes within the clusters. A summary of the steps applied to develop of the four zones is described in the following Section.

1.3 Objectives

The aim of this paper is to introduce a method for the clustering of comfort according to individual comfort perception and interpret the results in the context of adaptive building design and operation. The research questions are:

1. Can people be grouped according to their comfort perception by assessing the variance and mean of thermal perception indices (i.e. TSV, TPV, and TCV)?
2. Can these groups (or clusters) be directly associated with neutral (Z_n), comfortable (Z_c) and preferred (Z_p) thermal zones? The intersection amongst these three zones is defined as a thermal zone (Z_t) which represents the generalised person-centric thermal perceptions in each group.
3. Can these groups (or clusters) be directly associated with demographic characteristics and the climatic adaptation processes for the individuals in each group?
4. Can Z_t inform adaptive building design and HVAC operation strategies in non-residential mixed-mode buildings?

2. Research Design

Within the IEA-EBC-Annex 69-Subtask C, a longitudinal field survey was designed to analyse the performance of buildings from the view of indoor comfort, occupant behaviour and energy use. The survey applied a mixed method approach. Objective and subjective, qualitative and quantitative data were collected and analysed. All ethical approvals were obtained including data sharing and data storage requirements. The case study buildings were all selected to be "low-energy" as defined by benchmarks to their local legislation and guidelines. The operation and occupancy of the buildings was well established (at least six months of continuous occupancy) prior to the study. Post-occupancy evaluations were undertaken to choose buildings that were performing well in terms of occupants' satisfaction. The final sample of case study buildings was a compromise between the study prerequisites and the availability of buildings with easy access, permissions to install monitoring equipment and participants for the surveys. The survey was finally deployed in 14 mixed-mode office buildings located in eight countries (Australia, Canada, China, India, Jordan, Republic of Korea, UK and USA). Most of the monitored offices were open plan spaces with individual booth desks. This ensured that the background conditions were consistent across the sample (i.e. location, access to adaptive opportunities/building controls). In few cases, occupants moved offices, desks or even buildings. These participants were included in the sample only for the period they remained at their initial location. The surveys and environmental monitoring lasted from 8 months (1 building) to 21 months (1 building) with the majority lasting 11 months (7 buildings).

2.1 Questionnaires

The questionnaire surveys included an initial, introductory meeting with the participants followed by "right-here-right-now" questionnaires completed at regular intervals throughout the course of the study (8 to 21 months). The introductory meeting provided the participants with contextual information about the study and instructions to the "right-here-right-now" questionnaires. A minimal amount of contextual information was given to the participants to avoid "information" bias; whereby participants may try to conform to the aim of the study and the researchers' expectations, leading to erroneous questionnaire responses. At regular intervals (e.g. weekly), the "right-here-right-now" comfort questionnaires were completed using different modes, including smartphone application, online questionnaires or paper survey. Survey reminders in some of the countries (e.g. UK) were sent on random days and at a random time. This randomisation was considered important in order to avoid introducing bias from specific activities and daily recurring indoor environmental conditions. The "right-here-right-now" questionnaires collected information on: clothing level, perceived thermal sensation (TSV) (7-point ASHRAE scale), perceived thermal comfort (TCV) (2-point scale; 'comfortable/uncomfortable'), thermal preference (TPV) (3-point scale; 'want cooler/want no change/want warmer'), adaptive opportunities in use and perceived indoor environmental conditions (e.g. air movement, noise, air quality, etc.). Although the questionnaires' modes of completion differed, the scales were all ordinal scales. The three scales (TSV, TCV and TPV) uncovered the relative ranking of participants thermal sensation, comfort and preference. The assessment of the scales' intervals between these ranks remain outside the scope of this study.

2.2 Environmental monitoring

Concurrently to the “right-here-right-now" questionnaires, indoor air temperature (T_a), mean radiant temperature (T_r), relative humidity (RH) and in some buildings air velocity (V_a) levels were monitored within each office space at a minimum sampling rate of 15 minutes. Table 3 shows the variables and their main descriptive statistics. The indoor data loggers in the open space offices were located in different zones (the number of zones was associated with the layout of the room and its dimensions) and at the height of the seated participants (~0.8m to 1.1m). Within the eight buildings, the sensors’ minimum accuracy levels adhered to ISO 7726 required accuracies [30]; T_a was ±0.5°C, T_r was ±2°C, RH was ±5% and V_a was $\pm(0.05+0.05V_a)$ m/s with a response time of 0.5 seconds. Example of datalogger deployed included HOBO U12 and MX1102, Sensirion SHT21 and Extech Instrument SD800. In addition, external weather conditions were monitored at a minimum sampling rate of 1 hour. Weather data from nearby local weather stations were preferred to that of city weather data; as these are usually derived from nearby airport locations that do not represent urban locations.

2.3 Data preparation

The initial dataset consisted of 11,484 surveys collected from N=1,909 participants in eight countries. The data preparation was undertaken with the statistical software package R [35]. Missing data from the questionnaires and unknown variables in the dataset were left blank. Four consecutive steps were applied in the data preparation, described as follows:

A. In the first step, the first survey responses of the participants were excluded from the sample to limit potential “information” bias. There is a minimum requirement of seven surveys to allow group differences to be estimated [42], as a result all participants who had answered 8 surveys or less were omitted from the sample. This resulted in the exclusion of 1,639 participants out of the initial 1,909 participants (86% of total), or 2,226 surveys of the initial 11,484 surveys (19% of total). However, it was deemed as necessary in order to address the paper’s research questions. The sample after this first step comprised 9,258 surveys from N=270 participants from seven countries around the world.

B. In a second step, an additional threshold was considered for the maximum number of survey responses per participant. This threshold was necessary because all the responses in the sample should be assessed against the same baseline and have equal weight. If the statistical analysis was based on random number of survey responses then it is expected that the results would not be representative of the total population. In this case, it is highly likely that the generalisation and transferability of the results would have been affected by the weight of specific survey responses from participants with a number of responses much higher or lower (i.e. 1.5 times the standard deviation (SD)) than the mean number of responses per participant. The maximum number of observations per participant was 246, the median was 24 and 75%

of the participants had completed in total around 33 surveys. Therefore 33 surveys per participant were considered to be an appropriate sample in the context of this analysis. The sample after the second step included 5,977 surveys from N=270 participants from seven countries.

C. The third step was to review the variables used in the analysis (see Table 3) and to exclude any missing and non-uniformly reported variables. The dataset from India used a different scale from the ASHARE 7-point Thermal Sensation Vote (TSV). Consequently, any TSV data from the Indian dataset were omitted. Thermal Comfort Vote (TCV) was not surveyed in Australia and South Korea. In addition, four variables (T_{op}, V_a, T_{outDay} and TCV) were not surveyed in the USA sample. It was decided to omit the whole dataset from the USA from the analysis as it was missing many key variables. The sample after step three had 5,581 surveys from N=258 participants in six countries.

D. The fourth step reviewed the range of the variables used in the analysis. Following Humphreys and Nicol [31] approach, no limits were placed on the ranges of the variables. The ranges were compared with those given in ISO 7730 - Section 4.1 on determining the PMV index (see Table 3) [9]. Recordings of T_{op}, RH, V_a and Clo were found to be within the ranges suggested by ISO 7730 [9]. Two recordings of Met were below 0.8 met and three recordings of Met were above 4 met. The corresponding surveys were excluded. The final number of surveys in the analysis was 5,576 from N=258 participants and six countries (Australia, Canada, India, Jordan, South Korea and the UK).

2.4 Dataset

Following the data preparation and cleaning procedures the dataset used in the analysis comprised of 5,581 surveys from N=258 individuals. Table 1 shows the climate classification, the building type, the sample size and the period of data collection for each studied building. The climate was derived from the Köppen-Geiger classification [32]. This classification originally defined climate zones according to the type of prevailing local vegetation. The current classification also considers regional air temperature and precipitation data. Six buildings are in a temperate climate (“C” in Climate column, Table 1) representing 2,476 surveys from 120 participants. One building is in a hot semi-arid climate (“B” in Climate column) representing 932 surveys from 33 participants. One building is in continental climate (“D” in Climate column) representing 2,167 surveys from 105 participants. The number of surveys across seasons were as follows: 1,170 surveys in the Spring (21%), 1,587 surveys in the Summer (28.5%), 1,230 surveys in the Autumn (22%), and 1,588 surveys in the Winter (28.5%). It is noted that the sample has a balanced distribution across the seasons, however some of the climates represented in the study do not have large seasonal variability (e.i. Bsh). All the studied buildings were offices with mixed mode ventilation. This ventilation operation mode was providing heating in winter and peak cooling only when required in summer.

Table 1 Description of the studied mixed-mode buildings

Country	City	Climate*	Bdg. Code	Bdg. Type	Year built	No of participants	No of surveys	Duration of the study (months)
AU	Wollongong (Sydney)	Cfa	C01	University research office	2013	24	716	11
CA	Vancouver	Cfb	C02	Office building	2011	24	611	12
IN	Ahmedabad	Bsh	C05	University research office	2015	33	932	21
JO	Amman	Csa	C06	Office building	2011	18	182	11
JO	Amman	Csa	C07	Office building	2014	9	85	11
KR	Seoul	Dwa	C08	Office building	2012	105	2167	9
GB	Southampton	Cfb	C09	University research office	2010	27	533	12
GB	Southampton	Cfb	C10	University research office	2015	18	349	12

* Climate described as Köppen Climate Classification subtype [32]

Each survey had 182 variables. These variables we grouped into four categories; (1) building information, (2) participant information, (3) responses to "right-here-right-now" surveys and associated environmental variables, and (4) calculated variables. To address the paper's research questions, the chosen unit of analysis was the individual participant. This choice leads to a question of metric, i.e. which metric(s) should be used to quantify and qualify a participant's perceived thermal comfort? The metrics may be a measure of the central tenancy (mean or median) or a measure of the spread (standard deviation or interquartile range) [34]. This paper will use conventional measures of scale which include participants' mean ($\bar{x}$) and standard deviation ($_{SD}$). Interestingly, many of the variables had similar mean and median, which implies that the distributions of these variables were symmetrical (see Table 3).

Table 2 Characteristics of the participants and individual control opportunities

Bdg.	No. of participants	% of participants with perceived individual control						% of participants per office type			Age (%)			Sex (%)	
		Door	Window	Blind	Heating	Cooling	Fan*	Single office	Small shared office **	Large shared office	<30	30 to 39	>39	F	M
C01	31	26	26	-	23	42	13	3	3	94	48	39	13	26	74
C02	24	71	54	54	67	67	-	71	8	21	-	12	88	79	21
C05	34	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C06	60	67	73	70	88	88	-	17	20	63	-	-	-	-	-
C07	50	96	86	64	96	96	4	20	64	16	-	-	-	-	-
C08	131	100	100	100	-	-	100	-	-	100	11	28	21	53	47
C09	41	12	32	32	-	-	-	7	12	73	34	15	41	71	20
C10	28	36	57	57	21	29	-	57	36	7	43	25	21	39	50

* Ceiling, pedestal and/or desk fan ** Small shared office; i.e. 2 to 4 person office

Table 3 Review of the variables and comparison with ISO 7730 ranges [9]

	Variables	Median	Mean	SD	Minimum	Maximum	ISO 7730 minimum	ISO 7730 maximum
1	(T_{op}) Operative temperature (°C)*	24.0	24.4	2.3	18	31	10[a]	35[a]
2	(RH) Relative humidity (%)	44.0	46	13	16	86	0	100
3	(V_a) Mean air velocity (m/s)	0.08	0.08	0.06	0	0.79	0	1
4	(Met) Metabolic rate (met)	1.2	1.2	0.3	0.7	5.4	0.8	4
5	(Clo) Clothing insulation (clo)	0.7	0.7	0.2	0.3	1.7	0	2
6	($T_{outSurv}$) Outdoor air temperature at the time of the survey (°C)	17.3	17.3	11.4	-9.8	43.1		
7	(T_{outDay}) Outdoor air temperature on the day of the survey (°C)	15.3	15.4	10.6	-7	35.5		
8	(TSV) Thermal sensation vote	0.0	0.1	1	-3	3		
9	(TCV) Thermal comfort vote	1.0	1.2	0.4	1	2		
10	(TPV) Thermal preference vote	0.0	0.1	0.6	-1	1		
11	(PMV) Predictive Mean Vote	0.15	0.2	0.6	-3	3	-2	2
12	PPD (%)	8.6	14.3	13.3	5	99		
13	SET (°C)	25.9	26.2	2.5	18.9	37.3		

* assuming that $Va < 0.1$ m/s [13]

Interestingly, a large number of participants perceived that they had control over the building systems even if they worked in large shared office spaces. Previous studies have shown that the perceived control is a result of multiple perceptual and personal factors such as the cultural background [24] [33] [39]. In general, individuals comprising the sample in this study believe that they had control of the door, windows and blinds in their workspace (see Table 2).

2.5 Analysis methods

Similarly, to the data preparation, the data analysis was undertaken with the statistical software package R [35]. The five consecutive steps undertaken in the analysis are as follows:

a) Descriptive statistics of the variables were calculated for the sample and each individual (Table 3). The results were used to quantitative and qualitative review the thermal sensation, comfort and preference variability within and between participants (Section 3.1 of *Results*).

b) In the second step, a K-means clustering algorithm was applied to group individuals according to the mean and the standard deviation of their TSV, TCV and TPV responses. The final number of clusters was established by reviewing plots of within group's sum of squares for the number of factors extracted for each variable (i.e. scree plot). This analysis led to classifying each participant in a cluster for each of the three variables, TSV, TCV and TPV. The intracluster agreement between the TSV, TCV and TPV clusters was compared with the use of scatter plots and the Goodman & Kruskal's gamma following the method proposed in [19] (Section 3.2 of *Results*).

c) In the third step, neutral (T_n), comfortable (T_c) and preferred (T_p) temperatures were estimated for each participant, the entire sample and the four TSV clusters (Section 3.3 of *Results*). Following the method introduced by de Dear and Brager [12], T_n was calculated for each participant by applying the following three steps:

    1. T_{op} was binned into half-degree (◦C) increments (T_{op_bin}). Then, the mean TSV was estimated for each half-degree (◦C) interval (TSV_{bin}).
    2. A linear regression model was developed for the (TSV_{bin}) on the (T_{op_bin}).
    3. T_n was estimated by solving each participant's regression model for $TSV_{bin}=0$.

The paper introduces a new variable (T_c). When a participant reported to be comfortable the corresponding operative temperature at the time of the survey was drawn out. A participant's comfort temperature (T_c) was defined as the mean operative temperature when the participant reported to be comfortable.

Following de Dear and Brager's method [12], T_p was calculated for each participant by applying the following two steps:

1. Two probit models were fitted to TPV; one for "want warmer" and one for "want cooler" within each half-degree (◦C) bins of the operative temperature (T_{op_bin}).

    2. T_p was calculated as the point of intersection between the two fitted probit curves.

d) By grouping the participants' neutral (T_n), comfortable (T_c) and preferred (T_p) temperatures, a set of neutral (Z_n), comfortable (Z_c) and preferred (Z_p) thermal zones were established for the whole sample and for each TSV cluster as established in the third step of this analysis (Section 3.4 of *Results*). Then the intersection between these three zones was calculated and defined as the thermal zone (Z_t). The Z_t of the whole sample and of each TSV cluster were established.

e) In the fifth and last step, multinomial logistic regression was applied to explore the relationship between [demographic and adaptation factors] and [the probability of membership to the four thermal sensation clusters] [19] (Section 3.5 of *Results*). In this analysis the dependent variable was the thermal sensation cluster (1 to 4) and the independent variables were separated into two categories as follows:

- **Demographic factors**: sex (Female, Male, Other, Do not specify); age (<30, 30-39, >39, unknown); change between the climate of residence and the climate of origin (Change, No change).
- **Building adaptive factors**: window opening behaviour (opened, closed or on, off).

3. Results and discussion

3.1 Exploring thermal sensation, comfort and preference variability

The first step of the analysis was to review the variability of thermal sensation, comfort and preference indices for the entire sample. As a measure of the spread over the mean, the standard deviation was used to review variability. As shown in Table 3, $TSV_{SD}=1$ on a 7-point scale, $TCV_{SD}=0.4$ on a 2-point scale and $TPV_{SD}=0.6$ on a 3-point scale. The calculated standard deviation indicates that TSV varied relatively little compared to TCV and TPV for the entire sample. To review the variability within each participant's responses, the standard deviation of TSV, TCV and TPV were estimated separately for each participant. Participants' TSV_{SD} varied between 0 and 1.93, and the mean of participants' TSV_{SD} was 0.85. This result shows a large variability between participants' TSV_{SD}. The large variability indicates that only a few participants reported the same thermal sensation throughout the study (3% of the total number of participants) while most participants' thermal sensation votes varied. With regard to thermal comfort vote, participants' TCV_{SD} varied between 0 and 0.53, and the mean of participants' TCV_{SD} was 0.27. This result shows again large variability between participants' TCV_{SD}; with several participants reporting the same level of perceived thermal comfort throughout the study (33% of the total number of participants). Finally, participants' TPV_{SD} varied between 0 and 0.97, and the mean of participants' TPV_{SD} was 0.5. The large variability in individual TPV standard deviation points out that not many participants reported the same thermal preference throughout the study (8.5% of the total number of participants). These results are further discussed in the following Section 3.2.

The second step of the analysis was to review the relationship between TSV, TCV and TPV. The assumption was that TSV may not correspond to the expected TPV (i.e. a participant may feel warm (+2) and prefer no change (0)), and that TSV may not correspond to the expected TCV (i.e. a participant may feel warm (+2) and comfortable (0)). This assumption defies Fanger's assumption, namely that within a group, 80% of participants will be comfortable when feeling 'neutral' (i.e. TSV = 0 ±1) [41]. First, the analysis reviewed the relation between thermal comfort vote (TCV) and thermal sensation vote (TSV). As not all participants reported their thermal comfort, the sample was 1,739 surveys from N=129 participants. Participants found the environmental conditions "comfortable" in 72% of the surveys. While the environmental conditions were "comfortable", surprisingly, participants reported feeling of a "bit warm" (TSV=1) or a "bit cool" (TSV=-1) (495 surveys, representing 39.6% of the 'comfortable' subset). The strength of the association between TSV and TCV was assessed using the Goodman Kruskal gamma (G) with G=-0.015 (-0.102 to -0.072), which indicates a very weak association, which is to be expected as TCV is a binary variable. Then, the analysis reviewed the relationship between thermal preference vote (TPV) and thermal sensation vote (TSV) as reported by the participants in each survey (Figure 1). The central point (0,0) is interpreted for TSV as "feeling neutral" and for TPV as "want no change". The results show a large number of neutral votes (TSV=0) (n=553) that are associated with a preference of higher temperature ((TSV,TPV)=(0,1)). Interestingly, some participants feeling "slightly warm" (TSV=1) are associated with a preference of "no change" to their thermal environment (TPV=0) (n=747). While, some participants feeling "slightly warm" (TSV=1) would prefer it to be warmer (TPV=1) (n=238). The Goodman Kruskal test reviewing the strength of the association between TSV and TPV resulted to a gamma value, G=-0.62 (-0.66 to -0.60), which indicates a strong association; as TSV increased TPV decreased. The distribution of responses supports the hypothesis of a warm bias. The majority of the participants' responses have a central sensation tendency but the preference is towards higher temperature.

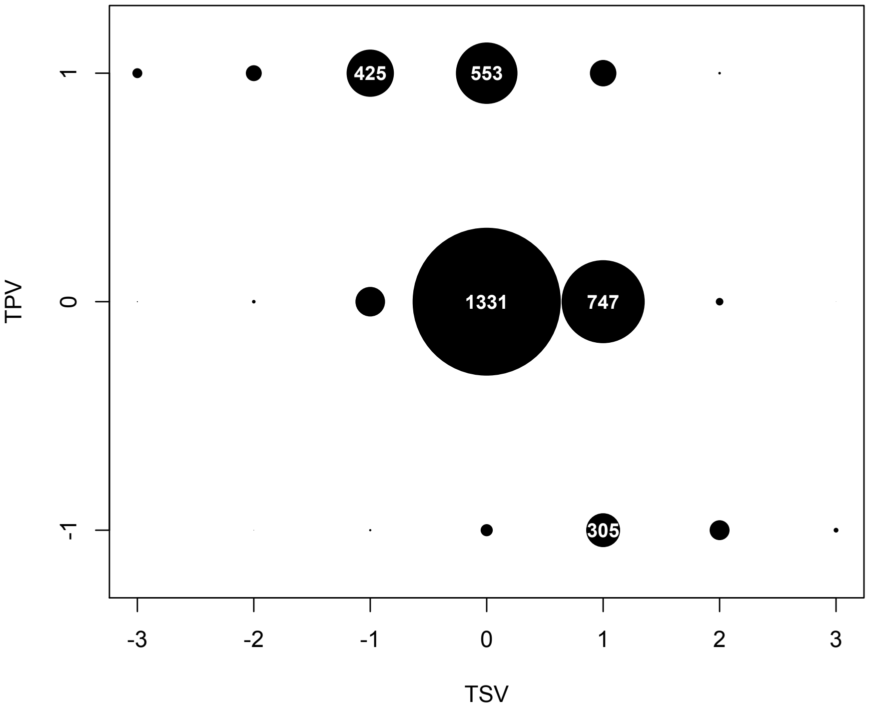

Figure 1 Comparison of agreement between thermal sensation vote (TSV) and thermal preference vote (TPV) results from the individual surveys (number of responses shown within circles and proportional to the area of circles)

To investigate further these inter-individual differences, participants' environmental conditions (T_{op}) and clothing levels (Clo) were reviewed. There was a statistically significant difference in T_{op} between participants as determined by one-way ANOVA ($F(240,4173) = 65.59$, $p<0.05$). Although a Tukey post hoc test revealed that 65% of participants' pairwise comparisons were not significant ($p>0.05$). Mean T_{op} varied between 20.2°C and 28.8°C; and T_{op} standard deviation varied between 0.1°C and 3.5°C. There was a statistically significant difference in Clo between participants as determined by one-way ANOVA ($F(254,5208) = 24.44$, $p<0.05$). Although a Tukey post hoc test revealed that 73% of participants' pairwise comparisons were not significant ($p>0.05$). Mean Clo varied between 0.3 clo and 1.4 clo, and Clo standard deviation varied between 0 clo and 0.6 clo. In summary, most participants were exposed to similar environmental conditions and had similar clothing levels. Participants inter-individual differences in TSV may be attributed to their own personal traits.

3.2 Clustering according to thermal sensation, comfort and preference.

The method introduced in this section departs from the commonly used approach, of averaging cross-sectional surveys at room or building level. From the study longitudinal research design, this analysis uses averaging of participants' surveys. This new method aims to improve the generalisation of the comfort models towards an algorithm that could be used with "smart" and adaptive buildings. Individual models have been generated from the participants' responses to the longitudinal "right-here-right-now" surveys. K-means clustering approach was used to group participants according to the mean and the standard deviation of the perceived thermal indices; TSV, TCV and TPV. For example, a participant may feel on average slightly cool and her/his thermal sensation varies by 1 unit (i.e. from cool to neutral) over the course of the study. This analysis allows to map between participants variability in TSV_{mean} and within participants variability (TSV_{SD}). Figure 2 shows the results of the K-means algorithm applied to the standard deviation of TSV and the mean of TSV. Four clusters were identified; Cluster 1 (black squares) represents participants who were feeling a bit cold (TSV_{mean}=-1.07) and their responses had a relatively large dispersion around the mean (TSV_{SD}=1.12, 8.3% of the total participants). Cluster 2 (red bullet points) contains the participants' responses with neutral thermal sensation (TSV_{mean}=-0.03) and a moderate dispersion around the mean (TSV_{SD}=0.70). This cluster contains 32.5% of the total participants, their responses had the smallest variation in TSV and several participants were feeling neutral throughout the study. Cluster 3 (green triangles) is mainly made up of participants that were feeling on average neutral (TSV_{mean}=0.06) but in reality, they were feeling equally warm and cold during the survey period, as their variability in TSV is relatively large (TSV_{SD}=1.28, 25.2% of the total participants). Lastly, Cluster 4 (blue rhombuses (diamonds)) represents participants feeling a bit warm most of the time (TSV_{mean}=0.63) with their responses showing a moderate dispersion around the mean (TSV_{SD}=0.77, 34% of the total participants). In summary, TSV Cluster 2, 3 and 4 represented similar number of participants, therefore future building systems should consider these different 'thermal traits.'

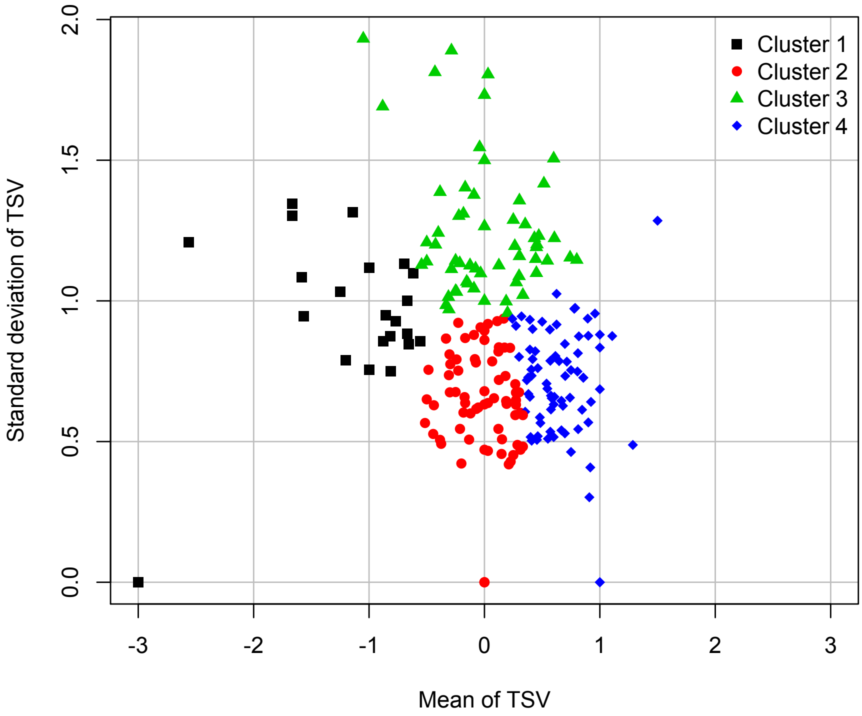

Figure 2 Scatterplot of the mean TSV and the associated standard deviation results from the individual responses. The K-means clustering resulted to 4 clusters; Cluster 1 (black squares): cold feel, high dispersion, Cluster 2 (red bullet points): neutral feel, moderate dispersion, Cluster 3 (green triangles): neutral feel, high dispersion and Cluster 4 (blue rhombuses): warm feel with low dispersion

In the case of TCV there are only three clusters. Figure 3 shows the scatterplot of the mean TCV of each participant and the corresponding standard deviation grouped by the K-means clustering method. The shape of the graph (Figure 3), as expected, follows the form of a parabola with low dispersion when the mean TCV equals to 1 (comfortable) or 2 (uncomfortable) and high dispersion at the middle (TCV_{mean}=1.5). The respondents could vote either 1 or 2 for their perceived comfort, meaning that a mean equal to 1 or 2 will have most of the responses being 1 or 2 respectively, hence low dispersion (i.e. almost all the responses are equal to the mean value). Cluster 1 (black squares) has participants who feel comfortable most of the time (i.e. low dispersion around the mean) (TCV_{mean}=1.01, TCV_{SD}=0.10, 51% of the total participants). Cluster 2 (red bullet points) contains participants who mostly were comfortable (TCV_{mean}=1.25, TCV_{SD}=0.44, 30% of the total participants). Cluster 3 represents participants who mostly were uncomfortable (TCV_{mean}=1.67, TCV_{SD}=0.47, 19% of the total participants).

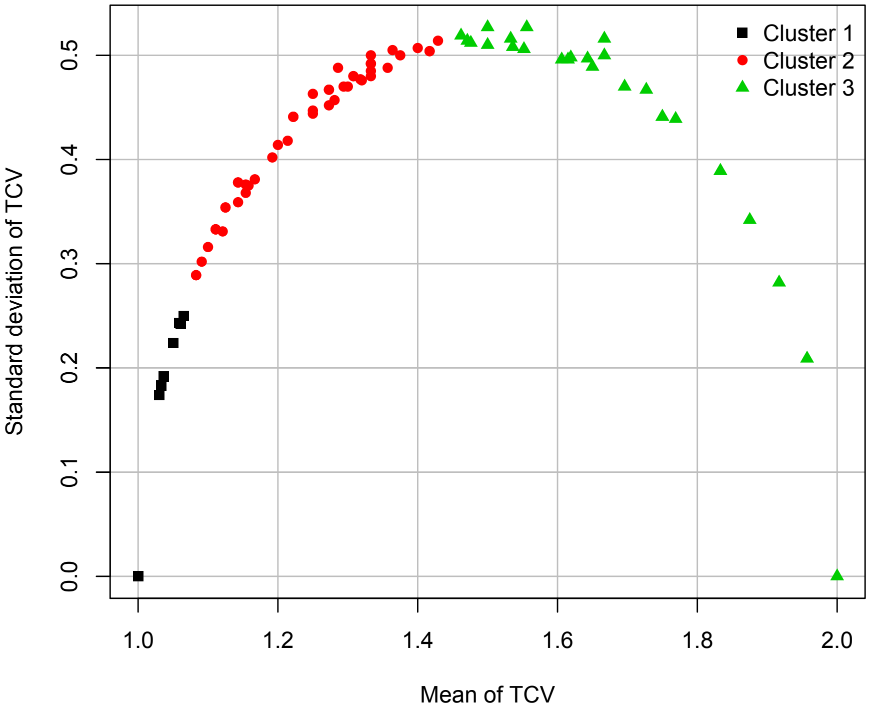

Figure 3 Scatterplot of the mean TCV and the associated standard deviation results from the individual responses. The K-means clustering resulted to 3 clusters with the lowest dispersion noted in Cluster 1 (black squares) where participants were comfortable most of the time. Cluster 2 (red bullet points): mostly comfortable, Cluster 3 (green triangles): mostly uncomfortable

Finally, Figure 4 summarises the results from the K-means algorithm for TPV clustering. As in the TSV clustering results, there are four groups created according to the standard deviation and the mean of TPV. The interpretation of the clusters is based on the likelihood of a participant expressing the same temperature preference during the study. Cluster 1 (black squares) represents participants that in general wanted cooler conditions (TPV_{mean}=-0.42, TPV_{SD}=0.57, 16% of the total participants). Interestingly, Cluster 2 (red bullet points) shows that there is a group of participants who most of the time did not want any change to the environmental conditions (TPV_{mean}=-0.01, TPV_{SD}=0.32, 23.6% of the total participants). Cluster 3 (green triangles) contains the participants that most likely had a large variation from "want colder" to "want warmer" in their responses (TPV_{mean}=0.15, TPV_{SD}=0.69, 35.4% of the total participants). Cluster 4 (blue rhombuses) contains the participants that most of the time wanted warmer conditions (TPV_{mean}=0.56, TPV_{SD}=0.51, 25% of the total participants).

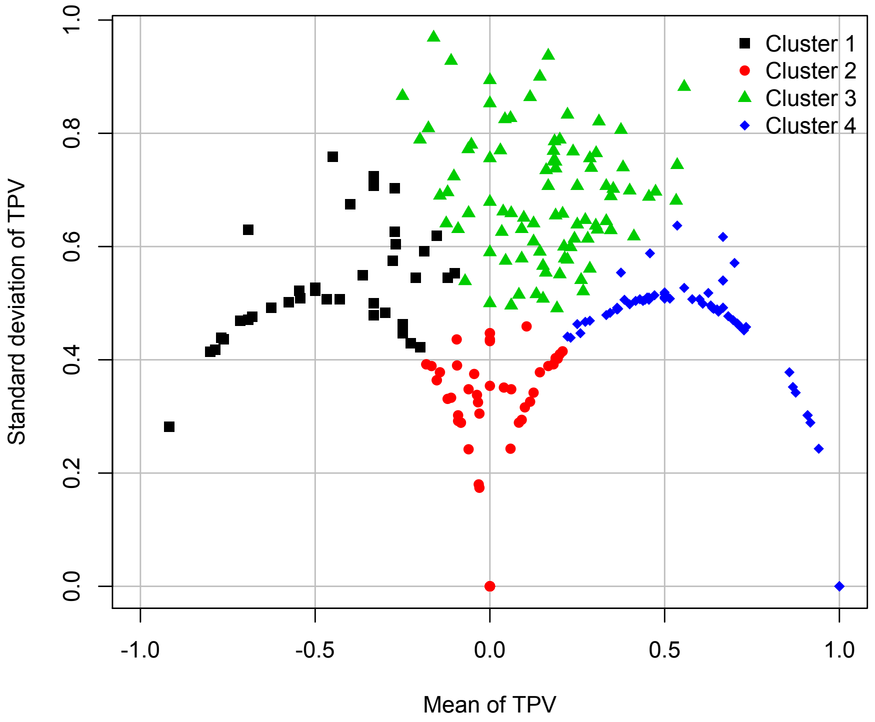

Figure 4 Scatterplot of the mean TPV and the associated standard deviation results from the individual responses. The K-means clustering resulted to 4 clusters; Cluster 1 (black squares): want colder, high dispersion, Cluster 2 (red bullet points): want no change, small dispersion, Cluster 3 (green triangles): want no change, high dispersion and Cluster 4 (blue rhombuses): want warmer, high dispersion

Figures 5 and 6 show the comparison between TSV, TCV and TPV clusters. Participants with neutral thermal sensation (TSV=0) and small dispersion around the mean (TSV cluster 2) are likely to be associated with TCV cluster 1, comfortable most of the time (i.e. low dispersion around the mean) as expected (see Figure 5). There is also a high association between TSV cluster 3 (neutral mean sensation with high variability) and TCV cluster 2 (mostly comfortable). This result shows that the personal sensation may vary despite the person's perception of being "comfortable". Thermal adaptation mechanisms (behavioural, physiological and psychological) are in continuous interaction and development [33].

The strength of the association between TSV clusters and TCV clusters was assessed using Cramer's V and it was found V= 0.3475 with k=3, which indicates a medium association [36].

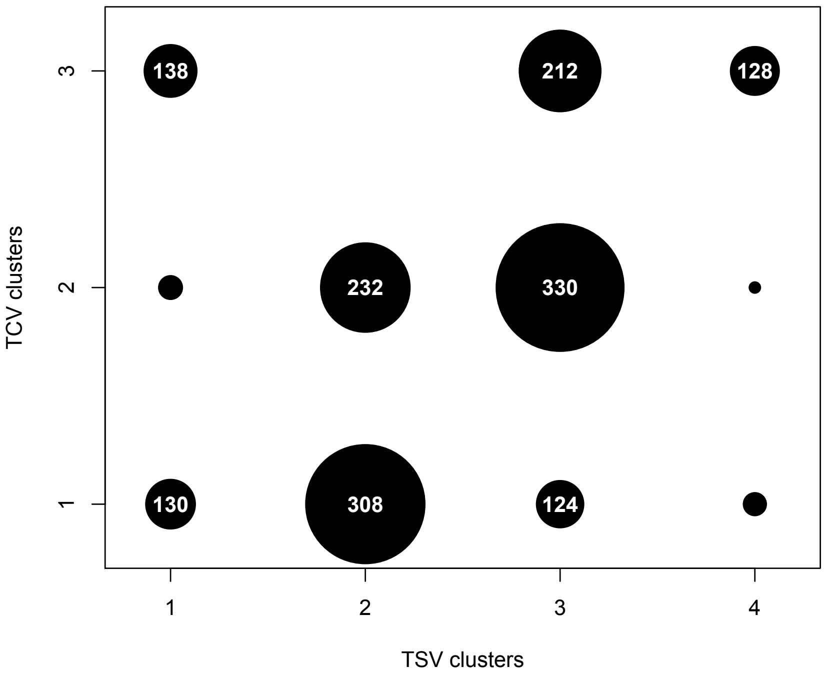

Figure 5 Comparison of agreement between the TSV and TCV clusters to reveal internal associations. (Number of responses shown within circles and proportional to the area of circles)

Regarding the association of the TPV clusters with the TSV clusters, Cramér's V is equal to V= 0.3064 with k=4 which again is representative of a medium association [36]. Interestingly, participants within the TSV cluster 2 (neutral thermal sensation, small variability) are equally associated with participants who wanted no change in the conditions (TPV cluster 2), participants with large variation in responses of warmer/cooler preference (TPV cluster 3) and participants that wanted warmer conditions (TPV cluster 4) (Figure 6).

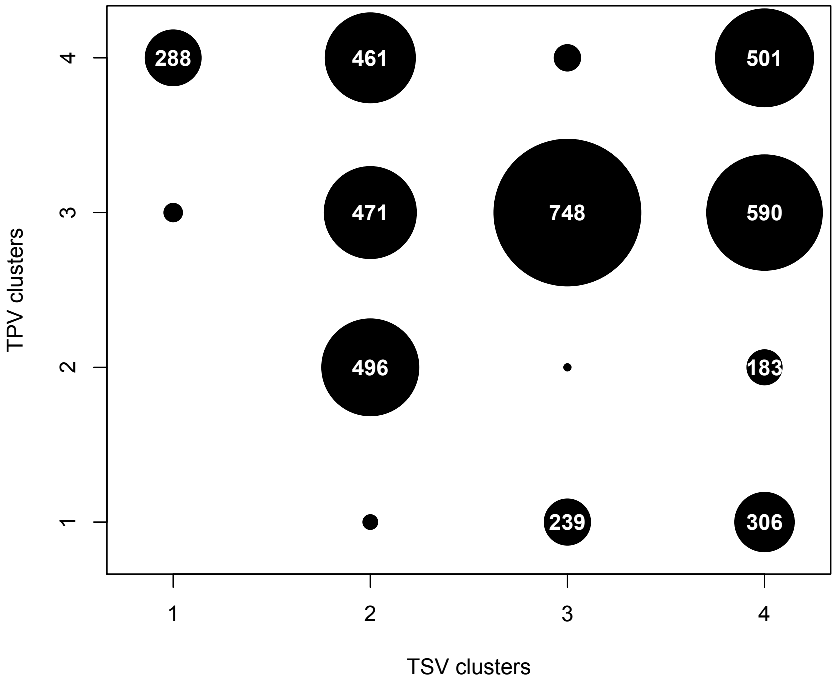

Figure 6 Comparison of agreement between the TSV and TPV clusters to reveal internal associations. (Number of responses shown within circles and proportional to the area of circles)

3.3 Estimating the individual neutral, comfortable and preferred temperatures

3.3.1 Estimation of neutral temperature (Tn) and the neutral zone, Zn.

A linear regression analysis of TSV_{bin} on T_{op_bin} was undertaken for the entire sample (R^2=0.79, $p<0.05$). As shown in Figure 7, the data fit closely to the regression line for T_{op_bin} values ranging from 21 °C to 27 °C. This range corresponds to TSV_{bin} values close to the neutral, central point.

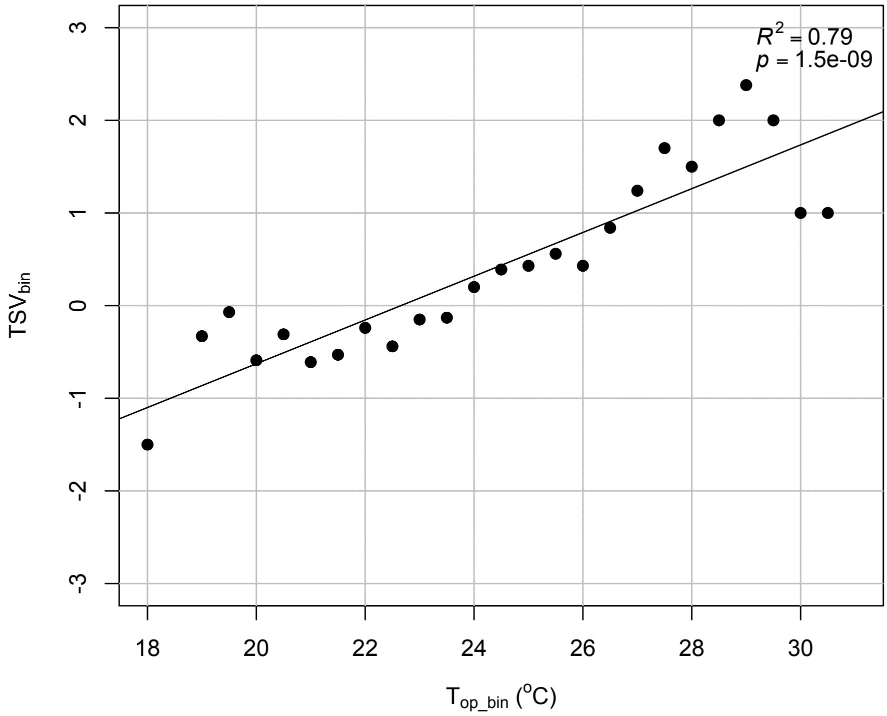

Figure 7 Linear regression analysis between T_{op_bin} and TSV_{bin} for the entire sample in the study (n=5,576)

Subsequently, linear regression models of the TSV_{bin} on T_{op_bin} were developed for each participant. Table 4 summarises the results from the regression analysis for the entire sample and each cluster separately. Only 21.7% of the individual regression models achieved statistical significance at α=0.05. This could be an indication of a large number of participants having experienced small indoor temperature ranges. The results between the TSV clusters show little percentage difference in the number of regression models that are statistically significant. The review of the models' slope suggests that TSV Cluster 3 participants ("neutral & high variation") are twice as sensitive to changes of the Top than participants from the other clusters. Kruskal Wallis ANOVA analysis showed that the T_n varies significantly amongst the four clusters ($\chi^2(3)$= 637.22, $p < 0.05$). A post-hoc test looking at multiple comparison between the groups revealed that there are significant differences between all clusters. The Z_n for the whole sample was defined from the cross-section of the TSV cluster regression results for each cluster (Figure 8). Figure 8 shows the neutral temperature range (T_n) calculated from the individual regression models for each TSV cluster. The final T_n range or neutral zone was estimated to be from 18.3 °C to 26.6 °C.

Table 4 Summary of the regression results of TSV_{bin} on T_{op_bin} used for the estimation of T_n and Z_n.

	Total sample	**TSV Cluster 1**	**TSV Cluster 2**	**TSV Cluster 3**	**TSV Cluster 4**
Number of participants	**258**	**22**	**71**	**58**	**74**
Number of participants with Regression Models Achieving 95% Significance	**56** (21.7% of total)	**5** (22.7% of total)	**16** (22.5% of total)	**17** (29.3% of total)	**18** (24.3% of total)
Mean model **Intercept**	**-11.04** (± 7.36)	**-4.41** (± 8.43)	**-4.58** (± 8.61)	**-9.46** (± 11.11)	**-5.53** (± 10.88)
Mean model **Slope**	**0.47** (± 0.32)	**0.14** (± 0.37)	**0.19** (± 0.35)	**0.41** (± 0.48)	**0.26** (± 0.45)
Mean T_n (°C)	**23.15** (± 1.4)	**24.9** (± 0.9)	**23.9** (± 1.1)	**23.3** (± 0.8)	**22.2** (± 1.1)
Z_n as Minimum, Maximum and **Range** T_n (°C)	18.3-26.6 **8.3**	23.8-26.0 **2.2**	21.8-26.6 **4.8**	22.4-24.6 **2.2**	18.3-23.2 **4.9**

Note: uncertainty is defined as ± one standard deviation.

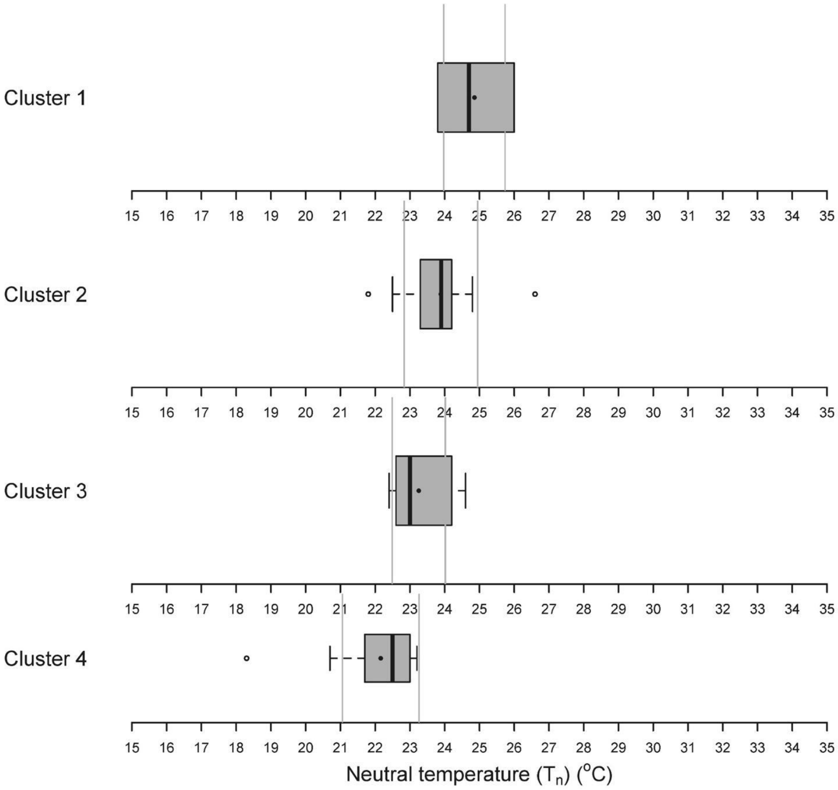

Figure 8 Neutral temperature (T_n) ranges for the four TSV clusters. The neutral zone (Z_n) for the entire sample was defined by the cross-section of the four T_n ranges (Z_n={18.3 °C to 26.6 °C})

3.3.2 Estimation of comfort temperature (Tc) and the comfort zone (Zc).

Responses from 128 participants were used in this analysis as TCV was not surveyed in Australia and South Korea. The TSV cluster analysis used data from 95 participants out of the total 128. The difference is explained by the survey from India results which excluded TSV from the reported variables but included TCV. Here, the Kruskal Wallis ANOVA analysis showed that T_c differs significantly amongst the four TSV clusters ($\chi^2(3)=50.331$, $p < 0.05$). However, the post-hoc test looking at multiple comparisons between the groups revealed that there were no significant differences between Clusters 1 and 3 and between Clusters 2 and 3. Table 5 shows the results from the regression analysis for the entire sample and each cluster separately. Following the procedure introduced in the *Analysis methods (Section 2.6)* the Z_c for the whole sample was estimated to be from 20.1 °C to 28.8 °C (Figure 9).

Table 5 Summary of the regression results of TCV_{bin} on T_{op_bin} used for the estimation of T_c and Z_c.

	Total sample	**TSV Cluster 1**	**TSV Cluster 2**	**TSV Cluster 3**	**TSV Cluster 4**
Number of participants	**258**	**22**	**71**	**58**	**74**
Number of participants for who TCV data were available from the surveys.	**128** (50% of total)	**18** (81.8% of total)	**27** (38% of total)	**38** (65.5% of total)	**12** (16.2% of total)
Mean T_c (°C)	**24.8** (± 2.6)	**23.1** (± 0.9)	**22.8** (± 1.0)	**23.0** (± 0.9)	**23.2** (± 1.4)
Z_c as Minimum, Maximum and **Range** T_c (°C)	20.1-28.8 **8.7**	21.3-24.3 **2.9**	20.5-24.9 **4.5**	20.8-24.6 **3.8**	20.1-25.0 **4.9**

Note: uncertainty is defined as ± one standard deviation.

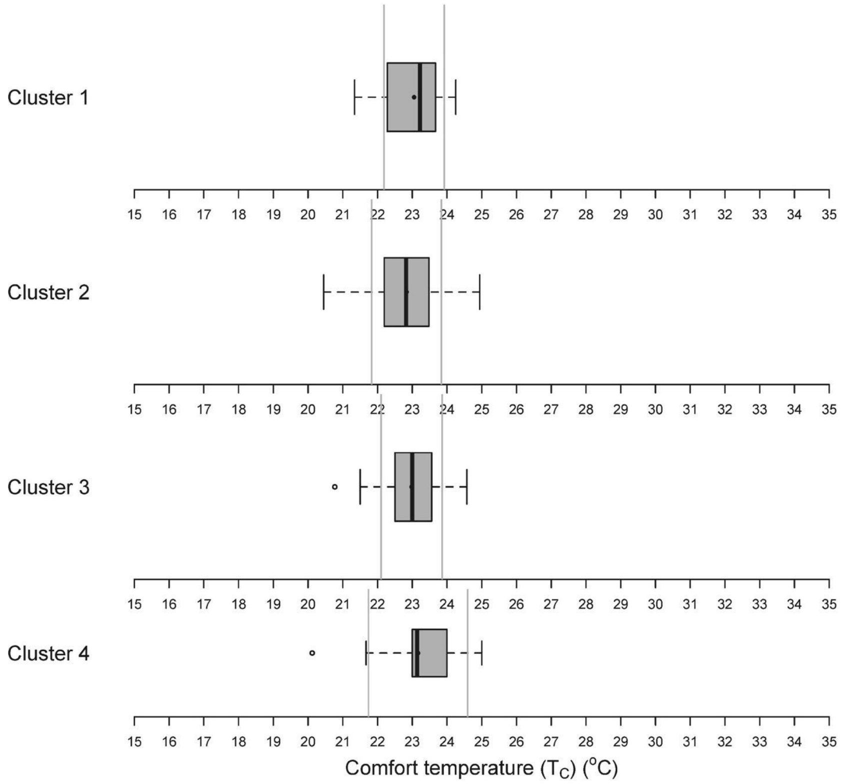

Figure 9 Comfort temperature (T_c) ranges for the four TSV clusters. The comfort zone (Z_c) for the entire sample was defined by the cross-section of the four T_c ranges (Z_c={20.1 °C, 28.8 °C})

3.3.3 Estimation of preferred temperature (Tp) and the preferred zone (Zp).

Following the procedure described in the Analysis methods (Section 2.6), only 5.4% of the sample achieved statistical significance at α=0.05. There were only 14 participants in total with probit models being at the 95% statistical significance levels (Table 6). In all TSV clusters, the total of the participants with probit models achieving 95% significance was 13. As in the previous results, there is one participant less in the TSV clusters than the entire sample, as the survey from India excluded TSV but included TPV as a variable. Since the sample sizes were very small, it was decided to use the temperature range of participants with probit models achieving 95% significance as the boundary conditions for the selection of T_p for each participant. In particular, if a participant's T_p was within the entire sample's temperature range from 21.5 °C to 34.9 °C (Table 6), then this participant was included in the analysis. A Kruskal Wallis ANOVA analysis showed that T_p in the four clusters differs significantly ($\chi^2(3)=191.8$, $p < 0.05$). Despite the ANOVA results, a post-hoc test looking at multiple comparison between groups revealed that there is no significant difference between Clusters 1 and 4. The final Z_p range for the whole sample was estimated to be from 21.5 °C to 34.9 °C (Figure 10).

Table 6 Summary of the regression results of TPVbin on Top_bin used for the estimation of Tp and Zp

	Whole sample	**TSV Cluster 1**	**TSV Cluster 2**	**TSV Cluster 3**	**TSV Cluster 4**
Number of participants	**258**	**22**	**71**	**58**	**74**
Number of participants with Probit models Achieving 95% Significance	**14** (5.4% of total)	**0**	**3** (4.2% of total)	**3** (5.2% of total)	**7** (9.5% of total)
Number of participants within the same range as the temperature range of participants with probit models achieving 95% significance	**145** (56.2% of total)	**11** (50% of total)	**41** (57.7% of total)	**38** (65.5% of total)	**40** (54% of total)
Mean T_p (°C)	**25.4** (± 2.8)	**26.1** (± 3.22)	**24.7** (± 2.1)	**24.4** (± 2.7)	**25.6** (± 2.8)
Z_p as Minimum, Maximum and **Range** T_p (°C)	21.5-34.9 **13.4**	22.6-34.3 **11.7**	21.6-31.5 **9.9**	21.7-34.2 **12.5**	21.5-34.9 **13.4**

Note: uncertainty is defined as ± one standard deviation.

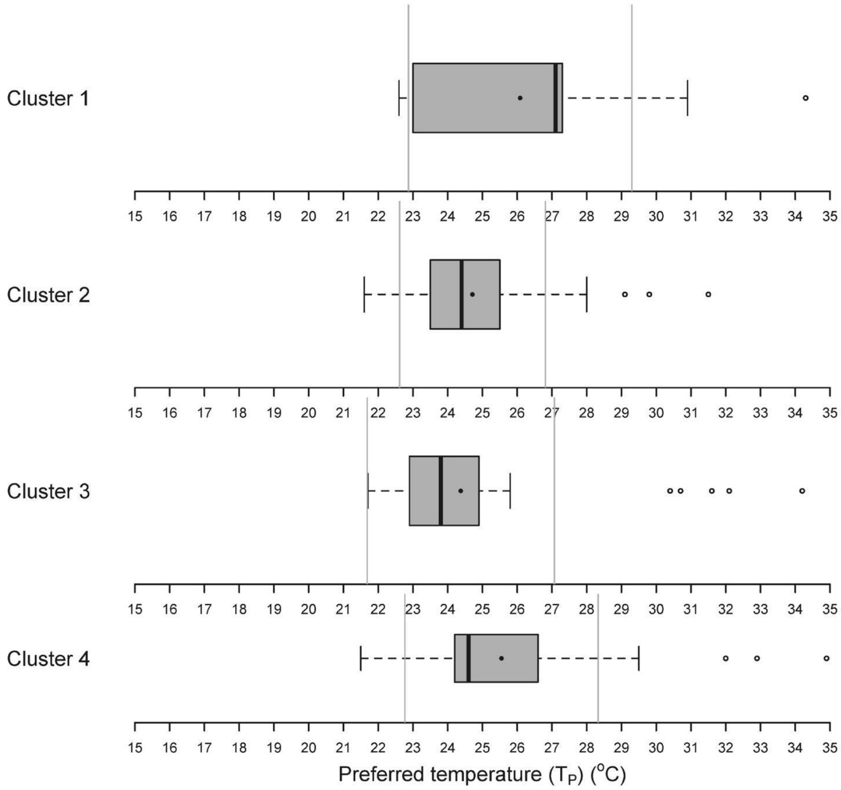

Figure 10 Preferred temperature (T_p) ranges for the four TSV cluster. The comfort zone (Z_p) for the entire sample was defined by the cross-section of the four T_p ranges (Z_p={21.5 °C to 34.9 °C})

3.4 Definition of the person-centric thermal zone Zt.

The person-centric thermal zone, Z_t, is defined by the intersection of the neutral (Z_n), comfortable (Z_c) and preferred (Z_p) thermal zones.

$$Z_t = Z_n \cap Z_c \cap Z_p$$

The thermal zones' characteristics used for the definition of the Z_t is shown in Table 7. The minimum and maximum temperature and the temperature range are presented for the entire sample and for the four TSV clusters. The preferred zone (Z_p) temperature has quite high maximum temperature values and the largest range consequently. The comfort (Z_c) and preferred (Z_p) minimum and maximum temperatures are higher than the respective neutral temperatures. The final Z_t range for the total sample was estimated to be from 21.5 °C to 26.6 °C (Table 7).

Table 7 Overview of the thermal zones' characteristics used for the definition of Zt

	Total sample	**TSV Cluster 1**	**TSV Cluster 2**	**TSV Cluster 3**	**TSV Cluster 4**
Z_n (Minimum - Maximum) and **Range** of T_n (°C)	(18.3-26.6) **8.3**	(23.8-26.0) **2.2**	(21.8-26.6) **4.8**	(22.4-24.6) **2.2**	(18.3-23.2) **4.9**
Z_c (Minimum - Maximum) and **Range** of T_c (°C)	(20.1-28.8) **8.7**	(21.3-24.3) **2.9**	(20.5-24.9) **4.5**	(20.8-24.6) **3.8**	(20.1-25.0) **4.9**
Z_p (Minimum - Maximum) and **Range** of T_p (°C)	(21.5-34.9) **13.4**	(22.6-34.3) **11.7**	(21.6-31.5) **9.9**	(21.7-34.2) **12.5**	(21.5-34.9) **13.4**
Z_t (Minimum - Maximum) and **Range** (°C)	(21.5-26.6) **5.1**	(23.8-24.3) **0.5**	(21.8-24.9) **3.1**	(22.4-24.6) **2.2**	(21.5-23.2) **1.7**

3.5 Probability of membership to the thermal sensation clusters

At the last step of the analysis, sex (Figure 11), age (Figure 12), change to climate of origin (Figure 13) and window opening behaviour (Figure 14) have been assessed for their effect on the probability of membership of an individual to a particular TSV cluster. These variables were selected because they have been identified as important by previous studies [38]. Table 8 shows the number of participants for three variables' categories, as well as T_{op} and Clo. Preliminary analysis reviews inter-TSV clusters differences. TSV cluster 1 has the lowest mean T_{op} but highest mean Clo value; while TSV cluster 4 has the highest mean T_{op} but lowest mean Clo value. Although the variations in mean T_{op} and mean Clo value between TSV clusters are small, the results may be interpreted as behavioural adaptation through clothing. While there are statistical differences between TSV clusters for T_{op} ($F(3,3532)=53.31$, $p<0.05$) and Clo ($F(3,4527)=71.71$, $p<0.05$), the range in mean T_{op} is only 1 °C and the range in mean Clo is only 0.2 clo with similar standard deviations for all clusters. In summary the four clusters have similar environmental conditions and clothing levels, yet the ranges in Z_t for each cluster vary (as seen above), which may be attributed to participants' own personal traits. Factors contributing to personal traits may be contextual (climate, building design and associated controls, economics, etc.), social (i.e. culture, organisation, etc.), physiological (i.e. age, sex, etc.) or psychological (i.e. habit, expectation, perception, etc.) [33] [38]. The following analysis reviews four factors (sex, age, change from country of origin and window opening behaviour) in an effort to unravel these personal traits.

Table 8 Summary of TSV cluster's variables

TSV clusters	**No. of participants**	**T_{op} (°C)**	**Clo (clo)**	**Sex**		**Age**			**Change from original climate**	
				F	**M**	**<30**	**30-39**	**>39**	**Change**	**No change**
Cluster 1	22	22.9 ±1.2	0.9 ±0.2	19	1	4	5	11	16	0
Cluster 2	71	23.5 ±1.5	0.7 ±0.3	36	26	12	21	18	46	1
Cluster 3	58	23.3 ±1.5	0.8 ±0.3	25	16	14	8	16	34	1
Cluster 4	74	23.9 ±1.5	0.7 ±0.3	26	44	5	23	19	64	0

Note 1: T_{op} and Clo are summarised by the mean ± one standard variation

Note 2: The sum of the number of participants in the variables 'Sex', 'Age' and 'Change from original climate' differ from the number of participants per cluster, as some participants did not answer the question.

Figure 11 shows that there are significant differences between males and females, especially for TSV clusters 1 and 4. The probability to belong to TSV cluster 1 (feeling colder and high variability) for male remains low and TSV for male is unchanged regardless of the increase in operative temperature; this is likely to be due to the small sample size, N=1. The probability to belong to TSV cluster 1 for female decreases with the increase of temperature. On the other hand, for TSV cluster 4 (feeling warmer and low variability), the probability has a positive correlation with the operational temperature for both males and females. Males have systematically higher probabilities across the T_{op} range. Further analysis applied chi-square tests to the numbers of male and female in pairs of clusters. Results showed that sex was significantly different between all four TSV clusters ($p<0.05$), with one exception; there was no significant difference between TSV clusters 2 and 3 ($\chi^2(1)=0.008$, $p=0.93$). As participants felt on average neutral, there was no difference in sex if their thermal sensations vote varied or not during the months of monitoring. Besides, the odds of a participant being female were 32.15 times higher if they felt on average slightly cold (TSV cluster 1) than if they felt on average slightly warm (TSV cluster 4).

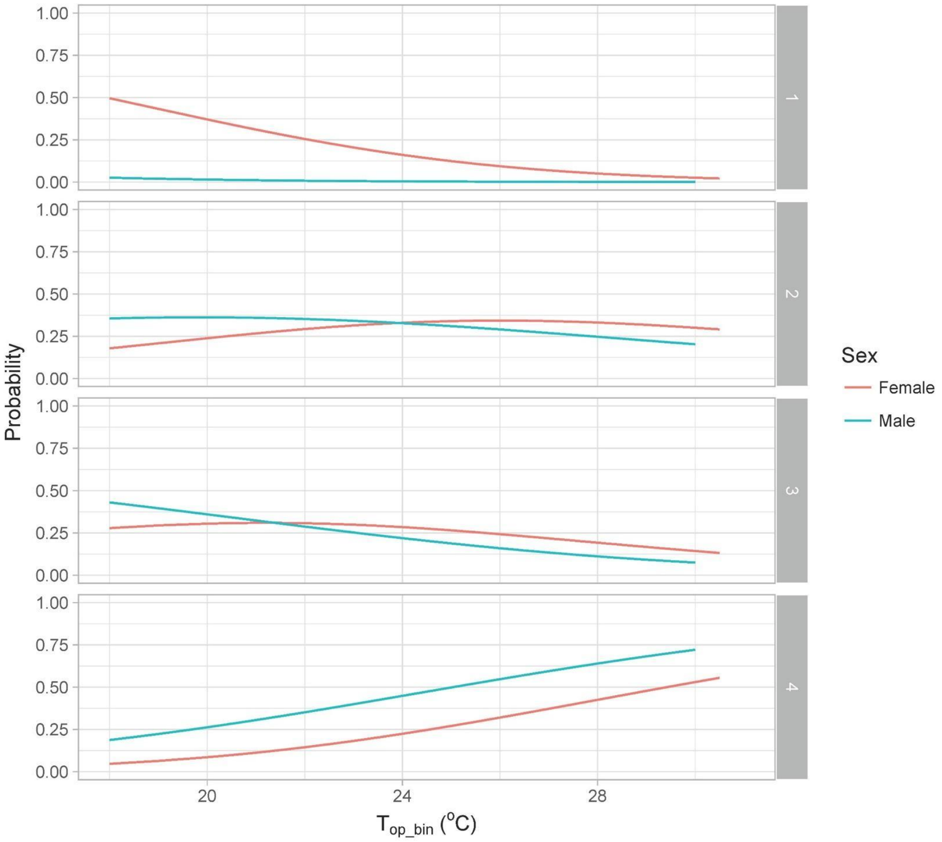

Figure 11 Probability of cluster membership in relation to sex of participants for the TSV clusters (number of clusters shown on right axis)

Regarding the relation between age and the probability of membership to TSV clusters (Figure 12), the probability of all age categories (<30, 30-39, >39 years old) decreases with the increase of T_{op} for the TSV clusters 1,2 and 3. For the TSV cluster 4 (feeling warmer and low variability), the trend reverses with the probability increasing in high T_{op} values. Between the age categories, the participants in the 30-39 group have lowest probability than the other to be in TSV cluster 3 (feeling neutral and high variability), whereas they have the highest probability of membership to TSV cluster 4 (feeling warmer and low variability). Further analysis applied chi-square or Fisher's exact tests to the three age groups in pairs of clusters. Results showed that there was no significant difference in age groups between all four TSV clusters (p>0.05), with one exception; there was a significant difference between TSV clusters 3 and 4 ($\chi^2(2)=12.269$, p=0.002). The odds of a participant being <30 years old, compared to 30-39 years old, were 8.05 times higher if they felt on average neutral and their thermal sensation varied (TSV cluster 3) than if they felt on average slightly warm (TSV cluster 4).

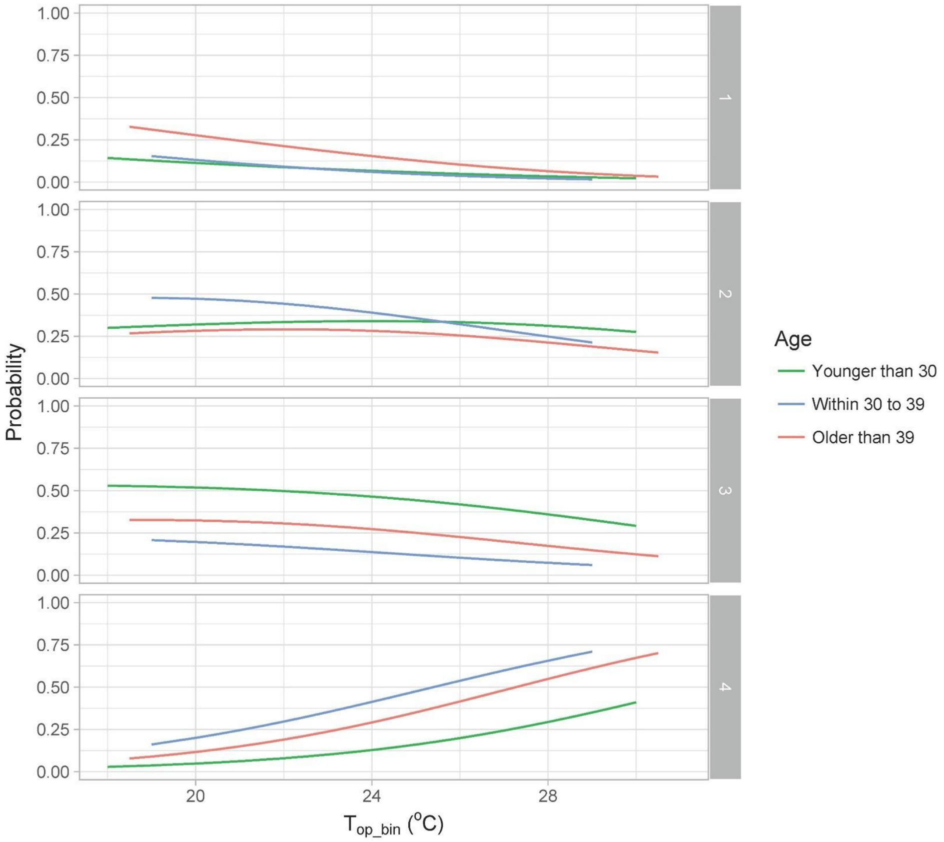

Figure 12 Probability of cluster membership in relation to age of participants for the TSV clusters (number of clusters shown on the right axis)

Regarding the relation between change from climate of origin and the probability of membership to TSV clusters (Figure 13), only two participants did not change climate. The analysis focuses on the trend of the participants that change location from their climate of origin. As expected, there is a higher probability of participants feeling on average slightly cold in TSV clusters 1 and a higher probability of participants feeling on average slightly warm in TSV clusters 4. Surprisingly, participants in TSV cluster 2 felt on average neutral with low variability throughout the range of T_{op} (18 to 31°C). While participants in TSV cluster 3 felt on average neutral with moderate variability on the cooler range of T_{op} scale. Future analysis may review the difference between the climate of origin of the participants and the climate they now reside (i.e. moved from a “hot” climate to a “temperate” climate).

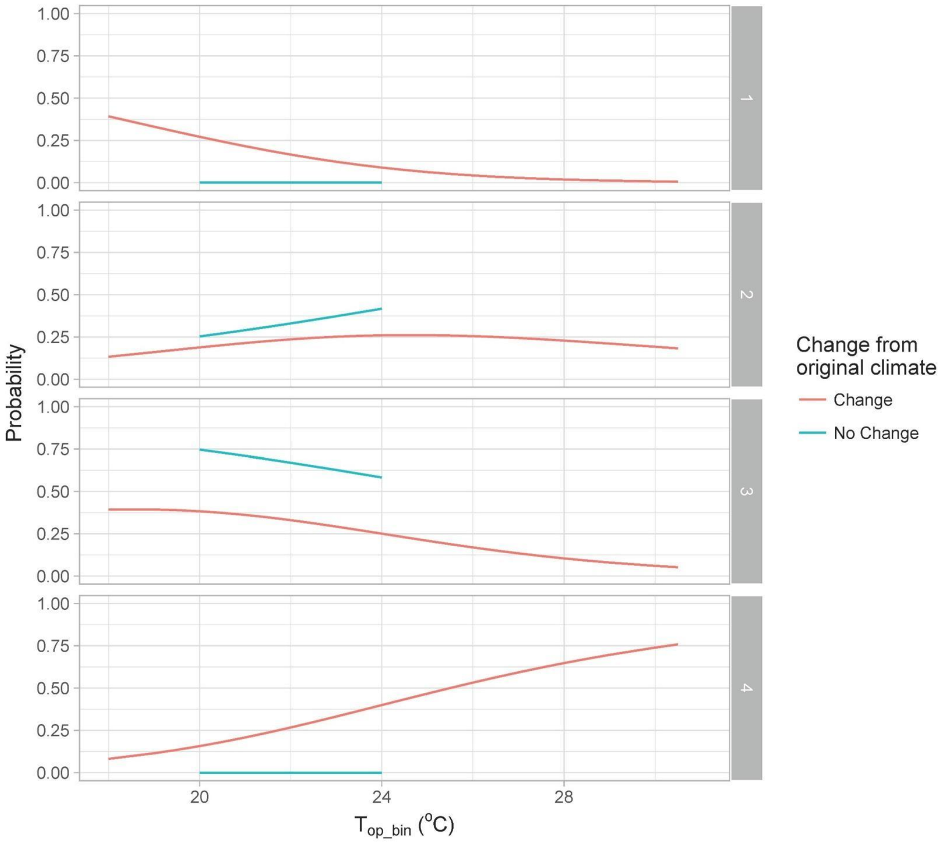

Figure 13 Probability of cluster membership in relation to any changes from the climate of origin to the climate of residence (climate of residence different than origin) of participants for the TSV clusters (number of the cluster shown on the right axis)

Finally, the probability of participants who opened the window to be in TSV cluster 1 (feeling colder and high variability) is close to 0 (Figure 14). At the same time the probability of participants who open the window to be in TSV cluster 4 (feeling warmer and low variability) increases as T_{op} increases. The probability for both "open" and "closed" windows in TSV clusters 1,2,3 is negatively correlated with the T_{op}. In general, the probabilities between those who opened the windows and those who closed the windows follow the same trend and have similar values across the clusters (Figure 14). For this dataset, the window opening behaviour does not seem to affect the cluster membership of an individual.

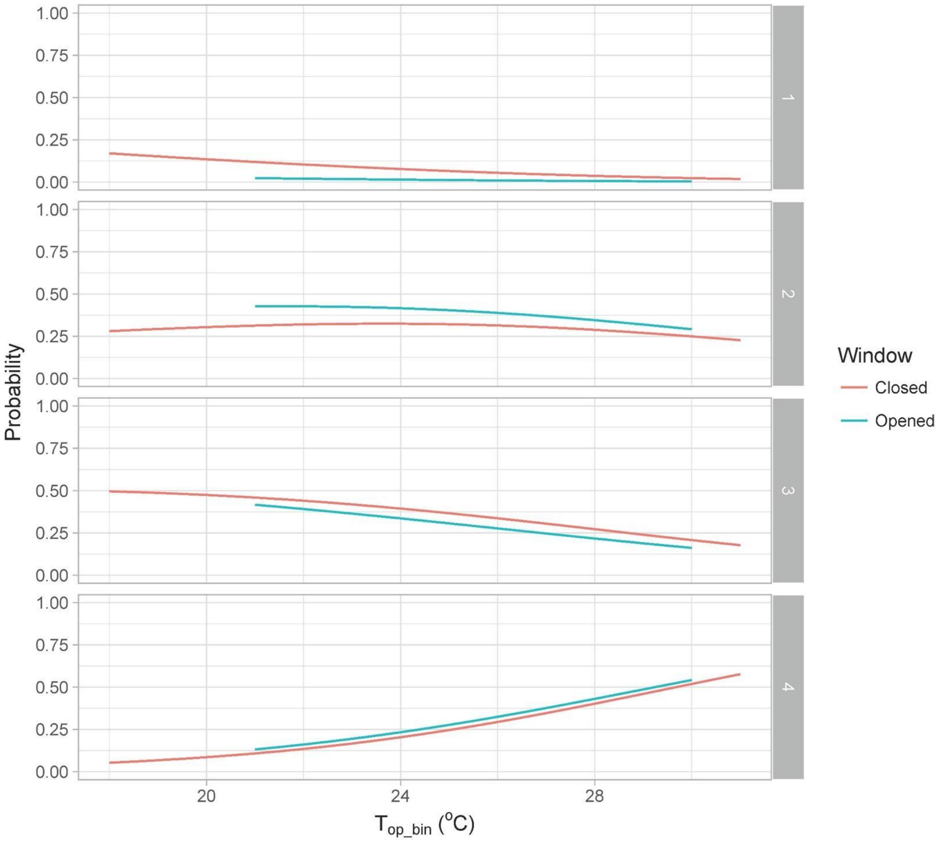

Figure 14 Probability of cluster membership for the window opening behaviour of participants for the TSV clusters (number of the cluster shown on right axis)

4. Conclusions

4.1 Key findings and new insights

This study is the first of its kind in the discipline of thermal comfort in adaptive buildings because it reports and analyses year-long observations from longitudinal surveys and concurrent environmental measurements in six countries around the world. Departing from a spatial averaging analysis of thermal comfort indices, this analysis reviews the individual thermal perception variations over time. Building on the longitudinal survey, the analysis applies a clustering approach to create groups of equivalent thermal sensation. Four thermal sensation groups of participants were identified: (TSV cluster 1) 8% of participants felt on average slightly cold and their votes varied moderately, (TSV cluster 2) 32% of participants felt on average neutral and their votes did not varied much, (TSV cluster 3) 25% of participants felt on average neutral and their votes varied moderately, and finally (TSV cluster 4) 34% of participants felt on average slightly warm and their votes did not varied much. These results reveal thermal comfort traits in an effort to contribute to the knowledge in the field of adaptive comfort mechanisms and their application towards adaptive buildings. Each one of these four thermal comfort traits may have different adaptive responses, in particular behavioural adjustments. These may be adjustments of/around the participants (e.g. clothing level, activity level, posture, food/drinks intake, etc.) [37], or adjustments of indoor environmental controls (e.g. window opening, shading, blinds, HVAC system, etc.). As reviewed in section 3.5, the four TSV clusters have similar mean clothing levels and window opening behaviour. Future research may explore differences in other behavioural adjustments and perceived controls. Besides, the analysis explored differences in psychological factors (sex and age) between TSV clusters. Interestingly, there were significant differences in sex; however there were little differences in age between TSV clusters. The odds of a participant being female were 32.15 times higher if they felt on average slightly cold (TSV cluster 1) than if they felt on average slightly warm (TSV cluster 4).

In addition, this paper introduced person-centric thermal zone, Zt, which incorporates the individual comfort characteristics into one metric that could improve the representativeness of personal comfort attributes and inform “adaptive” building design and HVAC operation strategies in mixed-mode non-residential buildings.

This analysis points out the unrepresentativeness of the scales in thermal comfort surveys. Results show that participants reported “Yes it feels warm” only to add “this is nice”. Such responses call into question the assumptions of the standard comfort indices that assumes thermal sensation neutrality as being comfortable. The results of this study show a “warm” bias.

The variation around the central point of the thermal indices (TSV, TCV and TPV) and the range of the calculated measures reinforce the argument about the existence of complex physio-psychological relationships that need to be further researched. This result may also be due to the scales and the discrepancy between the indices is likely to affect high temperatures and sensation of feeling warm without the same result being equally prominent to the cold side of the scales (e.g. none of the participants reported “feels a bit cold” and at the same time “want it cooler”). This implies that the neutral and preferred temperatures have a warm bias either because of physio-psychological factors or the particular sample’s characteristics. These results support the findings from previous studies [38] regarding diversity of thermal perception and the interrelations between psycho-contextual factors.

If we consider that adaptive buildings are buildings that adapt to the comfort requirements of the occupants –which themselves are based on adaptive mechanisms- then it becomes apparent that a successful adaptive building is one that facilitates the adaptive processes of the occupants and provides an environment where each individual can achieve comfort without this being restricted in a single temperature set point.

4.2 Internal and external validity

The internal validity of the study refers mainly to the research design. In the data collection, there were differences in the equipment used from each country. Differences in the calibration levels may introduce errors in measurement. Further bias might have been introduced by the positioning of the sensors within each building. Another source of bias might have been the selected method of administering the "right-here-right-now" surveys. Depending on the country, participants answered the survey using smartphone applications, online questionnaires and/or paper surveys. The difference in surveying methods may have resulted in a "respondent" bias. In the data analysis, the number of participants between country varied from 24 (Australia, Canada) to 105 (South Korea). Jordan had 27 participants, India had 33 participants and the UK had 45 participants. This variation in the number of respondents is likely to introduce perception, expectation and climatic bias according to the characteristics of the country with the most respondents, here South Korea. Finally, the duration of the surveys was from 8 to 21 months. This variation in duration might introduce a "seasonal effect" on the results and affect the robustness of the comfort indices due to the distribution of the data across seasons and the number of surveys used for each analysis respectively. Internal validity also refers to the analysis methods. The paper applies established methods to estimate T_n and T_p [12]. These methods bin temperature by half-degree (°C). Through this process information is lost, also changing the size of the bin may change the result. By binning the data, an ordinal variable (TSV) become a continuous variable, so linear regression may be applied. If the data was not binned, then ordinal regression should be applied.

The external validity of the study refers mainly to the generalisation and transferability of the results. The case study buildings were all mixed mode (concurrent or change-over mode of operation), therefore the findings may not apply to other types of buildings. In addition, the buildings were all office building and the results may not be transferable to other building uses that would have different adaptive opportunities. Nevertheless, the research design developed in this study may be applied in future research to further investigate individual comfort models in the context of adaptive buildings.

4.3 Future research

This study has identified the need for further research on the scales used for collecting subjective thermal comfort responses from longitudinal surveys. It has also been discussed that the driving factors of personal comfort are a complex system of physio-psychological interactions and mechanisms [39] [40]. The results support the findings from previous studies regarding the diversity in thermal perception between individuals [38], identifying a knowledge gap in the research on the interactions between the factors of individual thermal comfort and their relation to occupants' adaptive behaviour. Future research may explore how different personal traits and associated factors may affect adaptive behaviours in buildings.

In the context of adaptive buildings, it has been shown that the perceived control affects the thermal comfort of individuals [24] [39]. This study proposes that further investigation is required into the changes of individual adaptive mechanisms in relation to the occupancy background levels (e.g. different thermal sensation and preference when alone or in a group under the same environmental conditions and access to building controls). The results could reveal the existence of critical thresholds of occupancy for open plan, shared, work spaces which when exceeded could have adverse effects to the comfort and wellbeing of occupants.

Low energy, adaptive buildings with personalised controls are likely require the online training of adaptive algorithms. Future research needs to focus on the development and performance evaluation of adaptive, reactive and dynamic algorithms and systems that will continuously evaluate the zonal conditions and occupancy, and proactively adapt the environmental conditions according to historical data, automated occupancy detection, and the preferences of the people currently present in each zone.

Author contributions

S.G., L.B., F.A, C.B., S.K., R.M., H.P., R.R., F.T and R.U. contributed to field work of the case study buildings;

S.G., L.B., F.A, C.B., C.C., R.T.H., J.K., S.K., R.M., H.P., R.R., F.T and R.U. contributed to the quality assurance and processing of the case study raw data;

All authors contributed to the conception of the study design;

S.G. and L.B. contributed to the analysis and interpretation of the data;

S.G. and L.B. contributed to drafting the article and revising it critically;

S.G., L.B., R.T.H. and FT contributed to the approval of the final version.

Acknowledgements

This study was conducted within the framework of the International Energy Agency – Energy in Buildings and Communities Program (IEA-EBC) Annex 69 - Strategy and Practice of Adaptive Thermal Comfort in Low Energy Buildings.

S.G. and L.B. would like to thank the Sustainable Energy Research Group at the University of Southampton for supporting this work (www.energy.soton.ac.uk).

The work of F.A., R.T.H. and A.W. was supported by the Building Science Group, Karlsruhe Institute of Technology. F.A., R.T.H. and A.W. would like to give their appreciation to both the employers and the employees of the World Health Organization and Middle East Insurance Company for participating in the survey presented in this paper and for facilitating conducting the survey in their buildings.

C.C., C.B. and S.K. would like to thank the National Research Foundation of Korea (NRF) for supporting this work. (No. NRF-2017R1A2B4012122)

R.M. and R.U. would like to thank the BCIT Green Value strategies Fund for their financial support.

References

[1] Jazizadeh, F., Ghahramani, A., Becerik-Gerber, B., Kichkaylo, T., & Orosz, M. (2013). Personalized Thermal Comfort-Driven Control in HVAC-Operated Office Buildings. Computing in Civil Engineering, (June), 218-225.

[2] Tarantini, M., Pernigotto, G., & Gasparella, A. (2017). A Co-Citation Analysis on Thermal Comfort and Productivity Aspects in Production and Office Buildings. Buildings, 7(4), 36.

[3] Mofidi, F., & Akbari, H. (2019). An integrated model for position-based productivity and energy costs optimization in offices. Energy and Buildings, 183, 559-580.

[4] Ormandy, D., & Ezratty, V. (2012). Health and thermal comfort: From WHO guidance to housing strategies. Energy Policy, 49, 116–121.

[5] Xiong, J., Lian, Z., Zhou, X., You, J., & Lin, Y. (2015). Effects of temperature steps on human health and thermal comfort. Building and Environment, 94(P1), 144–154.

[6] Luo, M., Zhou, X., Zhu, Y., & Sundell, J. (2016). Revisiting an overlooked parameter in thermal comfort studies, the metabolic rate. Energy and Buildings, 118, 152–159.

[7] Humphreys, M. A., & Nicol, J. F. (2002). Adaptive thermal comfort and sustainable thermal standards for buildings. Energy and Buildings, 34(6), 563–572.

[8] Kučera, A., & Pitner, T. (2018). Semantic BMS: Allowing usage of building automation data in facility benchmarking. Advanced Engineering Informatics, 35, 69-84.

[9] ISO. (2005). ISO Standard 7730. Ergonomics of the Thermal Environment – Analytical Determination and Interpretation of Thermal Comfort Using Calculation of the PMV and PPD Indices and Local Thermal Comfort Criteria, ISO, Geneva.

[10] CEN. (2007). Standard EN15251. Indoor environmental input parameters for design and assessment of energy performance of buildings addressing indoor air quality, thermal environment, lighting and acoustics. Brussels.

[11] ASHRAE Standing Standard Project Committee 55. (2013). ASHRAE Standard 55-2013. Thermal environmental conditions for human occupancy. Atlanta, GA: ASHRAE.

[12] de Dear, R. J., & Brager, G. (1998). Developing an Adaptive Model of Thermal Comfort and Preference -RP 884 final report. ASHRAE Transactions, 104(1).

[13] CIBSE. (2015). Guide A: Environmental Design. Chartered Institution of Building Services Engineers.

[14] Indraganti, M., Ooka, R., Rijal, H. B., & Brager, G. S. (2014). Adaptive model of thermal

comfort for offices in hot and humid climates of India. Building and Environment, 74, 39–53.

[15] McCartney, K. J., & Fergus Nicol, J. (2002). Developing an adaptive control algorithm for Europe. In Energy and Buildings, 34, 623–635.

[16] Humphreys, M. A., Nicol, J.F., & Roaf, S. (2015). Adaptive Thermal Comfort: Foundations and Analysis. Adaptive Thermal Comfort: Foundations and Analysis. Routledge.

[17] Ning, H., Wang, Z., & Ji, Y. (2016). Thermal history and adaptation: Does a long-term indoor thermal exposure impact human thermal adaptability? Applied Energy, 183, 22-30.

[18] Teli, D., Gauthier, S., Aragon, V., Bourikas, L., James, P. A., & Bahaj, A. (2016). Thermal adaptation to high indoor temperatures during winter in two UK social housing tower blocks. Windsor 2016, (April), 7–10.

[19] Schweiker, M., & Wagner, A. (2018). Interactions between thermal and visual (dis-)comfort and related adaptive actions through cluster analyses. In BauSIM2018 - 7. Deutsch-Österreichische IBPSA-Konferenz Tagungsband (pp. 204–215). Karlsruhe, Germany.

[20] Daum, D., Haldi, F., & Morel, N. (2011). A personalized measure of thermal comfort for building controls. Building and Environment, 46(1), 3–11.

[21] Jazizadeh, F., Ghahramani, A., Becerik-Gerber, B., Kichkaylo, T., & Orosz, M. (2014). User-led decentralized thermal comfort driven HVAC operations for improved efficiency in office buildings. Energy and Buildings, 70, 398-410.

[22] Auffenberg, F., Stein, S., & Rogers, A. (2015). A personalised thermal comfort model using a Bayesian network. IJCAI International Joint Conference on Artificial Intelligence, 2015-Janua, 2547–2553.

[23] Li, D., Menassa, C. C., & Kamat, V. R. (2017). Personalized human comfort in indoor building environments under diverse conditioning modes. Building and Environment, 126(September), 304–317.

[24] Yun, G. Y. (2018). Influences of perceived control on thermal comfort and energy use in buildings. Energy and Buildings, 158, 822–830.

[25] Jung, W., & Jazizadeh, F. (2019). Comparative assessment of HVAC control strategies using personal thermal comfort and sensitivity models. Building and Environment, 158, 104-119.

[26] AlSkaif, T., Guerrero Zapata, M. and Bellalta, B. (2015) Game theory for energy efficiency in Wireless Sensor Networks: Latest trends. Journal of Network and Computer Applications, 54, 33-61.

[27] Chicco, G. (2012) Overview and performance assessment of the clustering methods for electrical load pattern grouping, Energy, 42:1, 68-80.

[28] Palm, J., Ellegård K. & Hellgren, M. (2018) A cluster analysis of energy-consuming activities in everyday life, Building Research & Information, 46:1, 99-113.

[29] Földváry Ličina, V., Cheung, T., Zhang, H., de Dear, R., Parkinson, T., Arens, E., … Zhou, X. (2018). Development of the ASHRAE Global Thermal Comfort Database II. Building and Environment, 142, 502-512.

[30] ISO, ISO Standard 7726. Ergonomics of the Thermal Environment. Instruments for Measuring Physical Quantities, ISO, Geneva, 2001.

[31] Humphreys, M. A., & Nicol, J.F., (2002). The validity of ISO-PMV for predicting comfort votes in every-day thermal environments. Energy and Buildings, 34(6), 667–684.

[32] Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World Map of the Köppen-Geiger climate classification updated. Meteorologische Zeitschrift, 15(3), 259–263.

[33] Brager, G. & de Dear, R. (1998) Thermal adaptation in the built environment: a literature review, Energy Buildings 27, 83–96.

[34] Gauthier, S., & Teli, D. (2018). Moving beyond averages: variations in reported thermal comfort. In Proceedings of the 10th Windsor Conference: Rethinking comfort. Windsor, UK: NCEUB.

[35] R Core Team. (2017). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

[36] Cohen, J. (1988). Statistical power for the social sciences. Hillsdale, NJ: Laurence Erlbaum and Associates.

[37] Gauthier, S. (2016). Investigating the probability of behavioural responses to cold thermal discomfort. Energy and Buildings, 124, 70-78.

[38] Schweiker, M., Huebner, G. M., Kingma, B. R. M., Kramer, R., & Pallubinsky, H. (2018). Drivers of diversity in human thermal perception – A review for holistic comfort models. Temperature, 5(4), 308–342.

[39] Schweiker, M., & Wagner, A. (2016). The effect of occupancy on perceived control, neutral temperature, and behavioral patterns. Energy and Buildings, 117, 246–259.

[40] Kim, J., Zhou, Y., Schiavon, S., Raftery, P., & Brager, G. (2018). Personal comfort models: Predicting individuals' thermal preference using occupant heating and cooling behavior and machine learning. Building and Environment, 129, 96–106.

[41] Wang, J., Wang, Z., de Dear, R., Luo, M., Ghahramani, A., & Lin, B. (2018). The uncertainty of subjective thermal comfort measurement. Energy and Buildings, 181, 38–49.

[42] Wilson Van Voorhis, C. R., & Morgan, B. L. (2016). Understanding Power and Rules of Thumb for Determining Sample Sizes. Tutorials in Quantitative Methods for Psychology, 3(2), 43–50.