Università Politecnica delle Marche
Scuola di Dottorato di Ricerca in Scienze dell'Ingegneria
Corso di Dottorato in Ingegneria Industriale

Sviluppo di un metodo innovativo per la misura del comfort termico attraverso il monitoraggio di parametri fisiologici e ambientali in ambienti indoor

Supervisor:

Prof. Gian Marco Revel

Co-supervisor:

Prof. Sara Casaccia

Ph.D. course coordinator:

Prof. Giovanni Di Nicola

Ph.D. Candidate:

Nicole Morresi

Università Politecnica delle Marche
Scuola di Dottorato di Ricerca in Scienze dell'Ingegneria
Corso di Dottorato in Ingegneria Industriale

Development of a novel procedure to measure human thermal comfort through physiological and environmental monitoring in indoor environments

Supervisor:

Prof. Gian Marco Revel

Co-supervisor:

Prof. Sara Casaccia

Ph.D. course coordinator:

Prof. Giovanni Di Nicola

Ph.D. Candidate:

Nicole Morresi

Acknowledgements

At the end of this PhD experience the best and warm thought that I always carry with me is the certainty that if I could go back, I would do this experience exactly from scratch. These 3 years have been very intense, stimulating, and plenty of personal satisfaction.

First, sincere thanks go to Professor Gian Marco Revel. He accompanied me daily on this journey. He has been always available to offer me his most valuable advice during all the stages of my research work. Thanks to him I learned that doing research is constant work but mostly passion for what you do.

A special thanks go to the now Prof. Sara Casaccia, who accompanied and followed me every day in this adventure. She has been always present in every moment of my PhD, guiding me through the process of becoming a young researcher.

A warm thanks are for Carlo. Side by side, he always supports me and reminds me that life is happier when you shared it with others.

A big thanks goes to my special colleagues: Serena, Riccardo Luca and Ilaria. We are not only colleagues, but also friends that share together most of their day.

And last but not least, I wish to thank all my family, who have always been supportive and curious about what I do here at university.

Abstract

Measuring human thermal comfort in indoor environments is a topic of interest in the scientific community, since thermal comfort deeply affects the well-being of occupants and furthermore, to guarantee optimal comfort conditions, buildings must face high energy costs. Even if there are standards in the field of the ergonomics of the thermal environment that provide guidelines for thermal comfort assessment, it can happen that in real-world settings it is very difficult to obtain an accurate measurement. Therefore, to improve the measurement of thermal comfort of occupants in buildings, research is focusing on the assessment of personal and physiological parameters related to thermal comfort, to create environments carefully tailored to the occupant that lives in it. This thesis presents several contributions to this topic. In fact, in the following research work, a set of studies were implemented to develop and test measurement procedures capable of quantitatively assessing human thermal comfort, by means of environmental and physiological parameters, to capture peculiarities among different occupants. Firstly, it was conducted a study in a controlled climatic chamber with an invasive set of sensors used for measuring physiological parameters. The outcome of this research was helpful to achieve a first accuracy in the measurement of thermal comfort of 82%, obtained by training machine learning (ML) algorithms that provide the thermal sensation vote (TSV) by means of environmental quantities and heart rate variability (HRV), a parameter that literature has often reported being related to both users' thermal comfort.

This research gives rise to a subsequent study in which thermal comfort assessment was made by using a minimally invasive smartwatch for collecting HRV. This second study consisted in varying the environmental conditions of a semi-controlled test-room, while participants could carry out light-office activities but in a limited way, i.e. avoiding the movements of the hand on which the smartwatch was worn as much as possible. With this experimental setup, it was possible to establish that the use of artificial intelligence (AI) algorithms (such as random forest or convolutional neural networks) and the heterogeneous dataset created by aggregating environmental and physiological parameters, can provide a measure of TSV with a mean absolute error (MAE) of 1.2 and a mean absolute percentage error (MAPE) of 20%. In addition, by using of Monte Carlo Method (MCM), it was possible to compute the impact of the uncertainty of the input quantities on the computation of the TSV. The highest uncertainty was reached due to the air temperature uncertainty (U = 14%) and relative humidity (U = 10.5%).

The last relevant contribution obtained with this research work concerns the measurement of thermal comfort in a real-life setting, semi-controlled environment, in which the participant was not forced to limit its movements. Skin temperature was included to the experimental set-up, to improve the measurement of TSV. The results showed that the inclusion of skin temperature for the creation of personalized models, made by using data coming from the single participant brings satisfactory results (MAE = 0.001±0.0003 and MAPE = 0.02%±0.09%). On the other hand, the more generalized approach, which consists in training the algorithms on the whole bunch of participants except one, and using the one left out for the test, provides slightly lower performances (MAE = 1±0.2 and MAPE = 25%±6%). This result highlights how in semi-controlled conditions, the prediction of TSV using skin temperature and HRV can be performed with acceptable accuracy.

Contents

List of Tables

List of Figures

Chapter 1.

Introduction

1.1 Background and Motivation

The point of departure of this work is that human perception of thermal comfort, which refers to the sensation of hot or cold that people experience every day, it is not perceived in the same manner among different humans. Given this assumption, it is expected that two people in the same room with specific thermal conditions, like air temperature, can experience different sensations: one can be pleasant, the other can be uncomfortable by feeling warm or cold. To prevent this situation, typically there are norms and indicators that should be respected to ensure the occupants a comfortable environment, in which neither hot nor cold sensation is experienced.

The thing is that, even if the standards imposed by the norms are met, people still experience discomfort and in particular "thermal discomfort". Broadly speaking, the potential impact that thermal discomfort has on occupants must be analysed from two points of view: in the first place, the prospect is that the well-being of the users is compromised or worsened, thus leading to a possible decrease of the users' state of health, in the long-term period. The second problem that arises concerns energy costs that must be faced to maintain adequate thermal conditions, which are often high, and it occurs that, expensive energy costs still do not provide thermal comfort anyway.

Buildings account for almost 30% of the world's total energy consumption, and 60% of this is due to the heating, ventilating, and air conditioning systems (HVAC). However, 89% of buildings cannot adequately meet thermal comfort demands for the indoor thermal environment, in which it is expected that more than 80% of the occupants are satisfied with the thermal environment [1]. These numbers above have created the specific need to renew the way of thinking thermal comfort assessment and building energy supply, in fact, the forthcoming trend is to move from a building-centered to a human-centered design.

Thermal environments assessment relies on past research performed on human subjects, but in extremely controlled, and therefore often unrealistic, environments. The measurement of thermal comfort, and in particular of its thermal sensation (TS), implies the creation of a proper thermal environment for the occupant; but the point is that occupants' thermal preference and sensation are dynamically changing in the real environment, because of the non-uniformity of the environment and to the change in the external weather [2]. Literature over the years offered a great number of studies that aim at estimating human thermal comfort, in terms of TS. The idea

behind many studies is to predict the number belonging to the 7-point scale of the American Society of Heating, Refrigerating and Air-Conditioning Engineers', which is the metric adopted to rank human thermal comfort; each sensation is associated with a value that ranges from -3 to 3, which corresponds to Cold, Cool, Slightly Cool, Neutral Slightly Warm, Warm, Hot [3]. The thermal sensation is therefore named Thermal Sensation Vote (TSV) and it is what researchers in the field of human thermal comfort are trying to measure with a variety of methodologies, experimental set-ups and different data types.

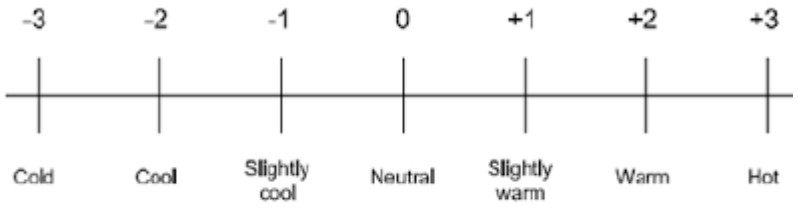


Figure 1. ASHRAE 7-points scale used to express the thermal sensation vote of occupants.

The innovative research aspect of human-centred design concerns the inclusion of physiological and personal parameters of the occupant into the measurement process, which are intended to capture the peculiarities of each occupant, leading therefore to a personalization of thermal comfort and the creation of personalized comfort models (PMC), which comes up one of the many aspects of the human centred design. PCM are identified as a prediction model based on the analysis on a single individual. Two issues need to be addressed when establishing a personal comfort model: improving the accuracy in predicting personal thermal comfort and minimizing interference to the occupant when the model is applied. Establishing a comfort model for an individual and choosing a limited number of sensitivity parameters to evaluate personal thermal comfort may be appropriate methods of ensuring the prediction accuracy and diminishing interference; however, literature is facing some challenges in identifying the most prominent parameters for PCM.

PCM allows to appreciate individual unique differences by monitoring each occupant's thermal response and these models are typically developed by means of AI algorithms, which include deep learning (DL) and machine learning (ML) algorithms, that are continuously trained and tested. In fact, PCM that rely on AI algorithms have the advantage of being continuously updated, in response to real-time data [3]. Researchers [15], [16], [17], [18], [19] have adopted a variety of machine learning methods to establish PCMs. The results show that the prediction accuracy of a PCM is much higher than that of the traditional PMV model when applied to individuals. Another aspect that should be taken into account is that occupants' thermal preference and sensation dynamically changes over time, and this is mainly due to the non-uniformity of the thermal environment inhabited by the occupant, or to a change in the outdoor

conditions. This last point is what makes the prediction of thermal comfort by means of PCM a challenge to be addressed.

1.2 Overview of the research

Recent developments in the field of wearable technologies, communications and computing, have made it possible to evaluate and measure the thermal comfort of occupants who are inside a building. However, despite the great variety of methodologies present in the state of the art, there are still gaps to be filled in terms of accuracy of the procedure and replicability in real-life contexts. In fact, due to the complexity and diversity of human behaviour and human's different perception of thermal comfort, extracting meaningful information and a common set of data that is suitable for every occupant is a complex task. The objective of this work is to tailor a methodology used for measuring human thermal comfort, into real-life context, rather than laboratory environments that are often far from reality. For this reason, a set of experiments that uses a dedicated measurement process and sensors network are proposed. In particular, this research aims to provide:

1. A methodology to be applied in real-life settings to improve the measurement of thermal comfort that could enrich the knowledge of the human-centric approach in buildings;
2. Improve the knowledge about the challenges that can be encountered when using wearable devices to measure physiological parameters related to human thermal comfort;
3. Determine the impact of the measurement uncertainty of the sensors network on the resulting measurement of thermal comfort;
4. Define the suitability of AI algorithms in the prediction of thermal comfort using a heterogeneous set of data.

1.3 PhD thesis context: the RenoZEB project

Figure 2. Logo of the project

Most of the activities described in this thesis work were developed within the RenoZEB project, "Accelerating Energy renovation solution for Zero Energy buildings and Neighbourhoods" funded by the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 768718. One of the scopes of the project is to investigate how the measurements of physiological parameters of occupant in buildings can be used for thermal comfort prediction

and improvement, in the framework of the work package 6 (WP6) entitled “Monitoring and Human-Centric Automated Control to maximize the value creation of nZEB renovation”.

The possibility of developing this approach is justified by available human-centric technologies, which are moving forward the measurement of human’s health, satisfaction, level of activity and even stress. This information is increasingly becoming of interest because they can be coupled to the usual empirical parameters, e.g., air temperature, humidity, illuminance, and can provide a full evaluation of everyone’s experience. The human-centric approach can be used to optimize traditional control systems for the enhancement of human comfort and satisfaction towards the daily indoor environment, such as in the context of this research activity. Considering these new developments, RenoZEB aimed at exploring this field and test innovative solutions whose results may help to increase the knowledge related to human-centric approach adopted through the deployment of wearable solutions.

Chapter 2.

State of the art

2.1 Thermal comfort measurement

First of all, in order to study and try to understand thermal comfort, it is necessary to start from its formal definition and the existing standards adopted to measure it. The American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) defines thermal comfort as "the condition of the mind in which satisfaction is expressed to the thermal environment"; the definition itself suggests that is a parameter related to the perception of the single individual and it is related to physical, physiological, and psychological factors. The importance of measuring human thermal comfort is due to its connection with human wellbeing; for this reason, in the next sections, this aspect will be deeply investigated. In addition, there will be a detailed analysis of the state of the art about the methodologies adopted for measuring human thermal comfort, its relationship with human wellbeing and the latest methodology adopted in literature to assess the problem of thermal comfort personalization.

2.2 Thermal comfort, well-being and IEQ

Thermal comfort has been deeply studied in years, given its related importance to subjects' well-being. The wellbeing dimension has many facets and thermal comfort was proof to have a high impact on it. Well-being is an old concept, studied by the scientific community in relation to human's health and the concept itself suggests that there is a focus on the single user, rather than a group of people. It is worth noting that well-being and thermal comfort are profoundly related to one more general parameter, which is the indoor environmental quality (IEQ). Literature has extensively reported the impacts that IEQ has on occupants' health and wellbeing, especially in indoor environments [4]. IEQ relies on different factors which are indoor air quality (IAQ), the thermal comfort, the lighting environment, and the acoustic comfort. If these parameters are not maintained at acceptable levels, it can have a negative impact on occupants' health and work efficiency [5]. Extensive research has proved that thermal comfort has a significant role in affecting wellbeing, health, and productivity; as an example, studies highlighted that thermal discomfort in buildings can cause psychological stress, depression, and anxiety, as well as poor physical health, expressed as heart disease, insomnia, a headache, and low arousal levels.

Thermal comfort is a topic of interest in many types of facilities like educational, office and residential buildings. As an example, [6] found that inadequate indoor air temperature has the highest effect on occupant's thermal comfort and therefore on productivity; another study that presents relevant outcomes is reported in [7]: results indicate that comfortable indoor thermal conditions can have beneficial impacts on workers' well-being and productivity, such as higher operational rates, lower production losses, fewer sick leaves, and reduced health-related costs.

Moreover, indoor thermal discomfort also affects human health and can be particularly critical in the case of susceptible subjects [8]–[10]. Depending on their age, people are subjected to diverse adaptive mechanisms: for example, elderly people's health can be compromised if their living environment is too cold or too warm. Prolonged exposure to warm conditions can trigger heart-related illnesses or heart failures, while a long exposure to cold conditions gives rise to the lowering of body core temperature, which may conduct to drowsiness, lethargy and even death. Therefore, maintaining adequate thermal comfort impacts the health of occupants, besides their comfort or satisfaction [11].

2.3 Thermal comfort models

The most used thermal comfort standards are based on steady-state models or adaptive models. The most famous steady-state model is the predicted mean vote (PMV) model also recognized in literature as the Fanger's model. This model aims to predict the mean thermal sensation of a group of people that are in the same environment. It is a steady-state model that refers to the heat balance model of the human body and is based on the computation of the PMV index, which is assessed in relation to 6 different parameters listed as follows: four parameters are environmental parameters (air temperature, air velocity, relative humidity and mean radiant temperature). The other two quantities are subjects' related factors: clothing insulation and metabolic rate. Their assessment is regulated by two other standards, ISO 9920 and ISO 8896, respectively. The resulting function is expressed in Equation 1:

$$PMV = f(t_a, t_r, v_a, p_a, I_{cl}, M) \tag{1}$$

Where t_a is the air temperature, t_r is the mean radiant temperature, v_a is the air velocity, p_a is the water vapour partial pressure, I_{cl} id the clothing insulation and M is the metabolic rate. Moreover, Fanger's model provide the Predicted Percentage Dissatisfied (PPD) index that estimates the percentage of people that experiences thermal discomfort.

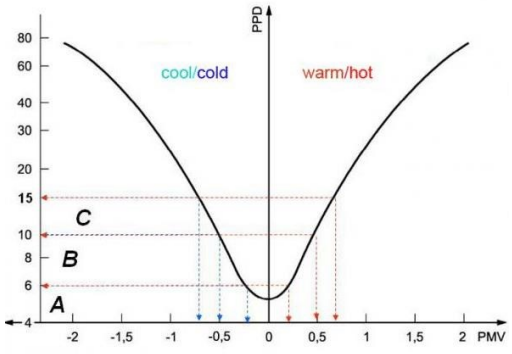


Figure 3. Relationship between the PMV and the PPD.

Using both indices, ASHRAE 55 dictates that thermal comfort can be achieved based on an 80% occupant satisfaction rate or more. The remaining percentage of people can experience 10% dissatisfaction based on whole-body discomfort and 10% dissatisfaction based on local discomfort/partial body discomfort (includes fewer factors than whole-body).

In order to comply with ASHRAE 55, the recommended thermal limit on the 7-point scale of PMV is between -0.5 and 0.5. ISO 7730 expands on this limit, giving different indoor environments ranges. ISO defines the hard limit as ranging between -2 and +2, for existing buildings between -0.7 and +0.7, and new buildings ranging between -0.5 and +0.5.The PPD can range from 5% to 100%, depending on the calculated PMV. These comfort values will vary depending on where the occupant is located in the building. In order for comfort ranges to comply with standards, no occupied point in space should be above 20% PPD.

The assessment of thermal comfort is usually reinforced by the surveys which compare the PMV values obtained from the model, with those obtained from the real sensations perceived by the occupants. It was observed that the PMV model underestimate the real thermal sensation of the people as reported in [12]. A possible explanation of this problem is that PMV index is not intended for considering psychological and behavioural adaptations. Thus, adaptive models have begun to spread.

Adaptive models are based on collecting simultaneously data from the thermal environment and from the thermal response of the subject, to assess the indoor thermal conditions and to determine the influencing parameters that could satisfy occupants' preferences. The adaptive approach refers to statistical studies conducted in buildings, and it is mainly focused on naturally ventilated buildings. It has been observed that people are more tolerant than what the Fangers' model suggests. The adaptive approach presents a linear correlation between comfort temperature of the occupants indoor and the outdoor air temperature. The main idea is that occupants inside the building are no longer passive agents, as it appears in the PMV model, but are considered as

active agents that interact at all levels with the environment. This model is based on the ability of the occupants to adapt to environmental conditions, acting on the variables that affect them.

In recent years, attention has been increasingly focused on the fact that the models described (adaptive and PMV) do not take into account the psychological aspect of the subject, and risk to overshadow the personal aspect that deeply influences the perception of thermal comfort. In this regard, new line of research that is developing more and more, is focused on PCM, whose aim is not to predict the average response of a large population but to focus the prediction of thermal comfort of the single subjects. PCM are developed in literature by introducing brand new parameters inside the model: in particular, they are heterogeneous set of parameters that include environmental, behavioural and physiological parameters.

2.4 Sensors for the Measurement of Physiological Signal Related to Thermal Comfort

To estimate the thermal comfort, expressed by the thermal sensation vote (TSV), there many studies in the literature, that have been conducted based on various physiological parameters. In particular, the trend is the migration towards a less-invasive equipment, that the user is able to wear and conducting at the same time the daily activities.

There are currently several types of sensors and devices on the market which are used to measure physiological signals that can be related to thermal comfort. Thermal comfort can be evaluated from multiple physiological signals, i.e. electrodermal activity (EDA), electroencephalography (EEG) and electrocardiogram (ECG) signals. EDA, EEG and ECG signals can be measured using biomedical devices but also by low-cost wearables already available on the market. The difference between the two categories of sensors lies in the parameters measured, accuracy, cost, and the sensors being comfortable to wear for a prolonged time. EDA provides the measurement of changes in skin conductivity, which is associated with the activity of sweat glands that, in turn, reflects the activity of the central nervous system [20]. It is used in various thermal comfort-related studies in combination with other physiological parameters (e.g. skin temperature, heart rate) to develop customized models to discover how humans react to different external environments [13]–[16].

Human thermal comfort is also analyzed using EEG signals. These signals are collected by placing a helmet equipped with electrodes on the user's head. The spectral power of EEG can be used to build models that discriminate different feelings associated with thermal comfort [17]. The user's thermal sensation can be correlated to indices obtained from EEG. Given this assumption, there are researchers that are developing systems based on the brain-computer interface (BCI) for the control of conditioning system to obtain optimal thermal comfort

conditions [18]. However, EEG sensors are expensive and are not suitable for daily monitoring, since the user manages to wear them for up to three hours [19].

ECG sensors, too, are often used to study human thermal comfort. ECG records the electrical activity of the heart using electrodes placed on specific locations of the body. The parameters that can be extracted from ECG for the evaluation of thermal comfort are the heart rate (HR) and heart rate variability (HRV). However, ECG equipment can be a source of discomfort if worn for a long period. ECG traces can be also extracted from 24-hour Holter monitors, which are portable ECG devices that record the heart's rate and rhythm for a period of time of at least 24 hours. The advantage of measuring HRV and HR is that they can be also extracted from alternative wearable sensors that are less invasive. Wearable sensors, e.g. multiparametric chest belts, smartwatches, smart bands, represent a good trade-off between accuracy, intrusiveness and user acceptance because they have become part of an individual's daily routine [20]–[23]. They are also less expensive than EDA, EEG and ECG technologies.

2.5 Measurement of Heart Rate Variability for Thermal Comfort

Human thermal comfort is handled by thermoregulation, which is the process that allows the body to maintain its core internal temperature and is designed to restore the body's homeostasis. This process is managed by involuntary mechanisms that take place in the brain's hypothalamus and therefore, recent studies are focusing on estimating thermal comfort of individuals through the analysis of specific human physiological signals. In this context, the focus is on the heart rate variability (HRV), obtained by measuring the difference in time between two consecutive heartbeats of the electrocardiogram (ECG) signal. An example of ECG trace is shown in Figure 4.

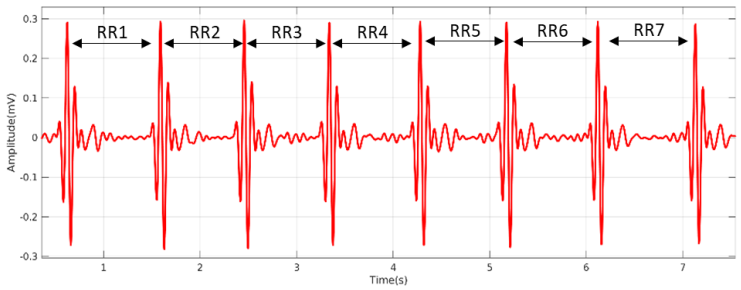


Figure 4. ECG signal and distance between each R wave that generates RR intervals.

However, recent findings have demonstrated that HRV can be measured in a less-invasive way, by collecting the photoplethysmography signal (PPG).

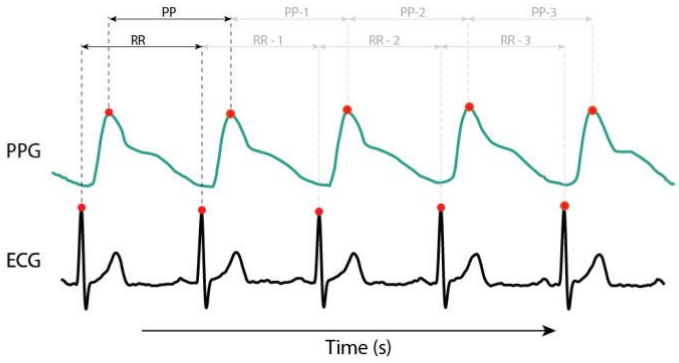


Figure 5. Example of a PPG signal (upper) compared with the correspond ECG signal (lower). Both methodologies can be used to measure RR intervals to derive HRV.

PPG is a low-cost technique that detects blood volume changes in the blood vessels of superficial tissue, therefore it is a non-invasive methodology that collects physiological signal directly from the skin surface. Therefore, PPG signal processing provides HRV signal. HRV itself is not the only contributor to understand thermal comfort but needs further manipulations to extract useful indicators related to this issue. In fact, the power spectrum analysis provides the elements for analysing the balance between the sympathetic nervous system and the parasympathetic nervous system. High frequency (HF) components of HRV (0.15- 0.4 Hz) are supposed to be due to the action of the vagus nerve, while low frequency (LF) components (0.04 – 0.15 Hz) originate from the sympathetic nervous system. Their ratio LF/HF expresses the balance between these two subsystems and therefore when there is the need for an adjustment to restore an unbalancing condition.

For this reason , HRV is typically analyzed in terms of time-domain and frequency-domain parameters after using specific processing of data [24], [25]. Time-domain parameters are statistical indices that evaluate the variability of the HRV signal collected and are useful to interpret the fluctuations during cardiac cycles. Whereas frequency domain indices are obtained starting from the power spectral density (PSD) of the HRV to decompose the variation of each HRV signal into its fundamental oscillatory components [26]. As anticipated before, most relevant components are obtained by computing the PSD in specific frequency bands: low frequency (0.04 - 0.15 Hz), high frequency (0.15 - 0.4 Hz). These frequency bands are considered to be linked to the thermoregulation mechanism. Thermoregulation is result of complex mechanisms that are modulated by mutual interactions between the sympathetic nervous system (SNS) and the parasympathetic nervous system (PNS). SNS is responsible for sweating and vascular constriction for heat generation, while PNS acts through the vagus nerve for vascular constriction. Frequency domain indices are linked to the activity of SNS and PNS. In fact, HF components are supposed to be due to the vagal activity, whereas LF components are originated from the SNS. The ratio between these components (LF/HF) expresses the balance between these two subsystems and is, therefore, subjected to variations in case of an external stimulus. Recent

works have developed different approaches to explore the relationship between HRV, its indices and thermal comfort perception. Literature works related to thermal comfort measurement involved in their research a variable number of participants. For example, [27] recruited six participants to explore the relationship between HRV features and different environmental conditions; also [28] in their work, predicts the user's thermal comfort states using HRV signal collected from a smartband, using a sample size of six participants. However, the sample size in thermal comfort studies can be also be reduced, as it is reported in [29] and [30], which tested their solution on a sample size of 4 and 1, respectively.

Literature has pointed out that there is a quadratic relationship between LF/HF and thermal discomfort. It has been demonstrated that LF/HF increases in an uncomfortable thermal environment (cold, warm) and decreases in a thermally comfortable environment, as reported in Figure 6 and Figure 7 [22] [28] [31].

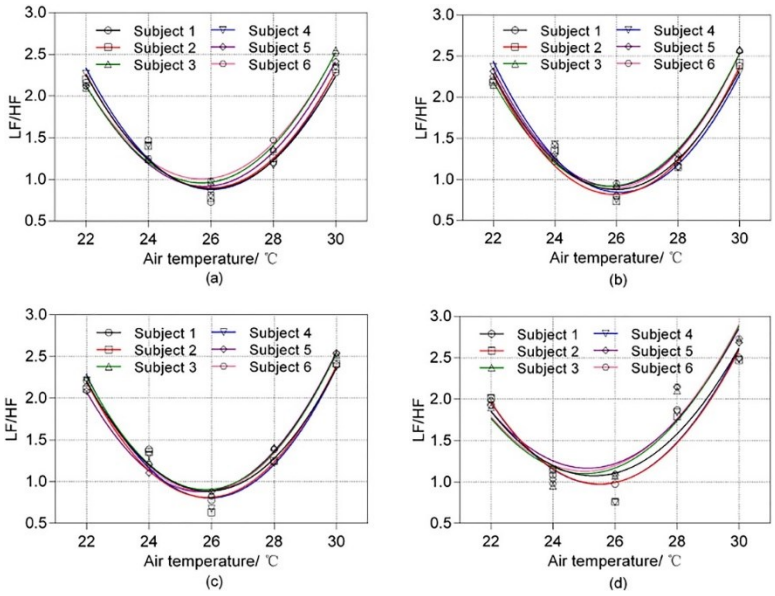


Figure 6. Changes of LF/HF under different air temperature.

Moreover, LF/HF is also subjected to a rise when humidity reaches high values (>80%) and when the air speed increases in case of low temperature [27], [28], [31]. To collect HRV data, different sensors are used in current literature: ECG is still the predominant methodology in many researches [22], [32]. However, works in literature are also exploring less invasive sensors like smartwatches or wristbands, which, being wireless, do not prevent users from moving freely. Empatica E4 wristbands are often adopted, however they are very expensive and therefore cannot be provided to a large population of users [25], [30], [33]. In this perspective, HRV-related works are developed through the use of commercially available smartwatches, which are less expensive and provide reliable accuracy [34], [35].

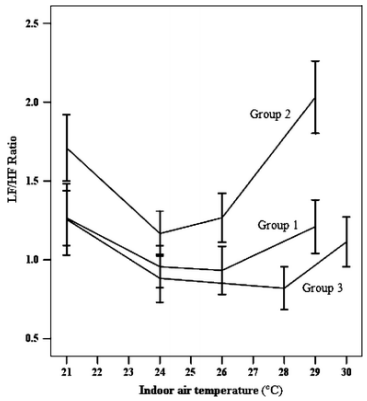


Figure 7. Low- versus high-frequency (LF/HF) ratio for variation in heart rate as evaluated by the ECG over 5 min in three groups of subjects at different indoor temperatures. For each group, n = 11; for group 1 the temperature changed from 21 to 29°C, for group 2 the temperature changed from 29 to 21°C, and for group 3 the order was 21, 28, 30, and 24°C. LF/HF ratio means the ratio of absolute power in low- and high- frequency bands.

2.6 Measurement of skin temperature for Thermal Comfort

Skin temperature can be retrieved in two different ways: by using contact measurement such as sensors placed directly on the skin at different locations, or in an unobtrusive way, by using thermal images acquired by means of non-contact infrared sensors.

Contact measurements include the employment of wearable sensors or sensors directly placed upon human skin. For example [36] proposed two different methodologies to predict the thermal comfort of occupants in terms of thermal state, using a wearable sensors. The first one is about a deep CNN that uses the hand skin temperature profile; the accuracy in the detection of the thermal state is 93.3%. The second methodology is based on a support vector machine (SVM) approach using six physiological features, that achieve the 90.6% of accuracy.

[37] proposes a fuzzy control method of the indoor temperature using physiological data (skin temperature and heart rate) to predict the thermal sensation. The idea is to show that thermal sensation prediction-based control can adjust the temperature set-point by monitoring physiological data from the subjects without interfering with their regular works. A more invasive methodology is presented in [3]: a prediction accuracy of 89.2% was obtained for the models based on the temperature of the point 2 mm above the wrist and the skin temperatures of the wrist and neck.

Among the unobtrusive measurements [38] presents non-intrusive thermal measurements to predict ASHARE 7-point thermal sensation; the system provides a real-time comfort estimation that is able to extract and compute facial temperature features to provide personalized comfort models. The study highlights that the methodology can reduce the error on thermal comfort by 64% over the traditional method (such as PMV).

[39] proposes a non-invasive approach for automatic prediction of personal thermal comfort and mean time to warm discomfort using machine learning techniques. The described prediction framework uses temperature information extracted from multiple local body parts to model an individual's thermal preference, with sensing measurements that capture local body part variance as well as differences between body parts. The proposed thermal models were tested on subjects' data extracted from an office setup with room temperature varying from low (21.11 °C) to high (27.78 °C). When all proposed features were used, personal thermal comfort was predicted with an accuracy higher than 80% and mean time to warm discomfort with more than 85% accuracy.

2.7 AI for the measurement of thermal comfort

Most of the work analysed in Section 2.6 and 2.7 have highlighted that AI approaches produce relevant results when physiological and environmental parameters are merged to forecast occupants' thermal sensation. In this work, [31] it is demonstrated that, depending on the thermal environment, it is possible to predict the thermal state of each subject by using HRV indices with an accuracy up to 93.7 %, thanks to the implementation of Random Forest (RF) and Support Vector Machine (SVM) algorithms. Another study [40] uses Extra Trees Classifier (ET) ML algorithm to predict participants' combined thermal sensation vote in response to cold environmental conditions: HRV features produce an accuracy of 73.04 % with an absolute error of 0.299; the accuracy increases up to 79.01 % when additional parameters (e.g. mean air temperature, mean radiant temperature, relative humidity) are added to the model. In addition to classification algorithms, to estimate each subject's thermal comfort level [41] uses ML regression models that produce an RMSE value up to 0.04± 0.01.

To the best of my knowledge, in literature there is a lack of studies that employ Deep Learning algorithms for thermal comfort prediction starting from HRV signals or in combination with skin temperature measurement. However, there are works that show various fields of application of HRV used to forecast some user-related physiological or psychological states. References [42], [43] trained a LSTM neural network to predict participants' health using HRV data alone. They investigated time domain, frequency domain and typical HRV measures and were able to predict mental health with a classification accuracy of up to 83% and 73%, with five- and two-minute HRV. They use a novel unified deep learning framework for sleep-wake classification with two heterogeneous sensors which include acceleration and HRV. In [44] Convolutional Neural Networks (CNN) are employed to detect whether a person is awake

or asleep starting from HRV features. These networks can identify HRV fluctuations and relate them to human physiological states. Also in [45] the authors claim that Deep Learning can help to reveal underlying patterns in the ECG trace which otherwise could not be observed and, for this reason, they used a 1-D CNN that employed HRV feature maps for stress state identification, obtaining an accuracy of up to 89.8% [46] also uses HRV as input to train a CNN for the prediction of sleep scoring.

2.8 Limitations in the state of the art

The study of the state of the art provided the necessary knowledge to establish the limits of the current level of research concerning thermal comfort measurement through personalized approaches that includes new parameters.

For example, most of the studies were carried out in very remote and highly controlled conditions, the results of which are often not applicable in real working contexts. Test users are often forced to wear a wide range of devices that cause severe emotional distress, which can negatively affect measurements of physiological parameters. In fact, it has been repeatedly found that the user's psychological state can impact the user's thermal sensation.

Often there are studies that combine heart rate and skin temperature, but still as far as my knowledge is concerned, there are no studies that measure thermal comfort by combining HRV and skin temperature; in addition, given the variety of types of data that are collected to obtain customized models of thermal comfort, the literature lacks studies that use AI starting from data such as HRV, its features or skin temperature.

Consequently, these gaps identified in the state of the art, provided the basis for developing this thesis work, which aims to measure the thermal comfort of users in indoor environments, through an innovative approach that includes the combination of a series of indoor parameters and physiological quantities measured from wearable devices and inserting the experimental campaigns in contexts close to the reality.

Chapter 3.

Implementation

This section is dedicated to the detailed description of the methodology developed and implemented to measure human thermal comfort, which has been deployed over the past three years through a series of dedicated experiments. Each experiment was developed with a different experimental set-up, which varies in terms of operational conditions of the test-rooms, participants, outdoor conditions and sensors employed.

The basic principle of this research work is explained in Figure 8: all the experiments that are presented aim to expose the users participating in them to a variety of environmental conditions with the aim of causing thermal stress. The customization of the measurement takes place thanks to the network of sensors for the measurement of physiological parameters: it will start with a more invasive and complex network, ending with a minimally intrusive smartwatch, which in everyday life performs the same function as a common watch, easy to insert into the daily routine of people. Participants in the early stages of the study will be more limited in the movements and type of activities conducted during the tests, to try to minimize errors due to user movements or due to other perturbations rather than environmental conditions.

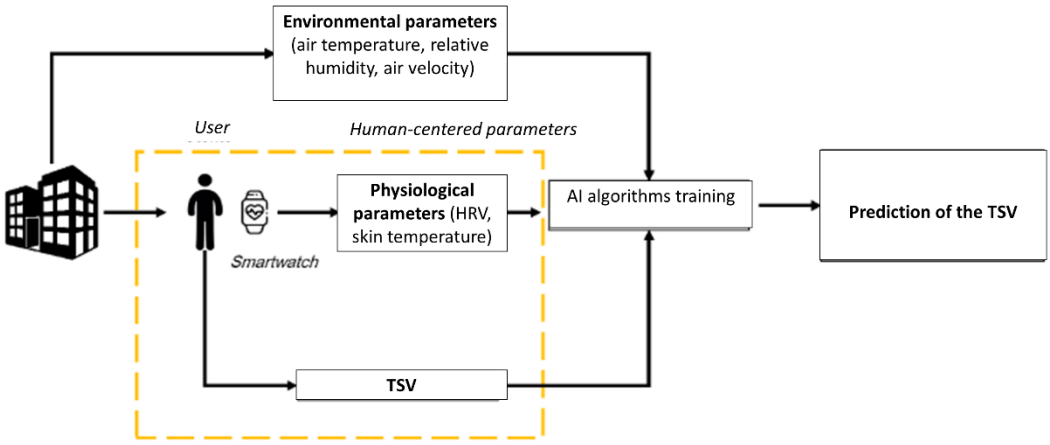


Figure 8. Technical workflow of the approach used to develop personalized thermal comfort model, using a human centric approach.

Data will be extracted and processed to create heterogeneous datasets, which will be used to train AI algorithms, capable of providing a measure of the TSV, which represents the user's thermal comfort. The timeline of the studies carried out is summarized in Figure 9.

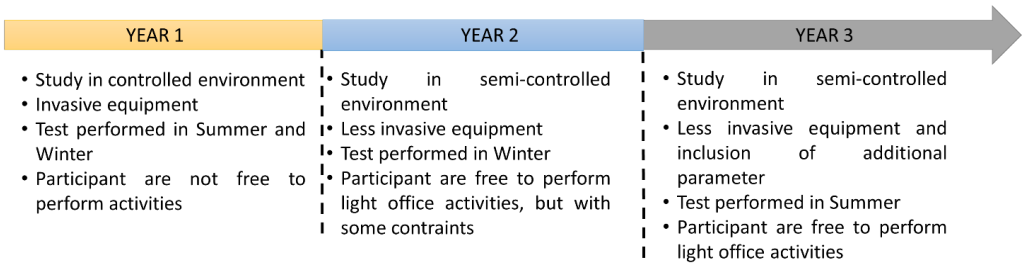


Figure 9. Workflow followed to develop the human centric approach for measuring human thermal comfort.

The first research is about human thermal comfort experiments conducted in controlled environment or test room: the definition of controlled environment, in thermal comfort studies, states that in controlled experiments all the parameters that belongs to the PMV model have to be controlled [47] and this is what occurs during the experiment conducted in the first research; controlled environments in literature are typically climatic chambers.

The second and the third research are conducted in a semi-controlled test-room, which means that at least one of the monitored quantitates is not controlled at all [48]. For example, clothing, relative humidity and air velocity are kept uncontrolled and the only parameter that is controlled is the air temperature. The decision to implement studies conducted in progressively less-controlled environments is because literature has repeatedly established that, in the context of a climatic chamber (controlled environment) it is possible to measure human thermal comfort integrating physiological parameters. What is missing, on the other hand, is the possibility of implementing the knowledge acquired in the climatic chamber into real-life settings, where the occupant is not bound by invasive sensors, and the thermal condition of the environment dynamically changes.

3.1 Development of a methodology for evaluating the correlations between thermal comfort and HRV in controlled environment

This paragraph presents the preliminary methodology that was first implemented to start working with the measurement of human thermal comfort, including the measurement of physiological parameters. The main idea was to develop a novel measurement protocol capable of assessing occupants' thermal comfort starting from physiological parameters collected while environmental conditions were varying, in a controlled environment such as a climatic chamber. The implemented methodology aims at comprehensively assessing thermal comfort perception of occupants to have a preliminary occupancy-related comfort perception scheme in indoor environment. The experimental campaign comprehends two test series conducted during heating and cooling seasons, respectively, for a total amount of 62 tests. All the participating subjects are volunteers in good health status, i.e. not affected by chronic diseases which could alter their environmental quality perception. The protocol includes the simultaneous collection of (i) physical environmental parameters, (ii) human physiological metrics, and (iii) subjective responses of the occupants. The measurement set-up includes different monitoring tools whose outputs are therefore synchronized and analyzed. More details about the measurement test procedure, implied tools and data analysis are given in the following sections.

3.1.1 Measurement set-up

The tests all took place in the same mechanically controlled environment. It is a house-like cubicle located within the Engineering campus of the Perugia University (Italy). The cubicle inner dimensions are 3 m × 3 m and the reduced volume allows to control the indoor environment by means of the installed air-conditioning system which is a heat pump with an inverter. The system provides three different levels of ventilation, i.e. low, medium, and high-speed mode, while no air change is provided given the short time duration of each experiment. The Southern wall has a rectangular window which is shaded during all the tests to keep a constant illuminance level provided only by the artificial lighting system of the test-room . According to collected data, the illuminance level provided during all the tests ranged between 237 lx and 389 lx, with an average value of 280 lx and a standard deviation of 25 lx.

The internal conditions were controlled by adjustments of the HVAC operation based on the real-time monitoring of indoor physical quantities. In fact, the space is continuously monitored by means of a fixed microclimatic station located in the center of the room recording data every minute, and the monitoring system datalogger is equipped with a display showing instantaneous

values. The monitored parameters are: air temperature at both 1.1 and 0.1 m [°C] from the floor, relative humidity [%], superficial temperatures of floor, roof, North and South-facing walls [°C], black globe temperature [°C], net-radiation between glazed and opaque surfaces [W/m2], air velocity [m/s], concentration of CO2 [ppm], and illuminance level [lux]. The accuracies of all involved sensors are reported in Table 1. All the sensors are compliant with ISO 7726 [49].

Table 1. Technical information of the sensors for environmental parameters monitoring.

Sensor	Parameter monitored	Accuracy
Thermal hygrometer	Air Temperature	± 0.1°C
	Relative humidity [%]	± 1.5%
Surface temperature sensor	Floor temperature[°C]	± 0.15°C
	Roof temperature [°C]	± 0.15°C
	Walls temperature [°C]	± 0.15°C
Black globe radiant temperature sensor	Mean radiant Temperature [°C]	± 0.15°C
Hot wire anemometer	Air velocity [m/s]	±0.05 m/s
CO_2 sensor	CO_2 concentration	±50 ppm (±2%)
Luxmeter	Illuminance [lx]	±5%

Predicted Mean Vote (PMV) and Draught Rate (DR) are also computed to get a complete thermal comfort evaluation according to formulas provided in current international standard ISO 7730 [50].

The physiological parameters of the tested subject are measured by means of three wearable systems:

- The subject wears a multi-parametric belt, BioHarness 3.0. from Zephyr, with an attached electronic module at the thorax level for the ECG signal acquisition (sampling rate 250 Hz, Heart Rate accuracy ±1 bpm, operating range 25–240 bpm) [23], [51]. Measured data are stored within the device and downloaded at the end of the test.
- A wireless neural headset with 14 electrodes gives the EEG signal with a sampling rate of 128 Hz per channel (operating bandwidth 0.16–43 Hz) [52].
- Finally, the EDA signal is measured through a BITalino acquisition board (sampling rate 100 Hz, operating range 0–1 MΩ) [53]. The subject wears two EDA electrodes on the index and middle finger of the non-dominant hand to reduce artifacts movements, while the acquisition board and the sensors modules are fixed on the same arm. Open-Source software allows waveforms acquisition through Bluetooth communication protocol.

Even if the adopted wearable systems have been selected in order not to be invasive for the subject, attention is paid to possible discomfort perception due to these systems, as better

specified in the following section on experimental procedure. Personal perception about sensors was always asked to each subject during the test sessions and eventual discomfort conditions were identified as outliers of bias sources during data analysis.

The subjective information of the subjects was collected directly in the test-room during different parts of the test, as it is described in detail in the following section. In this part of the survey, questions are developed according to ISO 10551 [54] which is focused on the thermal perception assessment. In particular, the sensation vote for each domain is given through a 7-points scale going from −3 to +3, where 0 corresponds to neutrality. Figure 10 shows the above described experimental set-up.

Figure 10. Experimental set-up: (a)microclimatic station, (b)EEG sensor EPOC+, (c) Zephyr Bioharness 3.0, (d)BiTalino sensor, (e) representation of the participant during the experiment.

3.1.2 Test procedure

Two different series of measurement were performed, one during the winter season and one during the summer season. The winter series includes 34 participants, while 28 participants were involved in summer tests, providing a total amount of 62 participants. This sample size was estimated according to literature findings: each participant will provide a series of responses (personal, physiological and environmental) that will provide an appropriate amount of data to be analysed, in line with literature findings [55]. Participants were voluntarily recruited; they were affected by no pathologies and the general anthropometric information are resumed in Figure 11.

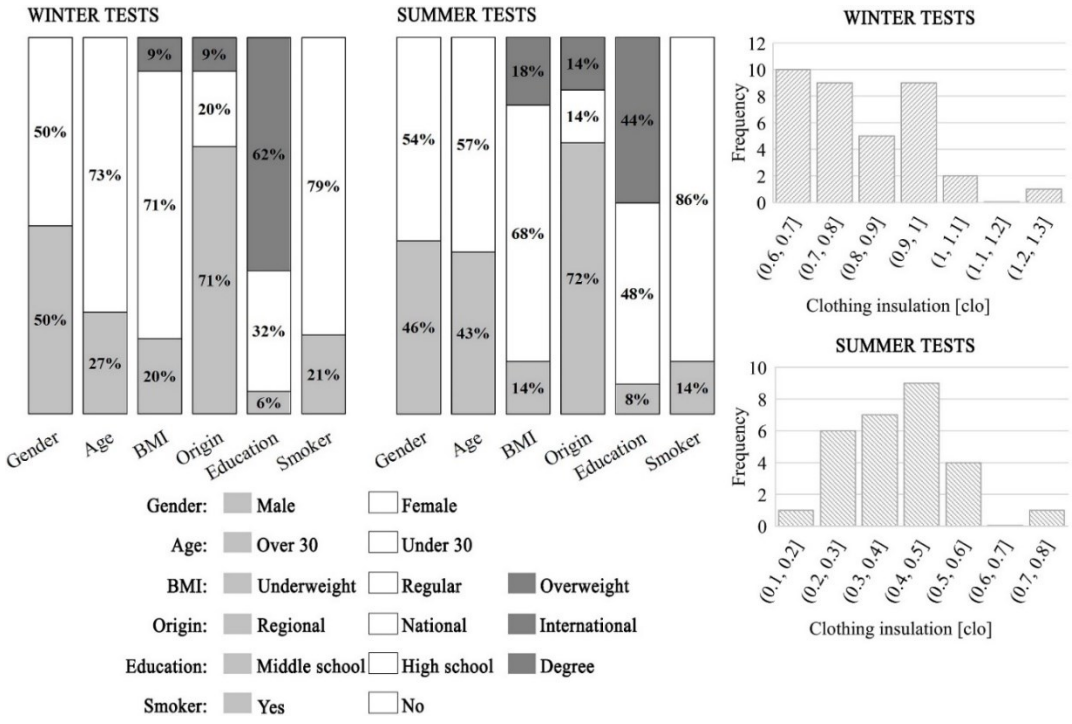


Figure 11. Anthropometric information of the participants during summer and winter tests.

The two experimental sessions consist of the same experimental set-up, i.e. same test-room, same monitoring sensors, but they are slightly different regarding the procedure, as reported in Figure 12.

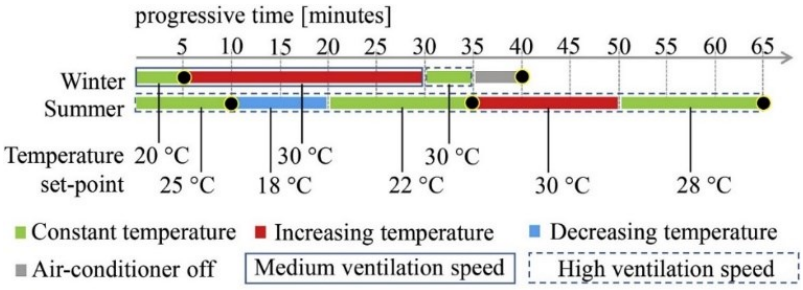


Figure 12. Description of the operational conditions of the two tests.

During both the campaigns, the subject is firstly exposed to a stabilization period of 20 min outside the test-room, in the conditioned spaces of the closest building. The space is conditioned at the same temperature settled in the test-room for the first part of the test which corresponds to the neutral state according to standards [50], [56], [57], i.e. 20 °C and 25 °C, winter and summer respectively. In particular, the initial values of air temperature are selected considering a classic level of expectations, i.e. environment belonging to II category, for subjects wearing typical clothes for the considered season, i.e. clothing insulation of 1.0 clo and 0.5 clo in winter and summer, respectively. The initial temperature set point in summer is 1 °C lower than the value suggested by standards, i.e. 25 °C instead of 26 °C, to guarantee reasonable time for the test considering the double temperature ramp, as specified later. Such initial temperature assumptions are necessary since no specific constraints are given to the subjects considering the personal attitude of everyone in wearing comfortable suites according to weather and his/her own personal preferences. Nevertheless, clothing information are collected in the first part of the survey, as specified in the above section, in order to check the validity of such assumptions during the data analysis process. In particular, the clothing thermal insulation of each tested subject is computed according to the provided list of worn garments and corresponding thermal insulation values provided in standard ISO 7730 [50].

During the stabilization period of 20 min before the test, the participant wears the chest strip for the ECG recording. Once in the test-room, the subject seats between the microclimatic station and the window, and the monitoring set-up is completed by connecting the headset and the electrode for the EEG and EDA signals record, respectively. These actions are made by an operator staying with the subject in the test-room for the whole test taking continuously notes of the expressed perceptions of the subject including possible complaints due to the equipment worn for the physiological monitoring. The operator has also the task to carefully check the actual indoor conditions as shown by the monitoring system in real-time to be compliant with the experimental procedure design. When the set-up is completed, the test starts.

Procedure here presented is specifically tailored to collect physiological signals and subjective responses under varying thermal conditions. These variable conditions are provided in a single test-room and thus subjects' responses are observed during both temperature ramps and at different constant temperatures. During the first tests series, acclimatization at constant temperature is provided in 5 min while 25 min are accounted to reach an air temperature increment of about +10 K. The short time of acclimatization is chosen to limit the experiment duration and possible participants complains due to worn monitoring equipment. Furthermore, 5 min is a reasonable time since initial thermal conditions are the same that the subjects experienced for 20 min before the test, and the space is gradually warmed up till reaching the final temperature of 30 °C. Once verified that physiological sensors do not bother participants, the duration of the second tests series is set as long as 65 min. Subjects' response is tested under two temperature ramps with a longer period of acclimatization, i.e. 15 min. Summarizing, during winter, the measurement test lasts 40 min in total, progressively: 5 min of acclimatization at 20 °C, 30 min of warming up with an air-conditioning set-point at 30 °C, 5 min at the same

temperature, but increasing the ventilation speed, i.e. from medium to high speed, and 5 min with the conditioning system switched off. The operator asks the subject to answer the third part of the survey at the end of the first period of constant air temperature at 20 °C, and at the end of the whole test. During summer, the test lasts 65 min in total, progressively: 10 min of acclimatization at 25 °C, 10 min of cooling down with an air-conditioning set-point at 18 °C, 15 min at constant temperature, i.e. 22 °C, 15 min of warming up with an air-conditioning set-point at 30 °C, 15 min at constant temperature, i.e. 28 °C. The ventilation speed is settled at the high mode for the whole test in summer. The operator asks the subject to answer the third part of the survey at the end of each period characterized by stable temperature, i.e. at 25 °C, 22 °C, and 28 °C. During the whole test and for both the seasons, the subject expresses any kind of changes in his/her environmental perception.

3.1.3 Data analysis

Due to the variety of monitoring systems adopted in the current experimental work, a first process of data synchronization is needed to provide exact correspondence among all the data time-series at disposal for the analysis, i.e. environmental data, EDA, ECG, and EEG signals.

Physiological raw data are therefore processed to reduce signal noise, get a smooth waveform, and extract useful features to be correlated to the measured environmental data and subjective responses. These features are extracted every minute which corresponds to the environmental data sampling rate. Among the different physiological parameters collected, only the ECG is considered, since it is the parameter commonly related to thermal comfort assessment in literature. Given this assumption, in this research a first analysis is conducted to examine the correlation between the trend of the LF/HF (see Section 2.5) with the trend of some environmental parameters (air temperature, CO2 concentration, relative humidity, mean radiant temperature) and the standard comfort indicator PMV.

The results of this analysis that will be shown in Section 4.2 suggested that environmental parameters are not enough to interpret human thermal comfort perception from a physiological point of view. Other drivers affect LF/HF, apart from environmental quantities, maybe linked to personal characteristics of the single subject. This behaviour suggests that the non-linearity of the problem should be treated with more complex and high-level algorithms. For this reason, an AI approach, focused on ML classification is introduced excluding environmental parameters and using only physiological quantities to predict the thermal sensation vote of the participants. This approach is applied to determine with a certain degree of accuracy the thermal sensation of the user that reflects whether user experiences comfort or discomfort without considering environmental quantities.

3.1.4 ECG signal analysis and validation procedure

The goodness of the proposed methodology is evaluated processing the ECG signal, the derived HRV and its related features, including LF/HF ratio, in relation to the TSV expressed by the participant during the test. A preliminary analysis of the subject's replies has been performed to exclude data potentially affected by biases. The main criteria used for excluding data are: i) presence of non-coherent replies (e.g. warm sensation and warmer preference at the same time) or ii) complaint expressed by the participant because of prolonged sitting position. In this way, a dataset, where discomfort experienced by users can be correlated mainly to thermal dissatisfaction, has been created. Therefore, 29 out of the 62 participants are included in this analysis. This procedure should ensure that the main drivers that contribute to a perturbation in the thermal comfort of the participant, is only the variation in the thermal environment.

For each participant, before extracting HRV, the raw signal is processed according to the procedure shown in Figure 13: initially a mean removal is performed, and the resulting signal is filtered with a bandpass filter [0.8–40 Hz] [58]. The signal passes through a 3rd order high-pass Butterworth filter with cut-off frequency of 0.8 Hz and the 3rd low-pass Butterworth filter with 40 Hz of cut-off frequency in cascade.

The following steps consists of dividing the ECG signal in 5 min consecutive windows and Pant-Tompkins algorithm for extracting the RR intervals (also named normal-to-normal intervals) [59]. This algorithm, recognized in literature as reference methodology, is used to denoise the signal and detect QRS complexes in the ECG signal. A QRS complex indicates the presence of a beat and therefore its detection is useful to compute RR intervals. A bandpass filter is firstly applied to reduce noise, to eliminate movement artifacts, 60 Hz powerline interference and baseline wandering. Then a derivative filter is applied to obtain information about the slope of the QRS complex, followed by the squaring of the signal which highlights better QRS complexes. Finally, signal passes through a moving integrator.

Then, the HRV is derived from the RR intervals measurement. In particular, time-domain measurements include: RMSSD defined as the square root of the mean squared difference of successive R-R intervals; SDANN which is the standard deviation of the R-R intervals; NN50 that represents the number of interval differences of successive R-R intervals greater than 50 ms; and pNN50 which is the ratio between NN50 and the total number of R-R intervals. Fast Fourier Transform (FFT) is implemented to obtain frequency-domain measurements by computing the power spectral density (PSD) of HRV. Three frequency-domain indices are extracted from PSD: the very low frequency (VLF) in the range [0.01–0.04] Hz, the low frequency (LF), and the high frequency (HF) in the range [0.04–0.15] Hz and [0.15–0.40] Hz, respectively. Finally, the computed ratio between LF and HF spectral density provides LF/HF [60].

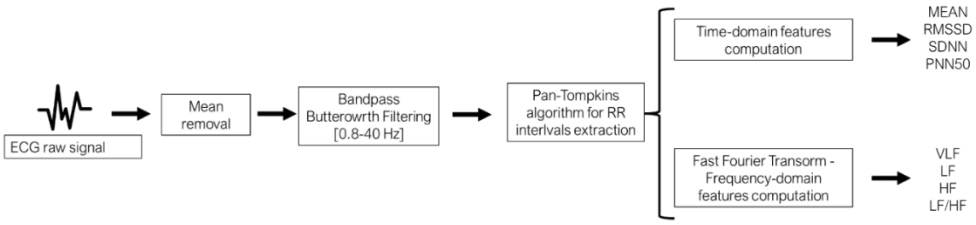


Figure 13. ECG signal processing methodology and HRV features extraction.

3.1.5 Analysis of environmental and physiological quantities

In this paragraph there is the description of the procedure adopted to investigate how LF/HF is influenced by environmental parameters. After the synchronization of the signals, Pearson's linear correlation is applied between LF/HF and each environmental parameter history, e.g. mean radiant temperature (tr), air temperature (ta), relative humidity (RH), PMV, CO2 concentration (CO2).

On this assumption, the correlation analysis conducted in this work first individuates windows of 5 min that slide of 1 min for the environmental and LF/HF signals. For each 5-min-window, the Pearson's coefficient between the environmental parameters and LF/HF is computed. At this point, a vector of Pearson's coefficient is provided for each participant's test and for each environmental quantity. The second step consists in considering the percentage of Pearson's coefficient that are greater or equal to 75% to evaluate which are the most influencing environmental parameters on the LF/HF signals.

3.1.6 Supervised ML analysis

To predict occupants' thermal comfort starting from the HRV and its indices, supervised machine learning (ML) algorithms are here implemented. In this study, six classification algorithms were selected: Linear Discriminant Analysis (LDA), K-nearest neighbours (KNN), Decision Tree (DT), Naïve Bayes (NB), Support Vector Machine (SVM), and Random Forest (RF) classifiers. ML classifiers are applied to four datasets to point out what are the HRV indices that classify with higher accuracy human's thermal sensation vote. The first dataset consists on LF/HF; the second dataset comprises time-frequency indices, the third dataset is made up by frequency-domain indices while the last one includes the entire set of estimated HRV indices. The HRV indices are used to train the ML algorithms to foresee the thermal sensation surveys used as label in the analysis. The datasets are built using the surveys and the related HRV indices of subjects

which have provided a realistic survey and excluding all the thermal perception assessment that never changed during the test.

The classification accuracy (A) of each algorithm is computed using a 5-fold cross-validation, where the dataset is partitioned into 5 randomly chosen subsets (or folds) of equal size. One subset is used by the classifier to validate the trained model using the residual subsets. The procedure is replicated by 5 times, so every subset is adopted only once for the validation. The accuracy of the model is the average accuracy of each fold.

3.1.7 Analysis of subjective parameters in relation to thermal comfort

Individual differences in environmental perception are preliminary analysed by coupling the expressed thermal sensation and the corresponding measured environmental data. In particular, thermal perception expressed by tested subjects in both the seasons, winter and summer, are analysed with respect to measured air temperature.

Thereafter, the dependency of the expressed TSV perception on subjective characteristics is investigated taking into account collected personal information of the subjects, i.e. information given in the first part of the submitted survey. This analysis aims to point out the existing correlations between the characteristics of the participants, assumed as dependent variables, and the perceived sensations expressed through the third part of the survey, i.e. the independent variable. Strength of tested hypothesis is expressed by the probability value, i.e. p-value.

3.2 Measurement procedure and data analysis for assessing thermal comfort in semi-controlled environment

The methodology discussed in the previous Section 3.1 have laid the foundations for the development of the research activity here presented and the tuning of the methodology according to its purpose. The work that will be presented in this chapter is different from the previous one, and in particular:

- Participants involved in the experimental campaign are working in a less controlled environment, closer to real-life settings;
- Participants are free to carry out light office activities and were not constricted to stay on the dedicated workstation without the possibility of moving for more than one hour;
- The experimental setup for the acquisition of physiological quantities was reduced to a minimally-invasive smartwatch that collects HRV;
- In the previous work, the sampling frequency of the environmental parameters was limited to one sample per minute, which made it necessary to compute the HRV features at the same frequency with a smaller number of samples. In this case, the methodology

used to extract the HRV features was revised by computing a greater number of samples, which in turn positively affects the ML computations;
- The dataset used to train the ML algorithms was modified, since it includes non-linear features of HRV and a different combination of time and frequency-domain parameters that has proved to provide better performance;
- The AI algorithms adopted for computing the thermal sensation vote are used for regression rather than classification, since TSV is manipulated differently.

Given the above assumptions, this second research activity developed within this PhD research, aims at demonstrating how the inclusion of some physiological parameters related to human thermal comfort measured through a minimally invasive smartwatch can help to improve the prediction of occupants' TSV vote, in semi-controlled environment. In particular, the methodology adopted aims at exposing each occupant to different environmental conditions generated by a variation in air temperature, relative humidity and air velocity while collecting physiological and environmental parameters.

3.2.1 Material and Methods

This section describes the experimental procedure developed to analyse human thermal perception in response to different external thermal stimuli. The test took place in KUBIK (Bilbao, Spain), Figure 14, which is a full-scale experimental infrastructure where research and development activities for energy efficiency purposes are carried out. The facility is mainly used to develop and test new solutions that could reduce energy consumption in buildings and, at the same time, improve the thermal comfort experience of occupants in indoor environments. The building's main functionality is its capability of creating realistic conditions for the purpose of analyzing energy efficiency, thanks to the intelligent management of its HVAC and lighting systems. A complete description of the structure is provided in [61]. The experimental procedure developed for the purpose of this study was performed on the first floor of the building; the planimetry of the floor is reported in Figure 15.

Figure 14. External appearance of the Kubik facility**.**

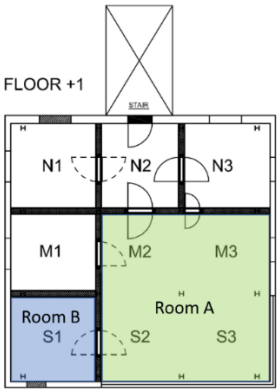


Figure 15. Planimetry of the floor in the facility where the tests were performed.

Respect to the previous work described in Section 3.1, more than one experiment was conducted: three different trials were conducted, all of which aims at exploring human thermal response to warm induced comfort, cold induced discomfort and transient environment conditions. The first trial aimed at creating cold-induced discomfort (Experiment 1), the second created warm-induced discomfort (Experiment 2), while the third test was performed by generating a transient temperature variation over time (Experiment 3). The three tests, which are described in detail in the following sections, were built in a similar way: they all included a first stage of acclimatization performed in Room A and a second stage conducted in Room B in which room temperature, relative humidity and air velocity varied. Room A was equipped with a dedicated workstation consisting of a table and chair to simulate light office activities. Room A was always kept at a neutral temperature according to users' sensation. The average indoor temperature recorded in room A was between 19°C and 21°C and the set-point of the HVAC system was 20°C; the test performed in room A was necessary to collect the baseline analysis before the user is exposed to a discomfort condition in room B. To this purpose, since the ground truth is the TSV, the test in room A started only if the user was in comfort or not. Room B was also equipped with a desk and chair to simulate light office activities. The temperature in Room B was set by an HVAC system controlled remotely from outside the room. Test Room B, Figure 16 , has three outdoor exposed elements, i.e. rooftop, west and south façades.

Figure 16. Picture of room B used to create a discomfort condition for the participants.

3.2.2 Participants

Ten volunteer participants (5 female and 5 male subjects) were recruited for Experiment 1 and Experiment 2, while 13 participants (7 female and 6 male subjects) were recruited for Experiment 3. The anthropometric information of the participants is summarized in Table 2. The participants were required to wear their everyday clothes. The reason behind this requirement is that people have a different thermal sensation when wearing different clothes, therefore, this condition helps to preserve the subjectivity of the test, because participants' perception is not subjected to a bias. The clothing thermal insulation I_{cl} (Table 2) was computed indirectly, by summation of the partial insulation values for each item worn by participants. The set of garments worn by participants was collected and the procedure suggested by ISO 9920 was adopted to evaluate the I_{cl} [62].All the participants gave written informed consent to use their personal data and were duly informed about the goal of the research.

3.2.3 Experimental set-up

Table 3 illustrates the environmental parameters recorded and the characteristics of the sensors installed in Room B for monitoring purposes. The measurement set-up of the sensors in Room B is displayed in Figure 17. ISO 7726 Standard was adopted to assess the thermal environment of Room B. Air temperature was measured by four thermocouples placed in different positions of Room B. The thermocouples (t_2, t_3, t_4) were mounted on the perimetral walls of the room, at 1.4 m from the ground (Figure 17). Thermocouple t_1 was instead positioned close to the

workstation the participants were sitting at during the experiment to record air temperature. Actually, three thermocouples were positioned close to the workstation at different heights (0.1 m, 0.6 m and 1.1m). However, for the purpose of this study, only the air temperature (t_a) measured at height 1.1m (t_1) was considered [49]. The anemometer was placed at 1.25 m from the window and 1.10 m from the ground; the relative humidity (*RH*) sensor was placed at 0.7 m from the floor and 1.25 m from the window. The globe thermometer was placed close to the workstation, 1.10 m from the floor. Each sensor collected data every 5 seconds. All the data collected were sent to an acquisition board that saves data locally.

Table 2. Anthropometric information of the participants involved in the tests and the experiments attended.

User	Gender	Age	Height (cm)	Weight (kg)	BMI	I_{cl} (clo)	Race	Experiment attended
1	F	26	165	64	24	0.80	Eu	1,2,3
2	F	33	179	70	22	0.70	Eu	1,2,3
3	M	29	187	108	31	0.67	Eu	1,2,3
4	M	35	173	70	23	0.62	Eu	1,2,3
5	F	25	168	57	20	0.76	Eu	1,2,3
6	M	27	178	72	23	0.72	Eu	1,2
7	M	37	172	68	23	0.67	Eu	1,2
8	M	39	189	75	21	0.73	Eu	1,2
9	F	32	161	56	22	0.67	Eu	1,2
10	F	47	156	59	24	0.73	Eu	1,2
11	M	27	190	88	24	0.66	Eu	3
12	M	31	186	75	22	0.83	Eu	3
13	F	24	178	56	18	0.80	Eu	3
14	M	32	183	90	27	0.83	Eu	3
15	F	27	165	62	23	0.60	Eu	3
16	M	27	173	75	25	0.68	Eu	3
17	F	34	179	67	21	0.85	Eu	3
18	F	28	175	63	21	0.91	Eu	3
Mean		31	175	71	23	0.74		
STD		6	10	13	3	0.09		

During the test, the participants were provided with a smartwatch that continuously recorded their HRV signal. The smartwatch model was a Samsung Galaxy Watch and the participants were asked to wear it on their non-dominant wrist. The physiological parameters collected were saved locally on the internal storage of the smartwatch and simultaneously sent to a smartphone via Bluetooth communication. For the purpose of this study, a dedicated Javascript application was implemented. The App collected each HRV sample in real-time and displayed it on a chart that was uploaded in real-time. This application was useful for verifying the correct functioning

of the smartwatch during the acquisition and saving an additional copy of the HRV signals on the internal memory of the smartphone.

Table 3. Technical information of the sensors employed to collect environmental quantities inside the test room B.

Sensor	Manufacturer	Model	Accuracy
Air Temperature	Thermo Sensor GmbH	PT100 (4 wired) 2113-1-074	± 0,1 ºC
Relative humidity	Ahlborn	FHAD46C41A	±2% of reading
Air velocity	Ahlborn	FVAD05TOK300	±1% of reading
Globe temperature	Ahlborn	FPA805GTS	± 0,1 ºC

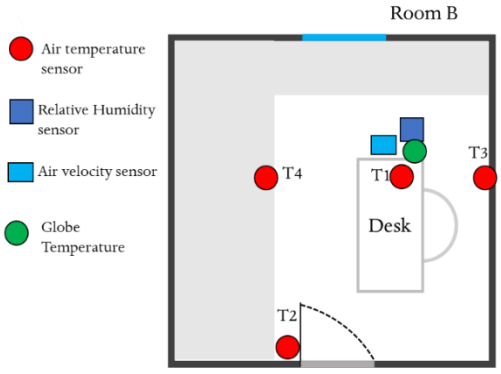
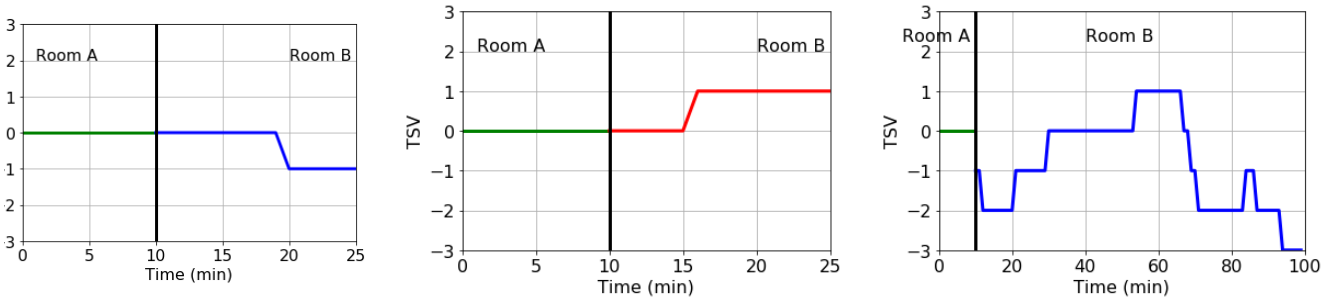


Figure 17. Measurement set-up in Room B including furniture and sensor location.

3.2.4 Thermal sensation vote

During the three experiments the participants were asked to express their thermal sensation vote (TSV) whenever they experienced a different TSV from the previous instant. The TSV expresses an occupant's thermal sensation; in this specific context, the ASHRAE 7-point scale was used, which is based on the measurement of how warm or cool an occupant feels. For this reason, during the different experiments, the participants were required to express a vote from -3 to +3, according to their thermal sensation. Each vote represented a particular sensation: cold (-3), cool (-2), slightly cool (-1), neutral (0), slightly warm (+1), warm (+2), hot (+3). TSV was not

collected with a specific frequency during the test, but it was expressed by the participants whenever there was a change in their TSV. Figure 18 (a-c) shows an example of the TSV collected in each experiment for a single participant.



(a) (b) (c)

Figure 18. (a) TSV trend during Experiment 1 for one participant. (b) TSV trend during Experiment 2 for one participant. (c) TSV trend during Experiment 3 for one participant.

3.2.5 Experiment description

The three experiments were performed on three consecutive days both in January 2020 and January 2021. During the three experiments the participants were sitting at a workstation and could perform light office activities (e.g. reading, working on the laptop). The tests were built as follows (**Table** 4):

Cold-induced discomfort (Experiment 1). Each participant started the trial outside the test-room, in a thermally comfortable environment (Room A). When the participant claimed to be thermally comfortable (which means that the TSV was equal to 0), the test started and lasted for 10 minutes. This initial part was necessary to collect the participant's baseline signal. The participant then moved inside the test-room (Room B), whose set-point temperature was set at 15°C and the window was kept open to vary air speed and relative humidity. The test in Room B lasted for 15 minutes.

Warm-induced discomfort (Experiment 2). Similarly to Experiment 1, each participant started the trial outside the test-room, in a thermally comfortable environment and the baseline signal of HRV was collected. When the participant claimed to be thermally comfortable (which means that the TSV was equal to 0), the test started and lasted for 10 minutes. This initial part was necessary to collect the participant's baseline signal. The participant then moved inside the test-room (Room B), whose set-point temperature was set at 26 °C. The test in room B lasted for 15 minutes. In this case the window was kept closed, but a fan system was used to air the environment and facilitate the diffusion of heat around the room.

Transient air temperature (Experiment 3). This experiment was different from the previous ones, because air temperature in Room B was not kept constant but varied over time. The profile

temperature was created as follows: Room B was previously set at 15°C for 5 minutes, then temperature set-point was set to heat the room up to 26°C for 5 minutes, then temperature was set back to 15°C. The temperature profiles created for each of the three experiments are showed in Fig. 5. This range of temperature variation, from +15°C to 26°C, was identified in line with a previous work that considered a variation from +20°C to 30°C [49]. The lower temperature was further decreased to induce a colder sensation environment compared to a comfort condition during winter season. In addition, the tests performed in the previous work were executed with a total duration of 40 minutes, in this work the total duration was lengthened up to 100 minutes.

The three experiments were conducted in a silent and quiet room, isolated from external stimuli that could act as an interference in the measurement of HRV. This procedure should guarantee that the greatest perturbation to the participants' condition is due to the variation in indoor temperature. In order to avoid possible motion artifacts, participants were instructed to limit wrist movements as much as possible and were advised to keep the hand of the arm with the smartwatch in a specified position indicated by a visual sign placed on the desk. Before starting each experiment, it was verified whether the participants were bothered by the sensor equipment and whether they felt comfortable in maintaining the required position at the desk.

Table 4. Description of the three experiments conducted.

	Room	Duration	Indoor Temperature
Experiment 1	A	10 min	Neutral (19 - 21 °C)
	B	15 min	Cold (15°C)
Experiment 2	A	10 min	Neutral (19 - 21 °C)
	B	15 min	Cold (26°C)
Experiment 3	A	10 min	Neutral (19 - 21 °C)
	B	5 min	15°C
			up to 26°C
		5 min	26 °C
			down to 15°C

To evaluate and assess the thermal environment, the specifications and methods of ISO 7726 Standard were followed. To evaluate the horizontal homogeneity of the air temperature, according to ISO 7726, the deviation between each air temperature value measured in one point and the mean value was compared with the multiplication of the required measure accuracy by the appropriate X factor equal to 4. The procedure was replicated throughout Experiment 3 in three different instants of time (t_1, t_2, t_3). The deviations are reported in Table 5. Figure 19 reports the trends of the environmental parameters measured using the configuration of Figure 17.

Each participant in Experiment 3 was exposed to the same environmental conditions, as shown in Figure 19, which display the average trend of each parameter during Experiment 3 and the maximum standard deviations in three different instants of time. Figure 19 also reports the average trend of the mean radiant temperature (t_r) and the operative temperature (t_{op}), computed from the formulas listed in ISO 7726.

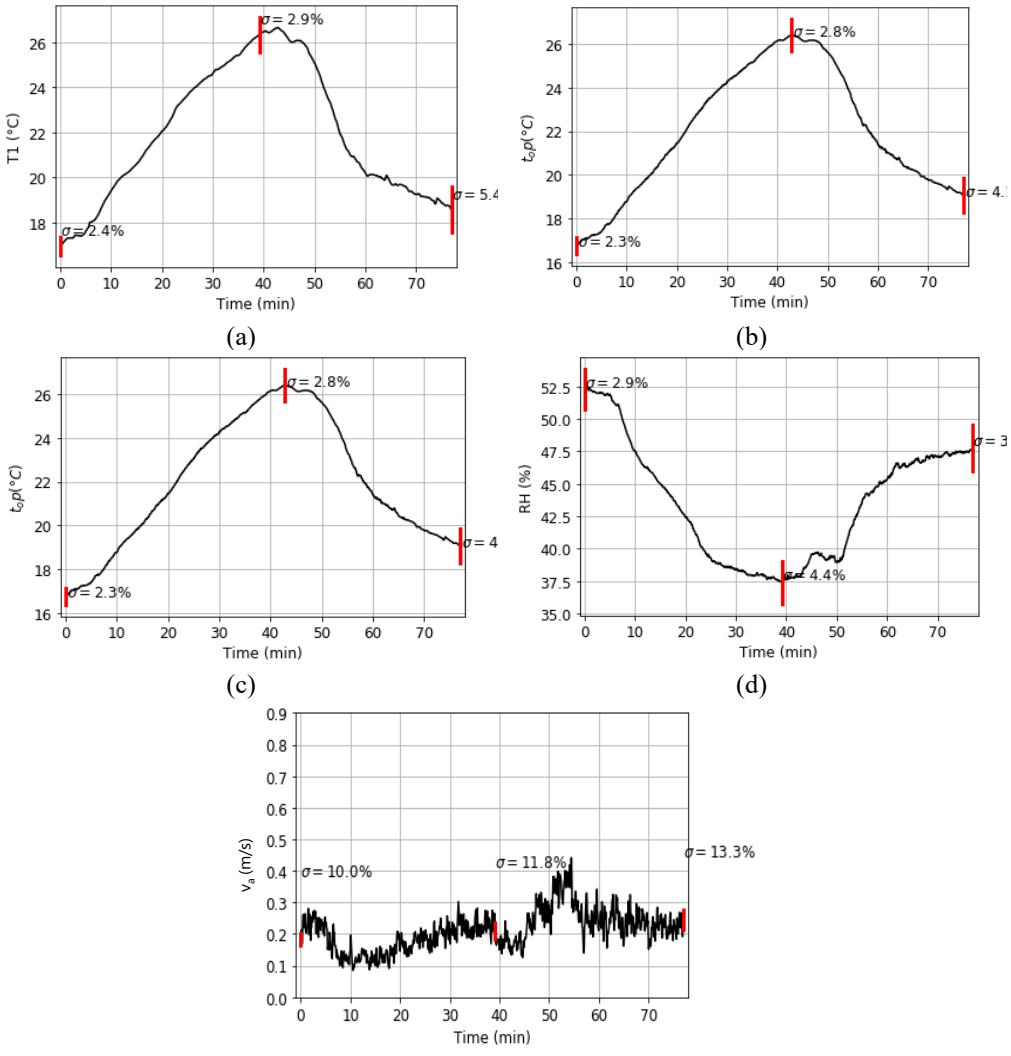


Figure 19. Average profile for each environmental parameter, with the maximum deviation computed among all participants. (a) Air temperature in point T1. (b) Operative temperature. (c) Mean radiant temperature. (d) Relative humidity. (e) Air velocity.

Table 5. Deviations of air temperature at the different locations in test room B at different time instants.

Homogeneity of the environment				
		(°C)		
time	T1	T2	T3	T4
t1	-1.1	-0.5	-1.2	-1.2
t2	-2.1	-0.5	-2.6	-0.1
t3	-2.2	-0.4	-1.9	-1.4

3.2.6 Data processing

This paragraph contains a detailed explanation of how each signal (physiological, environmental and personal) collected during the three experiments, was analyzed. For each participant, the environmental data, physiological data and the TSV collected during the tests were processed. Data processing was an essential step for the extraction of the proper set of features necessary to predict TSV.

3.2.6.1 HRV analysis

HRV signal extracted from the smartwatch undergoes a process of filtering and outliers' removal, since PPG-derived HRV can be deeply influenced by artifacts due to small movements of the arm or loose adherence of the smartwatch to the skin surface, which might generate unwanted sources of external light that negatively affect the PPG signal. In [34], which is a research activity conducted in the context of this PhD thesis, a preliminary experiment to compute HRV uncertainty in resting conditions was conducted. The uncertainty in the monitoring of HRV through the Samsung Galaxy smartwatch compared with a multi-parametric belt, BioHarness 3.0 was ±0.95%. Outliers were detected using a straightforward thresholding methodology consisting in the comparison of the actual HRV sample with the previous one and considering it an outlier if it differed from the previous value by more than 50% of the mean value of a time-window of one minute. The outliers detected were replaced with the previous uncorrupted interval [13], [25].

3.2.6.2 HRV windowing and indices computation

Once the outliers were detected, the resulting HRV signal was divided into time frames built as follows: the first time frame corresponded to 5 minutes of the HRV signal, since this duration is the minimum time recommended for short-term HRV series to compute the spectral analysis

[26]. After the extraction of the first window, a new window was computed by appending a new HRV sample interval of the signal, while the oldest sample was removed from the beginning of the window. The process was repeated until the end of the signal [63]. According to literature [64], several HRV features to be extracted from the HRV signal were identified. Time-domain HRV features ($f(HRV_t)$) are a collection of statistical and geometrical indices for the measurement of the variability in the HRV sequence that act as indices to interpret the oscillations of cardiac cycles. The ($f(HRV_t)$) statistical indices computed in this study are SDANN, RMSSD, MEAN, MEDIAN, PNN50, PNN25. In addition, HRV studies imply the use of frequency-domain features ($f(HRV_f)$), which are useful for the understanding of the stationarity or stability of the HRV signal. To obtain the frequency-domain analysis, the first power spectral density (PSD) was computed through the autoregression modeling-based method that has proven to provide better resolution.

Table 6. Description of the HRV features extracted.

Domain	**Time Domain *f(HRVt)***						**Frequency Domain $f(HRV_f)$**					**Non-Linear $f(HRV_{nl})$**		
Name	**SDANN**	**RMSSD**	**MEAN**	**MEDIAN**	**PNN50**	**PNN25**	**LF**	**HF**	**LF/HF**	**HF/LF**	**TP**	**SD1**	**SD2**	**SD1*SD2**
Description	Standard deviation of the HRV series	Root mean - squared error of the HRV series	Mean value of HRV	Median value of HRV	% of adjacent HRV samples differing by more than 50 ms	% of adjacent HRV samples differing by more than 25 ms	Low frequency band of the HRV power spectrum	High frequency band of the HRV power spectrum	Ratio of LF to HF	Ratio of HF to LF	Total Power of the HRV power spectrum	Poincarè plot index of the short-term HRV	Poincarè plot index of the long-term HRV	Combination of SD1 and SD2

Each frequency band was then computed: LF (0.04-0.15 Hz) and HF (0.15 - 0.4 Hz), LF/HF, HF/LF and the total power spectrum (TP). Non-linear features (f(HRV_{nl})) were also computed through the Poincarè plot. The Poincarè plot is a graphical representation of an HRV time series along the cartesian plane: the X-axis contains one HRV sample, while the Y-axis contains the following HRV sample. The Poincarè plot provides two additional features obtained by adjusting the point cloud of the figure formed into an ellipse, obtaining SD1 and SD2. The list of all the HRV features computed are shown in Table 6.

3.2.6.3 Thermal sensation vote processing

As previously written, during the experiment TSV was not collected with a predetermined frequency. The participants, in fact, communicated their vote as soon as they perceived a

different sensation with respect to the previous one. Therefore, the TSV vector was built with one sample per minute. Each TSV sample used to perform the analysis was obtained by building TSV windows in the same time interval of the corresponding HRV window, and the final value was obtained by averaging the TSV window.

3.2.6.4 Binary classification between comfort and discomfort

Three ML classification algorithms, selected from literature, were used to predict human thermal comfort expressed through TSV. The algorithms are Support Vector Machine (SVM), Random Forest (RF) and the Extra Tree Classifier (ETC). Experiment 1 and Experiment 2 were built in order to investigate whether it was possible to distinguish "Comfort", which results in TSV equal to 0, and "Discomfort", expressed in terms of TSV with the remaining values of the ranking different from 0. In Experiment 1 the participants were exposed to a neutral environment inside Room A and immediately after to a cold environment inside Room B. In Experiment 2, the same procedure was applied, but in this case Room B generated hot-induced discomfort. The level of activity of the participants during the experiment, the clothing insulation and the experimental set-up used were the same in both rooms, therefore, the only variable that changed over time was the air temperature, which in turn generated a different thermal sensation vote. This part of the study was therefore focused on the ability of an ML classifier to distinguish between comfort and discomfort of the participants starting from the LF/HF, HF and LF extracted from HRV, which are connected to the thermoregulation mechanism. The HRV values were normalized to include robustness to very small standard deviations of the features and preserve zero entries in sparse data. In this context, the TSV that has to be predicted was manipulated in order to have a binary classification as follows: in both Experiment 1 and Experiment 2 every TSV different from 0 collected was categorized as "Discomfort" while every TSV equal to 0 was categorized as "Comfort". One participant in Experiment 1 was excluded from the analysis because in both Room A and Room B the vote expressed was 0, so the participant did not experience any discomfort. 6 participants were considered for the analysis of Experiment 2, because the resulting TSV in Room B for 4 participants was always 0. To avoid overfitting of the three classifiers, the validation of the model was conducted by performing a k-fold cross-validation using 10 folds. The metrics for evaluating the performance of the classifiers is the accuracy (A) computed according to Equation 1:

$$A = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

where *TP* is the True Positives, *TN* is the True Negatives, *FP* is the False Positives and *FN* the False Negatives.

3.2.6.5 Prediction of thermal discomfort in transient environment in Experiment 3

Experiment 3 was conducted to investigate the human body response under transient environmental conditions obtained with small step changes in air temperature profiles. The temperature profile created in this experiment differs from the ones in Experiment 1 and 2, since it is generated by creating small variations of air temperature over time, which is what typically happens in office buildings. The complex mechanism of thermoregulation results in a strong non-linearity of the features extracted from HRV and this is the reason why ML techniques can provide support to find relationships between non-linear variables. The data obtained from Experiment 3 were processed in order to create 2 different datasets: the first dataset ($\mathbf{F_H}$) was built with HRV variability features (*f(HRVt)*, *f(HRVf)* and *f(HRVnl)*; the second dataset ($\mathbf{F_{H+E}}$) was built by joining physiological and environmental quantities. Before using ML, the t_a, *RH* and *v* recorded in Room B needed to be processed. This procedure was essential because Experiment 3 implied a transient air temperature that varied continuously during the test and, to proceed with the analysis, it was necessary to associate one sample of each parameter to each time window of HRV. For this reason, windowing was applied to each environmental quantity by delimiting each window with the corresponding time interval that delimits an HRV window. From each environmental window, the mean value was computed. This procedure can be explained by the fact that 5 minutes of HRV, and in particular of LF/HF, are strictly connected to the variation of the environmental quantities in the same time interval [65]. Five different regression ML and DL algorithms were used to predict the TSV of each participant in the transient environment: Support Vector Machine (SVM), Random Forest (RF), Multi-Layer Perceptron (MLP), 1-dimensional CNN and LSTM. All the algorithms were trained and tested using Python libraries keras and sklearn.

An MLP artificial neural network is made of an input layer, one or more hidden layers of artificial neurons and an output classification layer. The network is particularly indicated for nonlinear problems, because of its ability to learn new relationships between parameters by updating the weights of the connections between the neurons of consecutive layer [66]. For the purpose of this research, MLP was trained with a number of hidden neurons set to 8, which is strictly connected with the size of the training data dimension. The regularization parameter alpha (α) was determined using the grid search algorithm, whose function is to take a set of possible values of the chosen parameters and find the best combination [64]. The grid search algorithm provided $\alpha = 0.01$.

The LSTM model was created as follows: the network weights were optimized by minimizing the loss function using the "ADAM" optimizer. The LSTM hidden layer was made by 100 dimensions, a dropout layer to reduce overfitting, followed by a fully connected dense ReLu layer with 100 outputs and a final output layer to obtain one output for the regression. The model was trained for 15 epochs with a batch size of 50.

The architecture of the 1D-CNN proposed included 3 Convolutional layers, one pooling layer and an output layer that returns a single value to predict TSV. Each convolutional layer had 64, 32 and 16 layers, with the ReLu activation function and was followed by a dropout layer for regularization. After the convolutional layers and the pooling, there was a flatten layer that transformed the features learned into a one-dimensional vector that was then sent to a fully connected layer to make regression predictions.

The validation of each algorithms on the two datasets was performed by using the Leave-One-Subject-Out (LOSO) procedure, which consists in training the algorithm on all the subjects except one, which is used for the validation [27]. The metrics for the validation of regression algorithms were expressed in terms of Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE), computed as follows in Equation 2 and Equation 3.

$$MAPE = \frac{1}{n}\left(\sum_{i=1}^{n} \left| \frac{y_i - x_i}{y_i} \right|\right) \quad (2)$$

$$MAE = \frac{\left(\sum_{i=1}^{n} |y_i - x_i|\right)}{n} \quad (3)$$

Where n is the number of observations, y_i is the actual TSV and x_i is the predicted TSV obtained from the ML model.

3.2.7 Impact of the measurement uncertainty on the monitoring of thermal comfort through AI predictive algorithms

One of the main aspects that come up from the study in Section 3.1 and Section 3.2 is that the human centric approach for measuring thermal comfort is based on the measurement of a heterogeneous set of data, (i.e., environmental, physiological and personal parameters such as the TSV). Physiological quantities are represented by HRV and its derived measurements; environmental quantities comprehend air temperature, air velocity or relative humidity. Personal information is instead represented by the TSV.

Physiological, psychological and personal parameters are progressively included in the measurement setup of thermal comfort, to encourage a more customized environment carefully tailored to the distinct preferences of the occupants that live in it. As it was possible to appreciate is study of Section 3.1 and Section 3.2, personalized thermal comfort measurement typically requires an extensive range of sensing devices that make up a sensors network (e.g., thermocouples, anemometer, smartwatch). Since each sensor is properly characterized by

measurement uncertainty, it is significant to include in the measurement process how this uncertainty, associated with each parameter, affects the final measurement of thermal comfort.

This aspect is even more justified since, in all research fields, it is routinely accepted by the scientific community that the result of a measurement process promptly loses its meaning if an uncertainty value is not associated with it [67]. The traditional concept of measurement uncertainty must therefore be associated and applied to modern techniques, as in the case of artificial intelligence (AI) algorithms, which are finding more and more space in the measurement process. In support of this claim, literature reports that when small variations are applied to data, and variations are represented by the measurement uncertainty associated with the data collected, AI provides completely misleading results [68].

Moreover, the personalization of thermal comfort by means of heterogeneous dataset, is making necessary the employment of AI algorithms, such as ML or DL algorithms. The reason behind the use of AI for personalized thermal comfort measurement is that the measured output is dependent on several parameters, i.e., physiological and environmental parameters, that compose the measurement system.

The application of AI in the measurement of human thermal comfort in the built environment, is becoming preferable because the adopted dataset is composed of two categories of parameters, which comprise the environmental and physiological parameters. A third category is also added to this dataset, made up of subjective parameters, expressed by the TSV. The union of these quantities generates a complex and heterogeneous dataset, making necessary the employment of AI, which comprises models comparable to a black box [69]. There are strong non-linear relationships among the parameters included in this heterogeneous dataset, which suggests that this relationship should be explored with more complex and high-level algorithms, included in the AI field [65], [70].

AI measures its performance in terms of accuracy, which is expressed through different metrics, depending on the type of algorithm adopted, and the type of data being measured; for example, when the measured quantity is binary, or is made by discrete values, or class, performances are measured using accuracy, recall or precision; in case of AI used for measuring continuous quantities, the most used metrics are the mean absolute error (MAE), mean absolute percentage error (MAPE), and mean squared error (MSE). Of course, the accuracy of AI-based models, is strictly linked to the selected algorithm as well as the quality of the dataset, that is deeply connected to the uncertainty of the measured data [69], [71].

The method for the estimation of the uncertainty is described in the Guide to the expression of uncertainty in measurement (GUM) [72], which is based on the law of propagation of uncertainties (LPU). Given the strong non-linearity of the relationship that exists between the input quantities and the output quantities of AI algorithms, the evaluation of the uncertainty associated to the output quantity is hard to assess if the LPU should be used. When it comes to non-linear relationship, the GUM has proposed in its supplement [12], the evaluation of the

uncertainty by using the Monte Carlo method (MCM), which allows to assess the uncertainty even if the relationship among quantities is not linear, and the analytical equation between quantities is not known, but it can be assumed as black box, as happens in the context of AI models. The MCM represents a more practical alternative of the conventional LPU method, when it is not possible to effectively verify the hypothesis assumed by the LPU [71]. For this reason, this work studies the impact of the measurement uncertainty of a sensors network to measure thermal comfort, using AI models. To quantify and access the factors that affect the outcome of a measurement process through MCM, there are two methodologies that can provide this information: a sensitivity analysis and a measurement uncertainty analysis [73]–[76]. Both of them are now applied to AI models, the first one to analyze how much each uncertainty weights on the measurement of the model's outcome and the second to identify and quantify what are the main sources of uncertainties in the measurement [77]. This second part of the study deals with a procedure to evaluate the impact of the uncertainty of the input data on the thermal sensation measurement output, using the AI algorithm mentioned in Section 3.2.6. The study is structured as following:

a) First of all, MCM is applied to the raw HRV signals coming from different participants, to see how the uncertainty in the measurement of HRV impacts the computation of the HRV features, since they are fed into AI algorithms, to extract the TSV. The standard uncertainty, associated to the measurement of HRV, chosen for this analysis is ± 4 ms, computed from a previous study in which the smartwatch was compared to a reference method while participants were sitting at rest [24]. This preliminary analysis can be useful to establish if measurement uncertainty of the smartwatch, used to assess HRV and derive its features, has the same impact on each participant's data, by having similar measuring uncertainty for each

b) Secondly, MCM is applied to both physiological and environmental parameters used as input to the AI algorithms (RF and CNN), to evaluate how the measurement uncertainty propagates to the output, in black box methodology such as AI algorithms. The uncertainty used for the environmental quantities are the standard uncertainty provided from the datasheet of each device, while HRV is perturbed with different values of uncertainty that ranges from ± 4 ms to ± 100 ms. This range of values is chosen because literature has highlighted that, depending on the activity level performed by participants is wearing the smartwatch and it is completely in resting conditions, the uncertainty associated to the measurement of the HRV is ± 4 ms, while is greater than or equal to 100 ms when the user is performing a motion test [34], [35]. The research described in [34] and [35] were necessary since commercial smartwatches are not provided with precise datasheets that define the measurement uncertainties of the measured quantities. This is one of the problems being encountered in literature when it comes to working with low-cost sensors and often sensors that are not designed for research purposes. The results are used to

perform a sensibility analysis related to ± 4 ms of uncertainty, to examine the contribution of the uncertainties associated to the TSV measurement, in relation to the uncertainties of the input parameters.

The methodology described in Figure 20 is adopted: the first assumption is that each data coming from a measurement process is characterized by two quantities which are the data itself and the associated uncertainty. Collected data from sensors network are merged together to build the heterogeneous dataset, that comprises the set of data and the associated uncertainties; therefore, when AI is applied in the measurement process, the model takes as input also the uncertainty of the collected data. Therefore, it is expected that the results of the AI model derive from two main aspects: the first one is the intrinsic structure of the algorithm, while the second one is uncertainty associated with the input data. The evaluation of the impact of the uncertainty associated to the input data in the measurement output, when AI is applied, is a paramount result. There is one intermediate aspect that should be considered, which is the impact of the uncertainty of the AI models, that will be combined with the uncertainty of the measurement instrument. The last part of Figure 20 deals with two types of analysis which are the sensitivity analysis (SA) and the uncertainty analysis, performed by MCM.

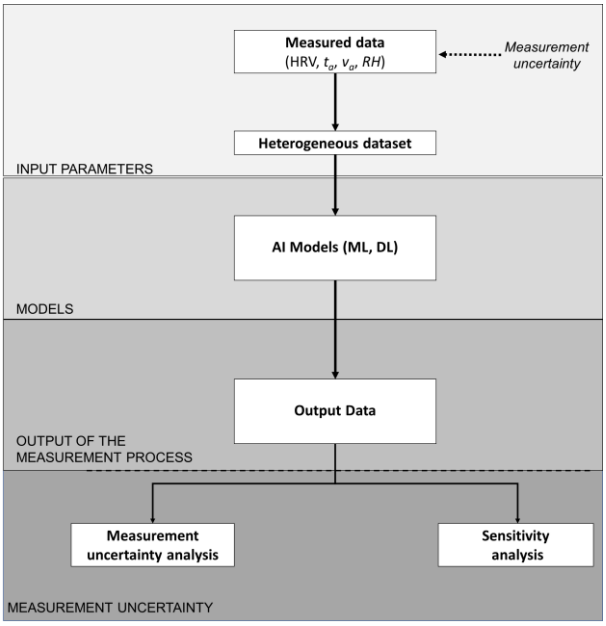


Figure 20. Conceptual description of the procedure adopted to study the impact of the measurement uncertainty in the context of AI model.

The procedure of the MCM consists in using data, acquired through real experiments, and perturbing them by assigning different measurement uncertainties or perturbation. Each input parameter is modified, by adjusting it with a different perturbation, one at a time, while the other

input variables are kept unchanged. Based on the obtained results, the effect of different measurement uncertainties on the prediction of the TSV is observed through the simulation of perturbed data. The described methodology summarizes the steps to evaluate the measurement uncertainty of different parameters, used as input variables (or predictors) in AI algorithms, applied for the measurement of personalized thermal comfort, expressed through the TSV. The GUM supplement provides the step necessary to perform the MCM. From a general point of view, MCM provides a general approach to numerically approximate the *cumulative density function (cdf)* of the output of a certain quantity $y = f(x)$. The main concept behind MCM is that every sample of the input quantity x_i, chosen from a predetermined distribution can be used. In this way, by taking a random sample of each input $x_{i,}$ from its related probability distribution function (*pdf*), it is possible to estimate a possible result of the output y and the associated uncertainty.

To explain how MCM is adapted in the context of this research, the following steps were applied:

1) The number of trials *M* of Monte Carlo trials is set to 200,000. M is the number of output quantity values that needs to be selected; in this study, it is chosen a priori. Usually, a number of trials equal to 10^6 are considered, which is the number of trials that should provide a coverage interval of 95%, as reported in the GUM supplement [78]. Since it is commonly accepted that, the higher the number of trials, the higher is expected the convergence of the results, the number of Monte Carlo trials (M) was set to 200.000, as reported from literature [79];
2) *M* vectors x_i , i = 1, …, *M*, were generated, by selecting randomly from the probability density function (PDF) of each input quantity [x_{HRV}, x_{t_a}, x_{RH},x_{v_a}] in order to realize a set of possible input that can be associated to the input quantity. The random sample is obtained from a gaussian distribution, with uncertainties described in Table 7.
3) For each vector generated in step 2, the corresponding output y (or TSV in this case) is computed, yielding to M vector output quantity values
    a. Point 3 is applied to estimate the uncertainty associate to the measurement of the HRV features;
    b. In addition, point 3 is applied, perturbing the features used as input in the ML and DL models.
5) The representation ***G*** of the distribution function for Y is computed, starting from the set of M output of Y;
6) G is used to compute an estimate y of Y and the covariance matrix u_i associated with y;
7) G is used to compute the appropriate coverage region for Y, for a stipulated coverage probability p;

Table 7. Characteristics of the measurement instrument used in the test.

Input Quantities	Manufacturer	Model	Type	Distribution	Standard Uncertainty
Air temperature (t_a)	Thermo Sensor GmbH	PT100 (4 wired)	B	Normal	± 0,1 ºC
Relative humidity (RH)	Ahlborn	FHAD46C41A	B	Normal	±2% of reading
Air velocity (v_a)	Ahlborn	FVAD05TOK300	B	Normal	±3% of reading
Heart rate variability (HRV)	Samsung	Samsung Galaxy Watch	B	Normal	± 4 ms

In the framework of this research, there will be reference to a generical model, with a number i of input x_i and one output y, which is the measured quantity. The input parameters used in this study to measure TSV, are referred to the experimental campaign presented in Section 3.2, in particular with reference to Experiment 3, in which participants were exposed to a transient variation of the indoor environmental quantities, which is the closer condition to real-life settings.

3.2.7.1 Monte Carlo approach on HRV features

In this first part of the analysis, MCM is used to estimate how the uncertainty of the device through which the HRV is collected, influences the computation of the HRV features, that will be lately used to evaluate human thermal comfort. HRV signal, which is made by the distance in time of two subsequent RR peaks of the ECG trace, is divided into time frames from which it is possible to extract some indices (or features). Each time frame is built as follows: the first time frame corresponded to 5 minutes of the HRV signal, which is considered the minimum time duration recommended for computing HRV spectral analysis. After the extraction of the first window, a new window was computed by appending a new HRV sample interval of the signal, while the oldest sample was removed from the beginning of the window; the process was repeated until the end of the signal. HRV features included in this study were already described in Section 3.2.6.1, Table 6.

In this section, there is a first analysis made specifically on the individual features of the HRV. In practice, we want to assess how it is the uncertainty measurement of each feature, when the smartwatch is affected by a specific uncertainty in the measurement of HRV; the procedure is summarized in Figure 21. First of all, one HRV window of 5 minutes is considered. This window will be perturbed with a random sample coming from a normal distribution with mean equal to 0, and standard deviation equal to (U), for 200,000 iterations. The perturbed HRV segment will be used to compute HRV features in order to establish the impact of the uncertainty. The final uncertainty is computed as the standard deviation of the resulting HRV, by a coverage factor k = 2. For the simulation study, 13 segments (one for each participant to Experiment 3) of HRV of a duration of 5-minutes, were extracted from the set of experimental procedure described

previously. The duration of 5 minutes is required because is the minimum time required to compute short term HRV features.

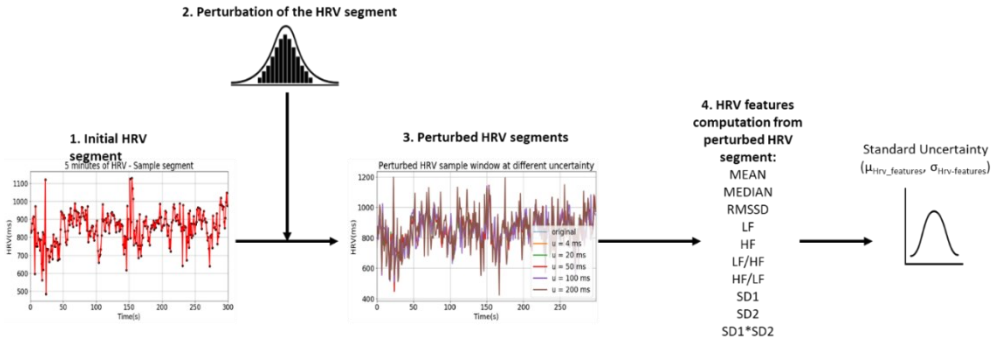


Figure 21. Description of the MCM adopted to evaluate the impact of the smartwatch uncertainty on the HRV features that will be used to measure and assess human thermal comfort.

3.2.7.2 Monte Carlo approach on AI models for measuring human thermal comfort

For this second part of the analysis, the set of HRV features, explained in Figure 22, was chosen [80]. Figure 22 also shows the approach used to apply MCM and AI: a set of input parameters was used to train RF regression algorithm and a CNN, to predict the TSV, output of the algorithm. Once the models are trained, the procedure consists in perturbing the input parameters (*HRV,* t_a*, RH,* v_a) one at a time, with different uncertainties, maintaining constant the other quantities, and applying the trained model to each perturbed set of features, to obtain the final measurement of the TSV. It is worth noting that the current analysis is conducted locally, by choosing one observation of the whole dataset; the final result is therefore associated to the chosen observation. Characteristics of the sensors used to collect environmental parameters, and the measurement uncertainty employed to perturbate the different parameters are showed in Figure 22.

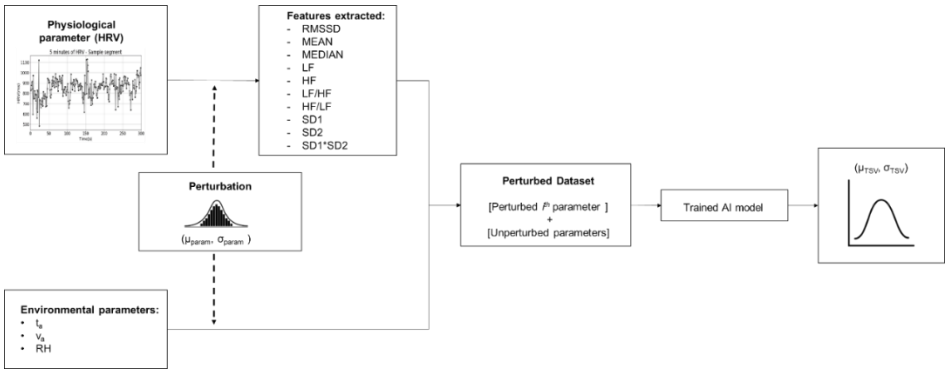


Figure 22. Description of the procedure adopted to apply MCM to the measurement of thermal comfort, with a heterogeneous set of data

3.2.7.3 Sensitivity Analysis

A sensitivity analysis (SA) is performed to identify and quantify what are the main sources of uncertainties in the measurement. Ideally, SA is the methodology for studying how the uncertainty of the output provided by a model, can be associated, qualitatively or quantitatively, to different uncertainties of the input parameters of the model [78], [81]. The GUM Supplement also provides the specifications for assessing the sensitivity coefficients; more in detail, it is explained that MCM is not sufficient to fully compute sensitivity coefficients but provides a methodology to compute the influence of each input quantities on the output quantity [78].

The procedure implies that one input quantity should be perturbed, and the remaining input quantities should be kept at their best estimates, in order to obtain the *pdf* of the output quantity, depending only on the variable perturbed. By using this procedure, the GUM supplement proposes an approach that can be representative as a generalization of the approximate partial-derivative formula; in particular, it is reported that the sensitivity coefficients can be approximated as the ratio of the standard deviation of the resulting model values and the standard uncertainty associated with the best estimate of relevant input quantity, as reported in Equation 4:

$$c_i = \frac{u_i(y)}{u(x_i)} \tag{4}$$

where c_i is the sensitivity coefficient, $u(x_i)$ is the standard uncertainty associated with the *i*th input estimate x_i, that contributes to the standard uncertainty $u_i(y)$.

3.3 Integration of skin temperature into thermal comfort measurement

The last part of this thesis deals with a further experimental session, created to include in the personalization of the thermal comfort of occupants, an additional physiological parameter, that is the skin temperature. As it was possible to deduce from the study of the state of the art reported in Section 2.6, in all the experiments in which the skin temperature was included as a parameter for the measurement of thermal comfort, relationships and correlations with the final parameter to measure were always found. The state of the art has also shown that currently there are no studies that combine HRV and its features, and the skin temperature in a single dataset to measure thermal comfort. Rather, many studies have focused on the single temperature of the skin in various regions of the body, or the combination of it only with the heart rate [29], [81]–[83]. It is reported in [84] that skin temperature alone can be adopted to evaluate individual thermal sensation at air temperature range between 30°C and 39°C and that in particular accuracy up to 93% can be reached in predicting two categories of discomfort (hot/very hot), using wrist temperature and Fisher discriminant analysis. The sensors used in this study are button-shaped thermometers directly placed on eight body parts. A combination of skin temperature and heart rate to develop personal thermal comfort models is used from [85]; skin temperature was collected by using iButton sensors placed on the wrist and ankle and demonstrated that ankle skin temperature is the most predictive parameter of the set of features. Skin temperature and heart rate are also employed from [86], but the scope is for detecting the effect of emotion state on occupant's thermal comfort. Consequently, since the literature has repeatedly found that both HRV and skin temperature taken individually can be indicators of the state of comfort or discomfort of the occupants, in the study that will be presented in the following sections, it is provided the description of an experimental campaign in which participants are exposed to changes in the environmental conditions of the test room, while a smartwatch and a minimally invasive sensor simultaneously record HRV and skin temperature.

The aim of this study is to be able to improve the measurement of the uncertainty of the TSV, compared to the studies presented previously, by including the skin temperature into the measurement process for assessing human thermal comfort, in real-life conditions. It is worth pointing out the similarities and differences with previous studies:

- The experimental campaign was conducted in the summer rather than the winter season; this requirement was introduced since it is reported that skin temperature in summer is more subjected to fluctuations that can lead to discomfort [87];
- The experimental set-up for physiological parameters is made of a smartwatch and a skin temperature sensor;

- To avoid that the participant may find themselves in uncomfortable situations, which can negatively affect the HRV measurement, participant is free to move without specific movement constraints to keep as still as possible the hand where the sensors are worn.

3.3.1 Measurement set-up

The tests took place in one room of the laboratory of Università Politecnica delle Marche (Ancona, Italy); the dimensions of the room are 4.8m X 2.9m x 3.0m and the planimetry is shown in Figure 23. The test consisted of monitoring a series of voluntarily recruited participants, while they are able to perform light-office activities; during the experiment, one operator was present to check that all sensors were functional and to control the internal conditions of the room, by means of a heater and a portable air-conditioning system.

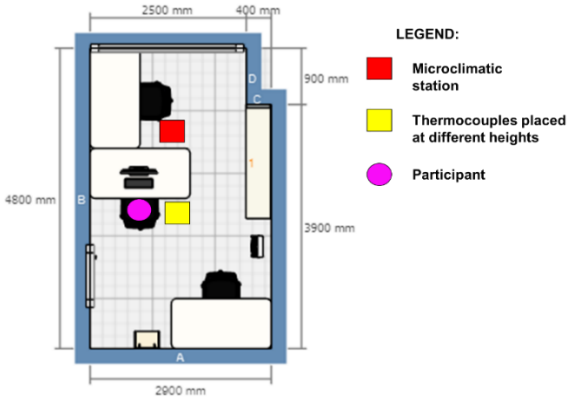


Figure 23. Planimetry of the test-room used for performing the experimental campaign, with the related dimensions and position of the equipment and the occupant during the experiment.

The monitored parameters during the test are: relative humidity (*RH*), air velocity (v_a), air temperature at 0.1m (t_L) , 0.6m (t_M) and 1.1m (t_H), black globe temperature. The accuracies of the sensors are reported in Table 8. All the sensors are compliant with ISO 7726. The physiological parameters of the tested participant are measured by using a wearable smartwatch and a sensor for the measurement of skin temperature. The participant wears a Samsung Galaxy Watch, whose accuracy in the measurement of HRV was tested in previous studies [34], [35]; the smartwatch was responsible for collecting HR and HRV, with a sampling frequency of 1 Hz (Figure 24). Collected data are sent in real-time to a laptop that acts as gateway, using a Websocket connection and then data are stored locally. Skin temperature was collected by means

of an iButton DS1922, by placing it on the wrist of the participant through a dedicated support and data are retrieved at the end of the test, by using the dedicated proprietary software.

Table 8. Technical information of the sensors for environmental parameters monitoring.

Sensor	Parameter monitored	Accuracy
Thermal-hygrometer	t_a	± 0.1°C
	RH [%]	± 1.5%
Black globe radiant temperature sensor	t_r *[°C]*	± 0.15°C
Hot wire anemometer	v_a *[m/s]*	±0.05 m/s
Thermocouples	t_L	± 0.5°C
	t_M	± 0.5C
	t_H	± 0.5°C

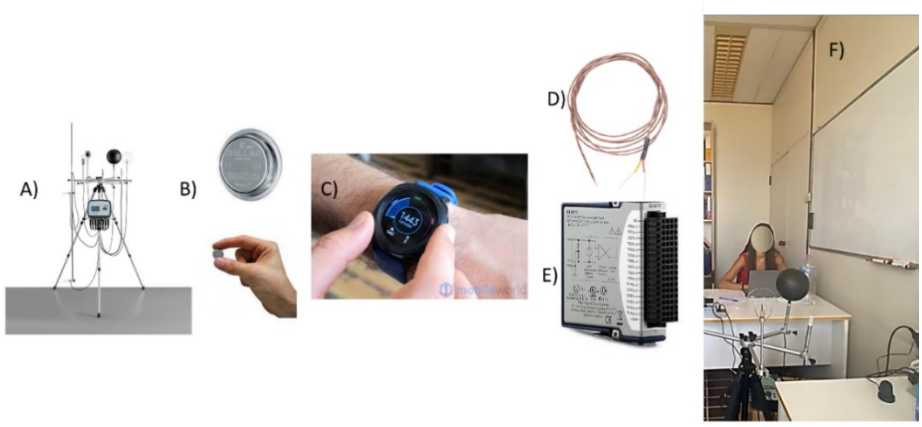


Figure 24. Measurement set-up of the experiment: (a) microclimatic station for recording globe temperature, (b) iButton D1922 for measuring skin temperature, (c) Samsung Galaxy Watch used to collect HRV, (d-e)) thermocouples used to measure air temperature at three different heights and the acquisition board of the National Instrument, (f) on-going test.

Table 9. Technical information of the sensors for physiological parameters monitoring.

Sensor	Monitored parameters	Accuracy
iButton	t_{skin}	± 0.5°C
Samsung Galaxy Watch	*HRV*	±4 ms

3.3.2 Measurement test procedure

The current work includes outcome from an experimental campaign that took place during July 2021. The experiment was conducted effectively on 15 participants. The involved participants were voluntarily recruited, and the personal information are summed up in Table 10.

Table 10. Personal information of the sample composition that participate to the experiment.

ID	Age (yrs)	Gender	Weight (kg)	Height (m)	BMI (kg/m^2)	I_{cl}
1	24	F	54	1.61	20.8	0.41
2	32	M	84	1.85	24.5	0.42
3	28	F	63	1.75	20.6	0.31
4	28	M	73	1.8	22.5	0.42
5	27	M	90	1.9	24.9	0.44
6	31	M	74	1.87	21.2	0.44
7	34	F	67	1.79	20.9	0.25
8	31	F	55	1.68	19.5	0.4
9	22	F	75	1.7	26.0	0.41
10	21	F	51	1.63	19.2	0.26
11	31	M	98	1.94	26.0	0.33
12	24	F	53	1.67	19.0	0.35
13	30	M	64	1.73	21.4	0.44
14	23	F	60	1.66	21.8	0.35
15	25	F	64	1.78	20.2	0.28
μ	27		67	2	22	0.4
σ	4		13	0.1	2	0.1

During the test, the participant is first exposed to a stabilization period of 10 minutes, inside the test-room, whose set-point is established at 25°C, which corresponds to the neutral state in summer according to actual standards [50]. Also, in this experiment, participants were required to wear their everyday clothes to preserve the subjectivity of the test. During this stabilization period, participant is required to fulfil a survey to track their personal information, including clothing information. In addition, during this stabilization period, the participant seats on the dedicated workstation and wears the smartwatch and the skin temperature sensor on the non-preferred wrist. The clothing thermal insulation I_{cl} reported in Table 10, was computed indirectly, by summation of the partial insulation values for each item worn by participants. The set of garments worn by participants was collected and the procedure suggested by ISO 9920 was adopted to evaluate the I_{cl}.

These actions are supervised by the operator which always stays in the test-room to check the correctness of the data collection, avoid possible malfunctioning of the sensors network and annotate unexpected movement of the participant that could lead to the measurement of misleading data. Then, the operator checks the thermal sensation vote of the participant, which is expected to be 0, since the test should start when user is at neutral condition. When all this information is collected, the test could start.

The procedure is specifically tailored to permit the user to perform its daily office activities that only include working on the laptop or reading a paper. The user is not forced to keep the hand in which there is the smartwatch in a fixed position, to avoid possible perturbation that can result in stressful condition due to the constraint of the experiment. In this way, it is possible to ensure that the only perturbation in the *HRV* and T_{skin} signal is only due to a change in the air temperature of the test-room.

After the preparation phase the test starts; more in detail, the test is built as following (Figure 26 and Figure 27):

- **Comfort segment**: 5 minutes in which the subject is free to perform office activities (e.g., read and write on a computer). The room temperature, set at 25°C, is just recorded, and it is not modified. The subject will be asked to communicate its Thermal Sensation Vote.
- **Discomfort segment 1:** 30 minutes in which the test-room is cooled down until 22°C.
- **Discomfort segment 2:** 30 minutes in which the test-room is heated up until 28°C.

It is important to specify that the participant is not informed about the procedure of heating or cooling down the test-room to avoid bias in thermal sensation. During the whole duration of the experiment the user is asked to fil a survey that indicates its TSV; also, in this experiment the TSV should be expressed whenever their TSV is different from the previous one.

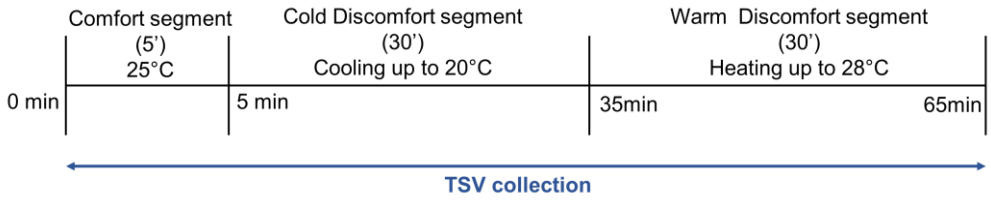
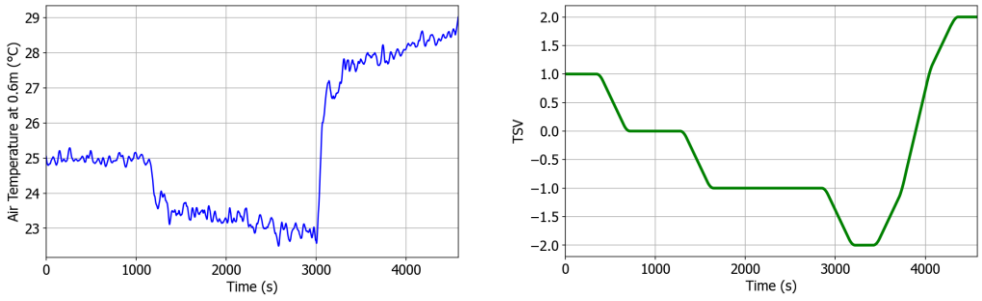


Figure 25. Experimental procedure.



(a) (b)

Figure 26. (a) Shape of the air temperature measured at 0.6m (t_2) during the experiment. (b) Shape of the TSV during the experiment, for one participant.

3.3.3 Data analysis

There are two approaches applied to build PCM: the first one is to develop a general model base on the LOSO approach; the second one consists of developing a model for each participant. The methodology will be useful for investigating and developing AI (DL and ML) algorithms that would predict the TSV of users, starting from HRV features, wrist skin temperature and environmental parameters.

HRV features were extracted from the sequence of RR intervals collected from the Samsung Galaxy Watch. RR intervals were firstly filtered in order to identify outliers; the methodology for filtering the RR sequence consists of detrending the RR interval time series and identifying as outlier a sample of the time series which differs from the previous sample of a thresholds of 20%. After identifying it, the sample recognized as outlier is substituted with the previous one. The filtered final RR sequence, also named HRV signal, is then used to compute HRV features.

After HRV processing, there is the creation of the dataset that should be used to predict and measure the TSV to estimate the thermal comfort of the participant. At this stage, the parameters that are possibly included in the dataset are a bundle of physiological and environmental data; the detailed description of the variable computed is reported in Table 11:

Table 11. List of all the variables extracted for each participant from the experimental campaign.

Environmental	**Physiological**					**Personal**
t_a	Skin temperature	t_{skin}				
RH [%]	Time-domain features	*MEAN*	*MEDIAN*	*RMSSD*	*SDNN*	
t_r *[°C]*		*PNN50*	*NN50*	*PNN20*	*NN20*	
v_a *[m/s]*	Frequency-domain features	*VLF*	*LF*	*HF*	*LF/HF*	TSV
t_L		*HF/LF*				
t_M						
t_H	Non linear -domain features	*sd1*	*sd2*	*sd1*sd2*		

3.3.3.1 Preliminary analysis on skin temperature and LF/HF

A preliminary analysis was carried out concerning the study of the correlations between LF/HF and skin temperature, that are the parameters that in literature are most correlated with thermal comfort, and the TSV measured during the test.

Linear correlation between the TSV and t_{skin} was computed, by means of the Pearson's correlation coefficient. Figure 31 shows the shape of t_{skin} against the TSV, for four randomly selected participants. The visual analysis, in fact suggests that the two trends are somehow similar.

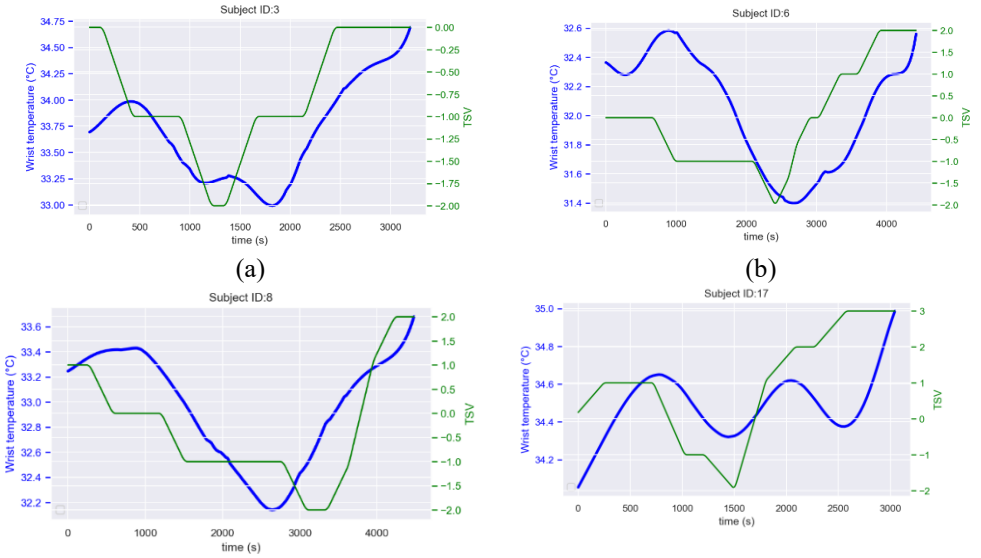


Figure 27. Shape of the skin temperature measured with the iButton against the TSV.

It has been observed that in real-life settings, the relationship that exists between LF/HF and TSV is not always easily deducible, so before moving on to the adoption of LF/HF as a feature to be used to train the AI algorithms, it was chosen to extract an additional parameter from the LF/HF, to highlight its relationship with the TSV. Firstly, a visual inspection of the data obtained was made, comparing the LFHF trend with the TSV trend; these data are reported in Figure 28, where the LF/HF and TSV curves are shown. For each participant, the linear correlation was evaluated by means of Pearson's correlation coefficient, and the resulting value is 14.06% (±29.2%).

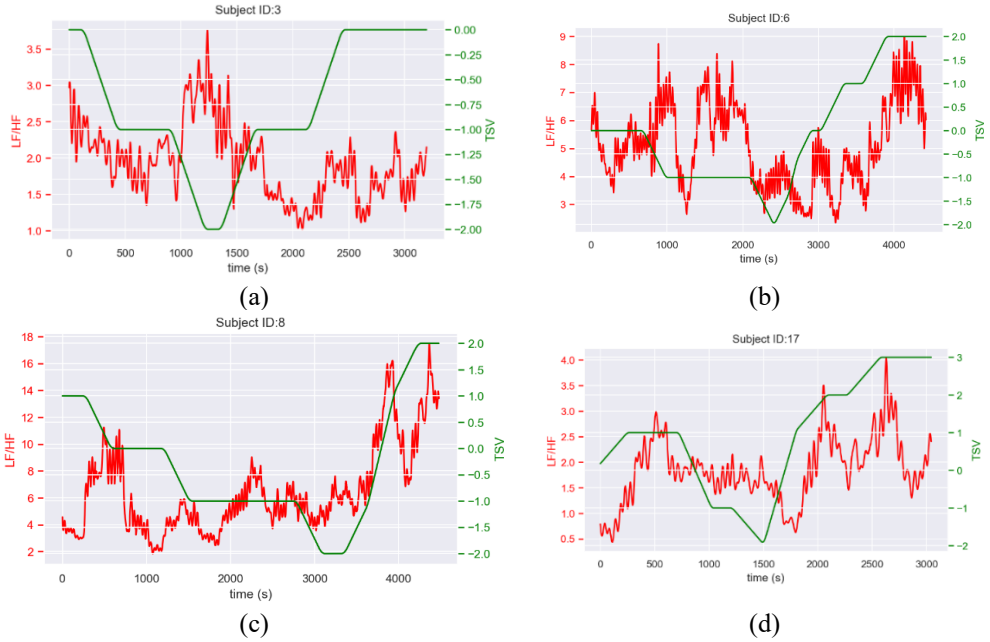


Figure 28. Shape of the LF/HF against the TSV.

Data of 4 participants were randomly selected: as can be seen from the figure, the LF/HF shows an oscillatory nature and often characterized by baseline drift that make it difficult to correlate it linearly with a more conservative quantity such as the TSV. Therefore, to make these two quantities comparable from a numerical point of view, it was decided to carry out further manipulations on the LF/HF signal to extrapolate its correlation with the TSV, as should be expected based on what is reported in the literature. In practice, it was decided to identify the instants of time within which the TSV undergoes a variation, that is a change in its value. the instants of time t1 and t2, that identify an time windows in which the TSV is constant, are identified; then the average value of LF/HF within the interval [t1,t2] is computed and it is associated to each sample of TSV for the window [t1,t2]. Pearson correlation between TSV and the new manipulated LF/HF are then computed. This new feature is included in the initial dataset created.

3.3.3.2 Feature selection

One of the challenges encountered in this research was mainly due to the necessity of individuating a reliable set of features that could possibly maximize the measurement of human thermal comfort in transient condition. To select the features, it is applied a procedure that combines two different techniques, Pearson's correlation and feature importance provided by Extra Tree Classifier [88]. The features that are less important and provide lower Pearson

correlation coefficient with the variable to be predicted, which in this case is the TSV, is dropped from the dataset.

Pearson coefficient (*R*) is used to compute the correlation coefficient between each feature of the dataset and the TSV; R ranges from -1 to 1: low correlation coefficient is applicable for value of *R* close to 0, while high positive or negative correlation is associated to *R* values of +1 and -1, respectively. Feature importance computed by means of Extra Tree Regressor, is based on the Gini Importance of each feature.

3.3.3.3 Model Training and evaluation

To assess the TSV of the participants that attend the experimental campaign, the TSV collected during the test were used to develop AI regression models that could estimate each participant TSV. Different algorithms were trained and tested with two different type of evaluation metrics:

- Firstly a PCM was developed, one for each subject. It is a person-tailored model that consists of training and testing each model uniquely on the data belonging to the subject under analysis. To evaluate the performance of the algorithm, the hold-out methodology was adopted, that consists on using the 70% of the data to train the algorithm and make the validation by testing the trained algorithm on the remaining 30% of the data that were not included into the training.

- The second approach consist of applying the LOSO approach, that was previously explained in Section 3.2.6.5; this method can provide a comparison with the previous study, to evaluate if indeed the combination of HRV and skin temperature, can overall provide an improvement in the performance of the algorithms.

In both cases, the metrics for evaluating the performances of the regression algorithms are the MAE and the MAPE. The algorithms chosen included ML algorithms such as K-Nearest Neighbors (KNN), Decision Tree (DT), Support Vector Machine (SVM), Adaboost (ADA), Random Forest (RF), Extra Tree Classifier (ETC); DL algorithms included Multi-layer Perceptron (MLP).

Chapter 4.

Implementation

4.1 Development of a methodology for evaluating the correlations between thermal comfort and HRV in controlled environment

The current section deals with an analysis of the complex dataset measured during tests performed following the experimental campaign that took place in the controlled environment, for measuring the thermal sensation of the participant. The following paragraph presents preliminary analyses on recorded ECG signals and extracted features which are correlated to monitored environmental parameters. These results are then compared to literature achievements in the field in a view of measuring procedure validation and providing the bases for future developments of the study. In addition, Section 4.2.1 shows results in terms of thermal comfort prediction from ECG features through supervised ML algorithms.

4.2 Preliminary analysis results

Results related to the analysis of the linear correlation between environmental and physiological quantities are here presented. LF/HF ratio is correlated to air temperature (ta), mean radiant temperature (tr), CO2 concentration (CO2), relative humidity (RH) and to PMV.

An example of the output provided for two subjects of the study is presented in Figure 29: Subject 1 which performed the test in winter, and Subject 2 which performed the test in summer. Pearson coefficient is computed for time intervals of 5 min; therefore, every coefficient expresses the linear correlation between LF/HF and one environmental parameter taken in the same time interval. Figure 29 (a) and (b) shows time trends of LF/HF computed from the ECG while Figure 29 (c) and d represent air temperature trend for the whole duration of the test. Figure 29 (e–f) presents values of Pearson coefficient computed every minute for each considered environmental parameter, through a grey-scale color palette. Lighter regions are associated to low Pearson coefficient while darker regions to strong correlations. Every row of Figure 29 (e) and (f) represents the outlined Pearson correlation coefficient between LF/HF and air temperature (R-t_a), mean radiant temperature (R- t_r), CO2 concentration (R- CO_2), PMV (R-*PMV*) and relative humidity (R-*RH*).

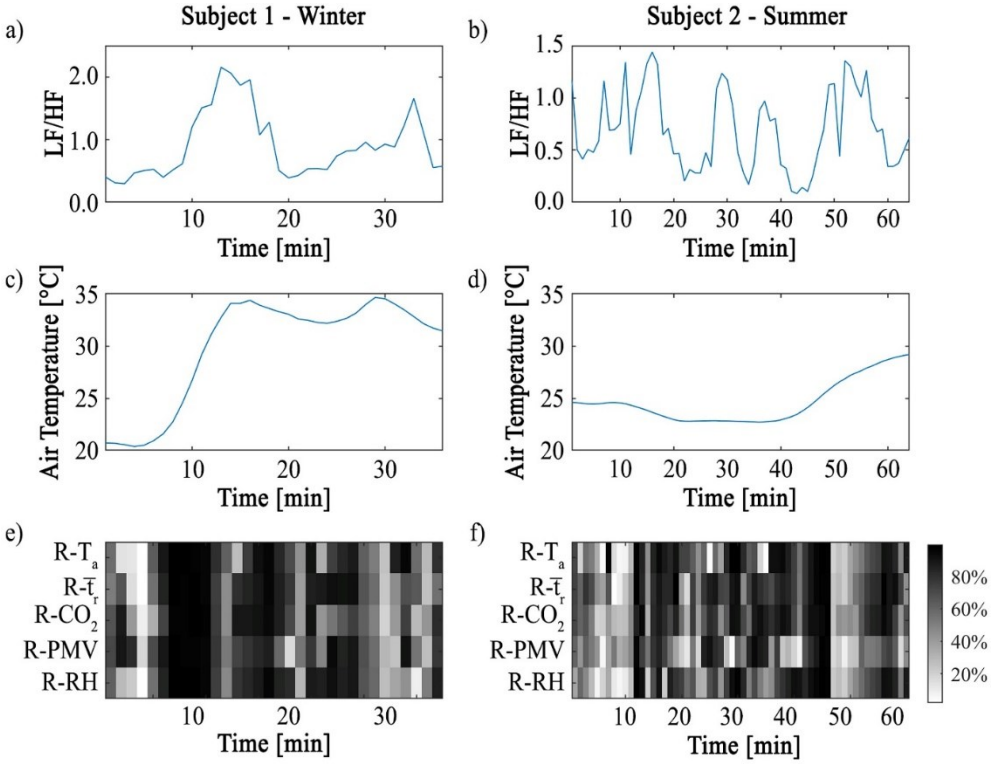


Figure 29. (a,b) Waveform of LF/HF ratio and (c,d) air temperature in time for subject 1 (a,c) and 2 (b,d); (e,f) Pearson correlation coefficient computed in 5 min time interval of LF/HF and the respective time interval in the environmental parameter.

Presented results concerning Subjects 1 and 2 are chosen to show that LF/HF trend is correlated to a variety of environmental parameters. In fact, subject 1 exhibits an LF/HF ratio increase between 10 and 15 min during the test that is strongly correlated with all the five parameters of interest, i.e., t_a, t_r, CO_2, *RH*, and *PMV*. In fact, increase of the LF/HF ratio derives from the simultaneous increase of the correlated parameters. In the same way, it is possible to interpret results of Subject 2: t_a and t_r do not exhibit significant correlation with LF/HF especially considering the first LF/HF ratio observed peaks, i.e., around minute 5 and 10. This can be probably explained due to the small variation of the environmental parameters in the first part of the test, while a higher variation is presented at the end. These results put in evidence that LF/HF trend, in dynamic environmental conditions, is not only strictly related on one parameter, e.g., t_a, but can be also influenced by CO_2 in the room, t_r, or *RH*. Computed Pearson coefficient between LF/HF ratio and both CO_2 and t_r, is generally above the 80% for the entire duration of the Subject 1 test. Moreover, R- CO_2 is higher in correspondence of the peaks in LF/HF ratio at minute 10, 29 and 40 of the Subject 2 tests. This is an important result showing that subject's

comfort is not only related to Ta which is relevant for the second peak together with CO_2 and t_r but not for the first one where the LF/HF ratio waveform increase for a R- CO_2 raising.

The analysis is repeated for every participant. Then, just the Pearson's coefficient values greater or equal to 75% are considered as significant for the final evaluation. Table 12 presents the percentage of LF/HF time interval that, correlated with environmental parameter, showed a Pearson's coefficient greater or equal to 75%. For example, 66.4% of the air temperature time intervals correlated with LF/HF scored a Pearson's coefficient higher than 75% in winter, suggesting that there is a relationship between LF/HF and air temperature. Same considerations can be done for CO_2 concentration, since 65.3% of the time interval has correlated more that 75%.

Table 12. Percentage of LF/HF signal correlated with environmental signal time interval that has a Pearson coefficient greater than 75%, in summer and winter.

		Winter tests				
		t_r - LF/HF	***t_a - LF/HF***	***CO_2 - LF/HF***	***RH - LF/HF***	***PMV - LF/HF***
Average	[%]	56.8	66.4	65.3	60.0	37.0
Standard deviation	[%]	8.1	0.5	5.0	0.6	13.9
		Summer tests				
		t_r - LF/HF	***t_a - LF/HF***	***CO_2 - LF/HF***	***RH - LF/HF***	***PMV - LF/HF***
Average	[%]	56.8	66.4	65.3	60.0	37.0
Standard deviation	[%]	8.1	0.5	5.0	0.6	13.9

Results concerning Subject 1 and Subject 2 have been chosen because they represent the linear correlation between environmental quantities and LF/HF. The same level of linear correlation with environmental quantities can be found on 13 out of 29 subjects. The remaining subjects turned out to provide a weak linear correlation. Since literature has shown that LF/HF can be related to thermal comfort [28], [89] and that the remaining subjects had no correlation between LF/HF and environmental quantities, it was supposed that there are more factors that drive LF/HF, apart from environmental quantities, linked to the behaviour of the single subject. The result of this analysis confirms the complexity in representing the human perception toward the environmental conditions, that could be better investigated using more complex modelling functions, typical of the ML. For this reason, the following section explains the analysis conducted using ML algorithms to achieve a more accurate model of the human perception based on physiological features with the aim of replicating the users' real perception recorded during the tests.

4.2.1 Machine learning results

Table 13 expresses the average prediction accuracy of the classification algorithms used to predict the thermal sensation vote expressed from subjects that participated to the study in relation with HRV indices. In particular, in the first row the six classification algorithms are trained only using the LF/HF as input data and the thermal survey as label. The average accuracy of DT, KNN, LDA and RF are close to 50% while NB and SVM increases up to 76%.

Table 13. ML classification accuracy computed as the average accuracies for each of the 29 subjects.

Mean ML classification accuracy [%]						
Indices	**DT**	**KNN**	**LDA**	**NB**	**SVM**	**RF**
LF/HF	52	52	55	75	76	51
Time	69	73	74	81	82	73
Frequency	63	68	64	79	80	69
All	72	68	77	82	84	79

The second and third row show the average accuracy of the algorithms trained with a dataset obtained from the aggregation of time-domain and frequency-domain HRV measurements respectively. Accuracy has increased in both cases in all the classifiers but also in this case NB and SVM have provided higher accuracy, up to 82%. Moreover, it has to be pointed out that all algorithms classify the thermal vote with lower accuracy in the frequency domain with respect to time-domain indices. Finally, the accuracy of the algorithm trained with a dataset obtained from the aggregation of all the computed HRV indices is shown in the last row. The higher accuracy is reached by SVM algorithm (84%) while KNN provides the lowest performance (68%).

Figure 30 compares time-trends of real expressed thermal sensation and predicted thermal sensation according to the developed model for one subject tested during summer as example. As final consideration, ML classification approaches, especially NB and SVM allow using the LF/HF to predict the thermal comfort vote of a user in an indoor environment even if better results are shown for the accuracy of time, frequency and aggregated indices.

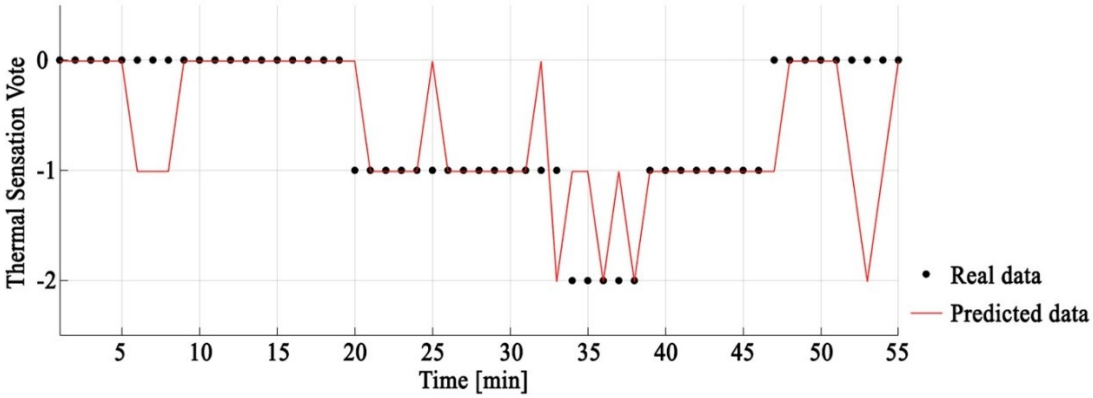


Figure 30. Thermal sensation vote predicted from SVM classification model (red line) against the real thermal sensation vote (black dots) for one subject out of the performed summer tests as example

4.2.2 Results of the analysis of subjective parameters in relation to thermal comfort

This section focuses on subjectivity of the perceived IEQ moving from a preliminary combined analysis of environmental data measured by the indoor microclimate station and expressed actual sensation votes. These are continuously noted by the operator throughout the whole test being translated into a 7-points comfort scale ranging between −3 and +3 where 0 expresses neutral conditions. The proposed analysis refers to data collected every minute for all the 62 performed tests, for a total amount of 1360 and 1128 data for winter and summer test series respectivelyPresented outcomes highlight how much IEQ perception varies among different subjects and thus occupants in general. Physical variables ranges corresponding to specific sensations are wide, e.g. up to 14.6 K and 8.9 K of operative temperature interval expressing thermal neutrality in winter and summer respectively, and almost overlapped.

In order to understand the reasons underneath the presented differences among expressed sensations combined to monitored environmental parameters, possible influences due to personal attributes of the subject are here investigated. In particular, the hypothesis here statistically tested assumes expressed IEQ satisfaction, i.e. fill-up survey part III, as dependent from personal characteristics of the subject gathered in survey part I. These characteristics are converted into markers highlighting two classes for each variable, as summarized in Table 2. Two different analyses are conducted by assuming groups of objective/subjective descriptors or single personal characteristics. Linear regression model is used to identify correlations between expressed IEQ satisfaction and grouped personal characteristics of the tested subject, i.e. subjective and objective descriptors (Table 2). Fig. 9 shows computed p-value which quantifies the goodness of the proposed model. The lower the p-value, the higher the significance of the tested hypothesis with a typically upper limit assumed equal to 0.05 to define statistically significant results. Only

few of the proposed correlations show up to be significant but such relations are not consistent throughout different tested seasons. The lowest p-value is 0.02 obtained for the hypothesis of objective personal parameters influencing the visual perception in winter at the beginning of the test. The other significant relation observed in summer still refers to the initial phase of the test, it has a p-value of 0.03, and concerns thermal comfort perception as depending on personal subjective attributes, i.e. Body-Mass-Index, education, smoking habits, and worn garments. Table 14 and Table 15 reports results obtained from the independent t-test conducted on thermal comfort sensation, assuming two classes for each personal characteristic. Significant relations are given by combining p-values lower than 0.05 and Cohen's d higher than 0.5 which expresses a medium effect size. These relations are highlighted in bold in the table. Some of the obtained p-values identify stronger relations even if these are never consistent throughout seasons or different phases of the test. Results confirm some of the most investigated correlations, e.g. gender and cloths are both related to thermal perception.

Table 14. The three single personal characteristics influencing expressed TSV perception and related p-value and Cohen's d, during winter tests.

WINTER TESTS				
	Initial test phase		**Final test phase**	
	Personal characteristic	***p*-value ; Cohen's *d***	**Personal characteristic**	***p*-value ; Cohen's *d***
Thermal comfort	Smoking habits	0.01 ; 0.11	Education	0.30 ; 0.36
	Gender	**0.05 ; 0.67**	BMI	0.43 ; 0.07
	BMI	0.36 ; 0.72	Age	0.47 ; 0.27

Table 15. The three single personal characteristics influencing expressed TSV perception and related p-value and Cohen's d, during summer tests.

SUMMER TESTS						
	Initial test phase		Middle test phase		Finale test phase	
	Personal characteristic	*p*-value ; Cohen's *d*	Personal characteristic	*p*-value ; Cohen's *d*	Personal characteristic	*p*-value ; Cohen's *d*
Thermal comfort	Garments	**0.03 ; 0.59**	Gender	0.07 ; 0.72	Gender	**0.02 ; 0.94**
	Education	0.12 ; 0.62	Smoking	0.11 ; 0.89	Age	0.11 ; 0.64
	Age	0.26 ; 0.44	BMI	0.18 ; 0.57	BMI	0.30 ; 0.43

4.3 Measurement procedure and data analysis for assessing thermal comfort in semi-controlled environment

4.3.1 Binary classification between comfort and discomfort

The binary classification between comfort and discomfort was considered the baseline analysis of the whole test. The aim of this part was to estimate whether it is possible to distinguish between comfort and discomfort condition caused by an external disturbance generated by a sudden change in temperature, air velocity and relative humidity. In the following paragraph, the binary classification was tested for the cold-induced discomfort and then for the hot-induced discomfort. The three ML algorithms extracted from literature, i.e. SVM , RF and ET, were trained and tested on the single participants to see whether HRV frequency-domain indices can be effectively used to distinguish if a person is experiencing cold-induced discomfort or warm-induced discomfort. This assumption is necessary in a future hypothesis of including the HRV measurement as a support tool for controlling indoor environment.

Nine out of the 10 subjects were considered for the cold-induced discomfort experiment, while one subject was excluded for having expressed no variation in the TSV when entering Room B. The variation in the TSV between Room A and Room B is essential to build the dataset for the ML classification, since this is the only way to develop a binary classification in which to relate physiological quantities and TSV. If the TSV of the participants remains 0 when exposed to discomfort in Room B, it is not possible to develop a model that classifies between two classes, because no variation of the output is recorded. From Table 16, it can be seen that the three algorithms performed at a high level of accuracy in distinguishing the two classes, with values up to 100% in the case of marked distinction and a mean accuracy of 92.2%. This result suggests that, with the support of ML classification algorithms, HRV features can be used to distinguish whether a participant is comfortable or in a discomfort condition. The same considerations can be made for the warm-induced discomfort classifier. Six out of the ten participants were included, since they provided a different TSV between Room A and Room B. As shown in Table 17, also in this case the accuracy reached a mean value of 92.9%, suggesting that the two indices (LF/HF, HF and LF) can be an indicator of warm or cold induced discomfort.

Table 16. Accuracy of the ML classifiers in the estimation of the cold-induced discomfort condition.

Accuracy (%)										
	User									
	1	2	3	4	5	6	7	8	9	Mean
SVM	90.6	78.2	92.3	100	73.4	99	86	98.2	100	90.9
RF	90.6	79	96.8	100	88.6	98.2	100	91.6	98.2	93.7
ETC	90.5	72	94.5	100	86	96.5	100	92.5	98.2	92.2
Mean										92.3

Table 17. Accuracy of the ML classifiers in the estimation of the warm-induced discomfort condition.

Accuracy (%)										
	User									
	1	2	3	4	5	6	7	8	9	Mean
SVM	-	80.2	86.8	100	85	-	100	-	77.7	88.3
RF	-	93.7	90.1	100	93.2	-	100	-	97.6	95.8
ETC	-	94.6	96.7	100	94.5	-	86	-	96.4	94.7
Mean										92.9

4.3.2 Prediction of thermal discomfort in transient environment in Experiment 3

The activity here presented pointed out the necessity to think about the relationship between physiological, environmental parameters and the TSV as a black box that can be implemented thanks to ML techniques. This is the approach used in Experiment 3, which merged different types of datasets that included various environmental and physiological features to predict the TSV of each participant. Among different algorithms tested, the best results were obtained with the RF algorithm, against, for example, the SVM algorithm, which provided lower performances. The algorithms were trained using a leave-one-subject-out validation and their performance was evaluated in terms of MAE and MAPE.

The algorithms were first trained and tested by using as input only the physiological features previously described in Table 6. Given the higher number of computed features, the model was trained multiple times, combining iteratively different subsets of physiological features to individuate the best combinations of parameters that could provide higher performance of the algorithm in terms of accuracy. With this procedure it was possible to identify that the combination of physiological features that led to higher accuracy was made of RMSSD, MEAN, MEDIAN, LF, HF, LF/HF, HF/LF, SD1, SD2, SD1*SD2, that led to the creation of FH. FH was subsequently combined with the three environmental parameters (ta, va, RH), thus creating FH+E, which is showed in Table 18. The performance obtained by testing each algorithm on the participant that was not previously included in the training is shown in Figure 31 (a-b). The average MAE and the average MAPE were obtained by averaging all the MAE values and the MAPE values obtained from testing the algorithms on the single test participant. In particular, when the algorithms were trained using the dataset FH+E, both MAE and MAPE decreased compared to the training made only with physiological HRV features (FH), except for the average MAE computed from the LSTM algorithm.

Table 18. Description of the dataset built for this analysis.

Dataset							
	F_H			F_{H+E}			
	f(HRV$_t$)	*f(HRV$_f$)*	*f(HRVnl)*	*f(HRV$_t$)*	*f(HRV$_f$)*	*f(HRVnl)*	
	RMSSD	LF	SD1	RMSSD	LF	SD1	t_a
Features	MEAN	HF	SD2	MEAN	HF	SD2	*RH*
	MEDIAN	LF/HF	SD1*SD2	MEDIAN	LF/HF	SD1*SD2	v_a
		HF/LF			HF/LF		

To compare the five algorithms, accurate results were obtained by applying RF algorithms, with a MAE and MAPE of 1.4 and 24% respectively, when the FH dataset was used. On the other hand, CNN and RF provided comparable performances when using the FH+E dataset, with MAE and MAPE of 1.2 and 21, respectively. These results highlight how the inclusion of some user-related features, such as HRV features, can somehow provide a prediction of the TSV of the participant, but, at the same time, demonstrate how the inclusion of some environmental quantities can lead to a better prediction of the TSV of the user in transient conditions, as it occurs in Experiment 3. MAE values, suggests that TSV measurement through personalized approach, can be made through physiological features, such as HRV, but further studies should be performed to try to improve the measurement, which are presented in section 4.3.4 of this PhD work.

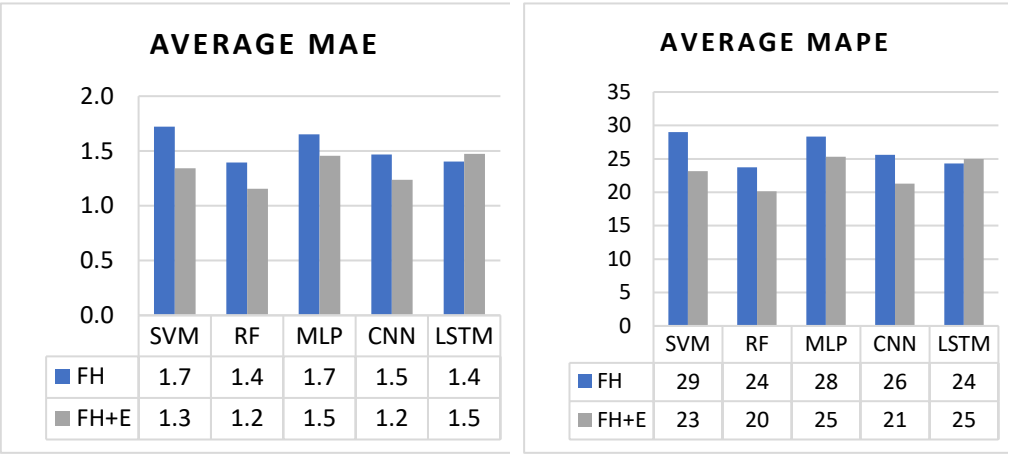


Figure 31. (a) Average MAE computed by averaging the MAE obtained from the prediction of each algorithm on the single user, left out of the training. (b) Average MAPE computed by averaging the MAPE obtained from the prediction of each algorithm on the single user, left out of the training.

Figure 32 shows the results obtained for one participant as an example; in addition, it also shows the trend of the PMV, which, although it is not used as a reference metric in this study and is not included for building the ML models, partly follows the trend of the real TSV. Furthermore, the figure shows that the trend of the predicted and the real TSV follows the shape of the PMV, although the predicted TSV deviates more from the PMV.

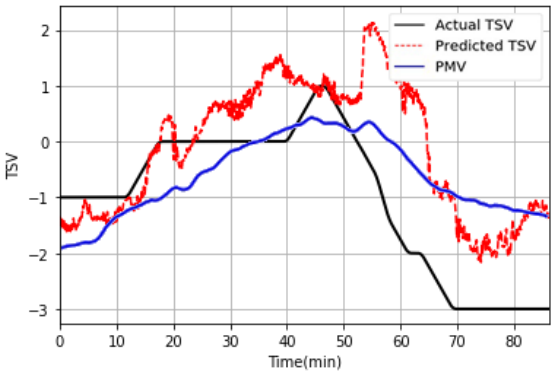


Figure 32. Real TSV and PMV against the predicted TSV obtained from the testing on one user, adopting the LOSO approach.

4.3.3 Discussion and limitations

The experiments conducted in the first research of this thesis work, described in section 3.1, are considered the starting point used to develop the procedure conducted in this work. In particular, [65] demonstrated that when the user does not perform any activity and environmental parameters (ta, RH and va) are more controlled, HRV features can provide accurate performance in the estimation of TSV. However, in this second part of this thesis, in which participants were allowed to carry out some light office activities, it was possible to demonstrate how HRV features are only able to discern between a comfort and a discomfort condition of the participant and provide lower performance in predicting the TSV in transient conditions compared to [65]. For this reason, the effect of the combination of HRV features with environmental quantities was examined to test whether a combination of the two could help in the prediction of TSV. On the other hand, [65] highlighted that human perception cannot be interpreted in a univocal manner on the basis of environmental parameters, but should be analyzed through a subjective point of view, which, in this case, is expressed by TSV.

For this reason, to predict TSV, it was decided to aggregate the physiological response of the participants (expressed by HRV features) that proved to be related to human thermal comfort with environmental quantities (FH+E). The evaluation of the performance of the algorithms suggests that features extracted from environmental parameters led to improved results compared to the use of HRV features alone. This result is compliant with literature, which suggests that environmental parameters can be helpful in the prediction of thermal comfort and that physiological measurements can be used as support features to develop personalized models [90]. Experiment 3 aims to simulate a real-life application in which users are able to perform light office activities while a smartwatch collects their HRV. By using FH to predict TSV, the MAE and MAPE values cannot provide satisfactory values of accuracy, which suggests that in real life conditions, different from the controlled environment created in [65], more complex mechanisms are present in the management of HRV. However, a step-forward in the prediction of TSV was made by adding environmental quantities to the HRV features, since the value of MAE and MAPE improved.

The research proposed was performed in order to explore whether it is possible to build ML models by collecting physiological data from a smartwatch, providing a methodology that could predict the TSV of the participants. The LOSO approach used for testing the accuracy of the ML algorithms aims at creating a generalized model that does not take into account gender-related differences but considers only TSV as the ground truth of the model. Although the results of the LOSO approach highlight that a punctual, real-time prediction of TSV was not always accurate for the current test, it is important to underline that the results obtained from ML algorithms can be used to predict the overall comfort of a user. This information can point out that the overall thermal comfort is evaluated over a long-term period, and it is not necessary to estimate a point-by-point trend. Given this assumption, this work tried to demonstrate how the methodology applied for the test, which involves a smartwatch and environmental sensors, can be further

explored to investigate its effectiveness in helping to predict the TSV of users in real-life contexts. The results obtained help to clarify that the aim of the entire work is to add physiological parameters in the measurement of thermal comfort to realize environments that are more tailored to the users that live in them. Thus, the analysis provided is helpful to investigate the capability of physiological quantities to predict the TSV of occupants, without taking the place of environmental quantities.

4.3.4 Results of the analysis of the impact of the uncertainty of measurement on the HRV features

4.3.4.1 Impact of the uncertainty on HRV features

In this section, results of the MCM used to analyse the propagation of the uncertainty in the measurement of HRV that propagates to the computation of the HRV features, are shown. The simulated results are used to obtain the frequency histogram of each HRV features that is computed with the simulation, to obtain the related uncertainty. For each HRV features the values of the standard uncertainty computed with the MCM are showed. The frequency histograms that come out of one simulation for one participant are displayed in Figure 33. Table 19 contains the uncertainty associated to the computation of each HRV features, expressed in percentage, among the 13 participants. It can be seen that the minimum measurement uncertainties is obtained when computing the MEAN feature (±0.01% of reading), while highest values are associated to the measurement of LF/HF and HF/LF (±0.7% of reading). Physiological quantities, such as HRV, can vary among participants and therefore it is interesting to understand the impact of the uncertainty on each feature, divided according to the related domain, among the different users, to see if it is possible to establish a generalized standard uncertainty associated to each HRV features, in relation to the value of the uncertainty of the device which measures it.

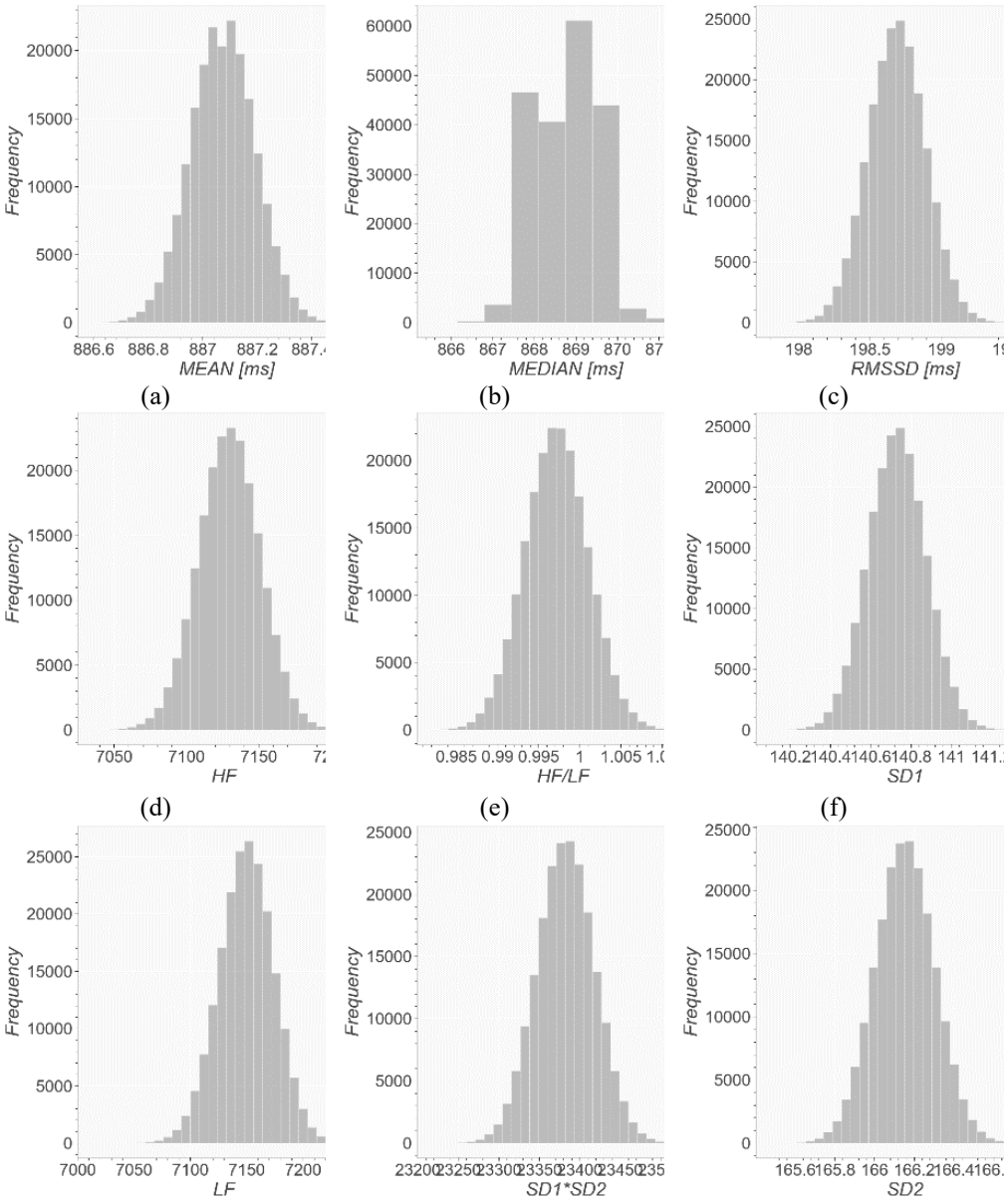


Figure 33. Histogram obtained from MCM with 200,000 iterations applied to the HRV features: (a) MEAN, (b) MEDIAN, (c) RMSSD, (d) HF, (e) HF/LF, (f) SD1, (g) LF, (h) SD1*SD2, (i) SD2.

The uncertainty associated with the features in the frequency domain is higher than the features in the time or the features in the non-linear domain. A reason that can explain this result is that

the frequency components are more sensitive to the uncertainty of the device (± 4 ms) because they physically reflect the oscillatory activity of the HRV; for this reason, it can be verified that a small variation of HRV can lead to a greater variation of features in the frequency domain.

Table 19. Measurement uncertainty computed for each HRV features used to estimate the TSV, among the 13 participants.

	Uncertainty (%)									
ID	**MEAN**	**RMSSD**	**MEDIAN**	**LF**	**HF**	**LF/HF**	**HF/LF**	**SD1**	**SD2**	**SD1*SD2**
1	0.01	0.10	0.09	0.34	0.30	0.4	0.4	0.10	0.1	0.2
2	0.01	0.08	0.09	0.31	0.25	0.4	0.4	0.08	0.1	0.1
3	0.01	0.29	0.07	1.30	0.81	1.5	1.5	0.29	0.3	0.5
4	0.01	0.11	0.08	0.53	0.55	0.6	0.6	0.11	0.1	0.2
5	0.01	0.25	0.06	0.93	0.58	1.1	1.1	0.25	0.2	0.4
6	0.01	0.10	0.08	0.30	0.28	0.4	0.4	0.10	0.1	0.2
7	0.01	0.16	0.10	0.57	0.59	0.6	0.6	0.16	0.1	0.2
8	0.01	0.04	0.13	0.26	0.12	0.3	0.3	0.04	0.0	0.1
9	0.01	0.12	0.08	0.31	0.32	0.5	0.5	0.12	0.1	0.2
10	0.01	0.24	0.06	0.92	0.58	1.1	1.1	0.24	0.2	0.4
11	0.01	0.09	0.07	0.31	0.25	0.4	0.4	0.09	0.1	0.1
12	0.01	0.30	0.08	0.66	0.79	1.0	1.0	0.30	0.2	0.4
13	0.01	0.06	0.10	0.26	0.19	0.3	0.3	0.06	0.1	0.1
μ	**0.01**	**0.1**	**0.1**	**0.5**	**0.4**	**0.7**	**0.7**	**0.1**	**0.1**	**0.2**

4.3.4.2 Impact of the uncertainty on AI models for measuring human thermal comfort

This section reports the results used to compute the measurement uncertainty of the TSV, using MCM applied to AI models, which are RF and CNN. Table 20 contains the results which represent the contribution of the uncertainty of each input quantity (HRV, t_a, RH, v_a) into the AI model (RF and CNN), to estimate the associated uncertainty in the output TSV. Each *pdf* of the input quantity is used, one at a time, according to Monte Carlo simulation, to obtain the resulting pdf of the TSV. The following table contain the results of a MCM conducted by perturbing each

parameter, one at a time, with the respective uncertainty $u_i(y)$, where i are the perturbed parameters.

Table 20. Results of MCM applied to compute the measurement uncertainty $\boldsymbol{u_i(y)}$ of the measurement of the TSV, both for the RF and the CNN algorithm. The uncertainty was computed considering a coverage factor k equal to 2.

	RF				**CNN**			
	HRV	t_a	RH	v_a	HRV	t_a	RH	v_a
Standard uncertainty	± 4 ms	± 0,1 °C	±2% of reading	±3% of reading	± 4 ms	± 0,1 °C	±2% of reading	±3% of reading
TSV Estimate (μ)	1	1.4	1.4	1.4	-0.5	-0.5	-1.2	-1.2
$\boldsymbol{u_i(y)}$ [TSV unit] (σ)	7.46E-14	0.2	1.78E-13	0.0105	0.015	0.015	0.126	0.029
$\boldsymbol{u_i(y)}$ [%]	7.24E-12	**14**	1.28E-11	0.8	2.825	2.9	**10.5**	2.5

The simulation regarding the impact of the HRV uncertainty of the TSV was made for ± 4 ms. Since HRV measurement through smartwatches are particularly prone to motion artefacts, more simulations were conducted by simulating different HRV uncertainties, which are U = [4, 10, 20, 50, 100] ms. As previously stated, this range of values was chosen since previous researchers have demonstrated that the measurement uncertainty of the galaxy Watch is ±4 ms in resting condition and it is ± 100 ms when the user is moving [34], [35]. The results of the simulations are showed in Figure 34: the contribution of HRV uncertainty increases after 50 ms, suggesting that up to this value, the uncertainty of the devices that measure HRV is still acceptable for assessing the TSV. On the other hand, environmental sensors that measures t_a, RH and v_a, are less subjected to variations in the uncertainty of the measurement of TSV. Environmental quantities produce different uncertainty on the TSV, which depends on different parameters. RF has highest uncertainty due to the t_a ($\boldsymbol{u_{t_a}(y)}$ = 14%), while CNN has highest uncertainty when RH is perturbed ($\boldsymbol{u_{RH}(y)}$ = 10.5 %). This different result is explainable if we consider that the two algorithms perform with different rules and can be considered as black boxes.

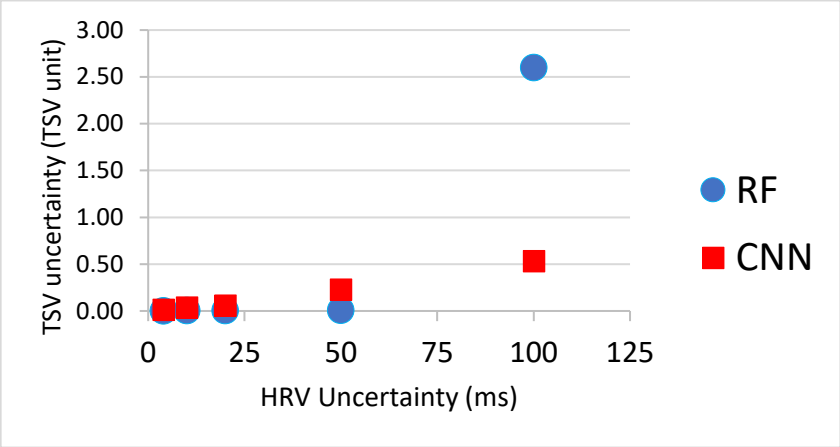


Figure 34. Trend of the measurement uncertainty of the TSV, in response to different HRV uncertainties measurement, obtained with 200,000 iterations.

4.3.4.3 Sensitivity Analysis

In Table 21 the partial uncertainty budget due to each parameter is given according to GUM uncertainty framework; the sensitivity coefficients were evaluated by Monte Carlo simulation. According to the table, SA highlights that air velocity and air temperature are the parameters that most affect the TSV prediction, respectively for RF and CNN. RH, and HRV perturbed with ± 4 ms are the less sensitive parameters in the measurement of TSV, while v_a is the parameter that mostly contribute to the TSV measurement uncertainty. This outcome is consistent with the test procedure, which implied that during the experiment, to generate thermal discomfort, window was opened, and participant reported to experience greater discomfort.

Table 21. Sensitivity coefficients computed for the two algorithms.

	HRV	t_a	RH	v_a
Standard Uncertainty	± 4 ms	± 0,1 ºC	±2% of reading	±3% of reading
Sensitivity Index RF	1.87E-14	1.8	4.61E-13	**5.8**
Sensitivity Index CNN	0.004	0.09	0.32	**16.03**

4.3.4.4 Discussion about the methodology

Depending on the kind of algorithm trained, there are different quantitative results in the uncertainty of the TSV, using an MCM, with 200,000 iterations, in relation to the environmental parameters. RF provides greater uncertainty when t_a is perturbed with an uncertainty of ± 0,1 ºC; on the contrary, CNN exhibits a higher uncertainty when RH is perturbed with ±2% of reading. Physiological parameter (HRV signal), when perturbed with ±4 ms of uncertainty, impacts the resulting TSV with 2.8% of uncertainty and > 0.001% when it comes with CNN and RF respectively. In addition, MCM on AI models for measuring human thermal comfort (Section 2.2) is applied to one observation of the entire dataset, and therefore it is a local analysis that is strongly related to the trained model. The overall methodology presented in this section can be applied to brand new model, that do not include an analytical equation, to compute the uncertainty associated to the output of the model.

4.4 Results of the integration of skin temperature into thermal comfort measurement

4.4.1 Preliminary analysis on skin temperature and LF/HF

First of all, the purpose of this section is to show the results of the experimental campaign conducted to measure thermal comfort starting from environmental and physiological parameters, integrating the measurement of HRV with the measurement of skin temperature. Since the literature has repeatedly stressed that skin temperature is an important predictor for the estimation of the occupant's TSV, it was first decided to estimate the degree of correlation of t_{wrist} with TSV in the context of this research. Pearson correlation coefficient (R) was estimated, and the result is that the average R among all the participants, expressed as mean ± standard deviation (μ±σ), is 63.9 ± 30.6; this value reports that there I medium-to-high linear correlation between TSV and t_{wrist}.

However, LF/HF, which is also reported to be correlated with human thermal comfort, do not have the same correlation with TSV. In fact, in this case R is 14.06% (±29.2%), suggesting a low-to-medium correlation with the TSV. For this reason the LF/HF signal extracted was modified, to examine the trends and evaluate the possible compatibility with the results obtained in the previous studies in the literature. An example of the analysis is displayed in Figure 35, which displays the trend of the manipulated LF/HF, named LF/HF_n.

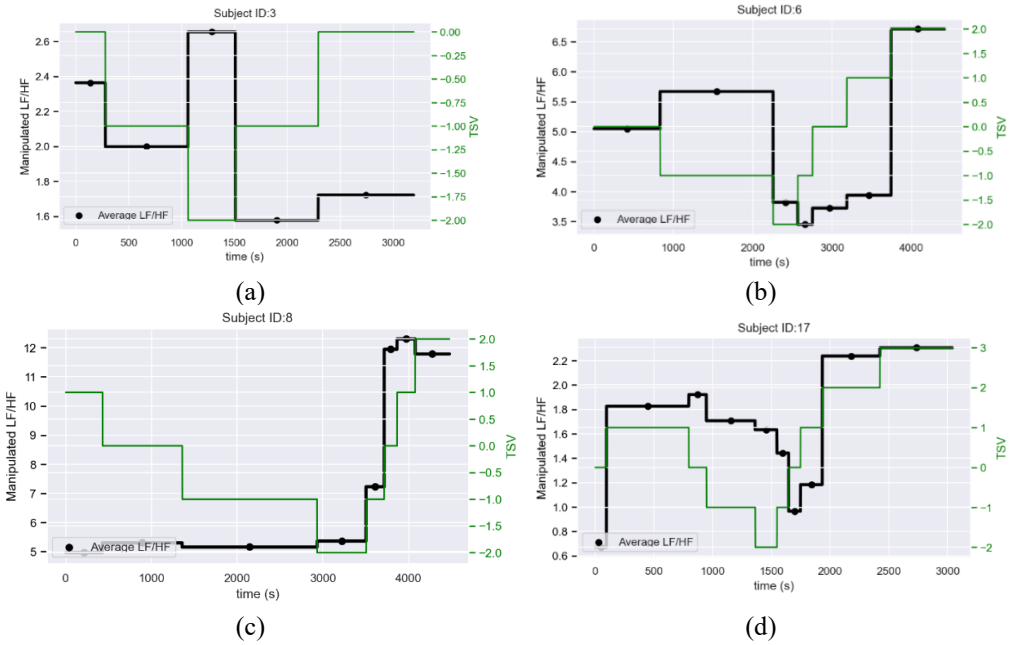


Figure 35. Shape of the manipulated LF/HF against the TSV.

LF/HF_n seems to follow better the trend of the TSV after being processed; to quantify the relationship between TSV and LF/HF_n, Pearson's coefficient for each subject is computed: the average *R* is 20.7% (± 46.2%), which is higher respect to R computed for the original LF/HF, which is 14.06% (±29.2%).

The medium correlation registered between LF/HF_n suggests that this parameter, that is related to the action of the ANS, is deeply related to many phenomena that occurs in human body, that cannot be isolated and excluded in real-life conditions; secondly, it is also worth noting that LF/HF_n represents an average of the original trend of LF/HF computed over a certain period of time; therefore, it should be hypnotized that when this parameter is put in association with human thermal comfort, the analysis that should be considered is not punctual but rather considering a longer period of time. Given the medium correlation with the output TSV, LF/HF_n will be included in the dataset that will train AI algorithms to predict and measure the TSV.

4.4.2 Feature selection

In this section, results from the correlation analysis by means of the Pearson's correlation coefficients (*R*) are displayed. Table 22 shows *R* computed between each parameter included in the initial dataset, and the variable used as ground truth of the AI algorithms, which is the TSV. Coherently with literature, environmental quantities such as t_L, t_M, t_H, t_{r1}, t_{a1} exhibit medium-to-

high correlation with the TSV; on the other hand, as expected in the context of this research, which is conducted by subjecting the participant to a dynamic and transient change of the environmental conditions, linear correlation between the physiological HRV features have a medium to low *R*. However, the most promising correlation with TSV are the one with $\mathbf{t_{skin}}$.

Table 22. Pearson coefficients in percentage computed between the TSV and each feature of the dataset, expressed as mean(μ) ± standard deviation (σ), among all participants.

	μ(%)	σ(%)
***R*(t$_{skin}$ - TSV)**	**63.9**	30.57
R(VLF-TSV)	12.9	27.54
***R*(LF-TSV)**	14.8	27.5
***R*(HF-TSV)**	1.6	31.84
***R*(LF/HF-TSV)**	14.1	30.36
***R*(LF/HF$_n$ -TSV)**	20.4	47.41
***R*(HF/LF-TSV)**	-11.2	27.16
***R*(TP-TSV)**	17.7	30.39
***R*(MEAN-TSV)**	13	43.76
***R*(SDNN-TSV)**	17.1	24.19
***R*(NN50-TSV)**	13.5	35.47
***R*(PNN50-TSV)**	13.5	35.47
***R*(NN20-TSV)**	5	32.8
***R*(PNN20-TSV)**	5	32.8
***R*(RMSSD-TSV)**	11.6	34.73
***R*(MEDIAN-TSV)**	14	43.44
***R*(SD1-TSV)**	11.58	34.73
***R*(SD2-TSV)**	16.38	24.3
***R*(SD1*SD2-TSV)**	16.88	27.17
***R*(t$_{a1}$-TSV)**	**69.32**	32.41
***R*(RH-TSV)**	-42.78	33.99
***R*(v$_{a1}$-TSV)**	-9.551	33.85
***R*(t$_r$-TSV)**	**67.21**	39.65
***R*(t$_H$-TSV)**	**71.1**	29.55
***R*(t$_M$-TSV)**	**70.65**	28.57
***R*(t$_L$-TSV)**	**70.21**	29.13

R computed in Table 22 are the result of the average R coefficient for each participants; the fact that for many parameters there are values of σ comparable with the value of the μ, it suggests that, in addition to not having linear relationships between parameters, there is an intra-subject variability in the relationships that linked the TSV and physiological parameters. Consequently,

the selection of the features carried out by applying the methodology that uses the estimation of the predictor importance, is made on the entire dataset composed of the aggregation of the data of all the participants. In this way, the chosen dataset will represent an average behavior of the subjects, which allows us to take into account both subjective differences and to be able to generalize the applied methodology.

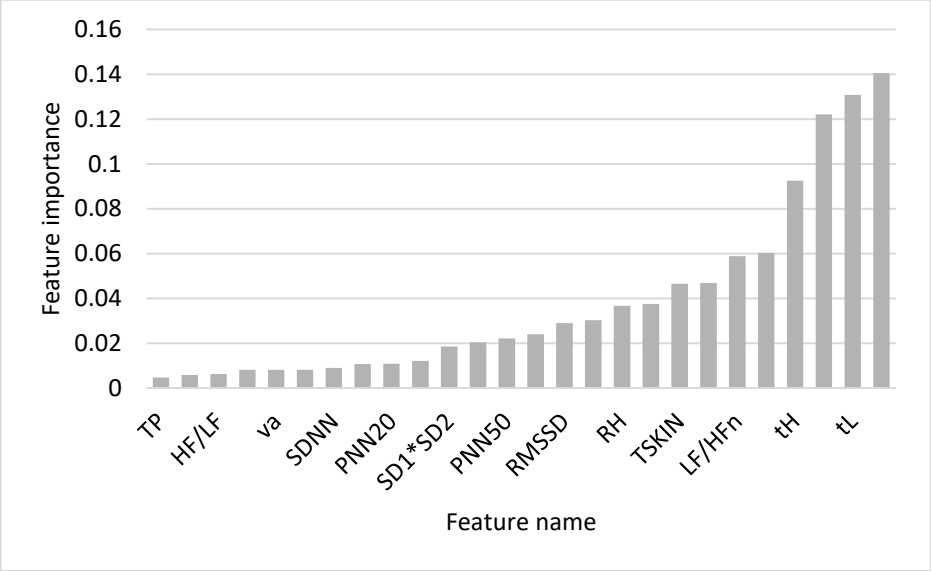


Figure 36. Plot of the importance of each feature of the initial dataset computed by means of the Extra Tree Regression algorithm.

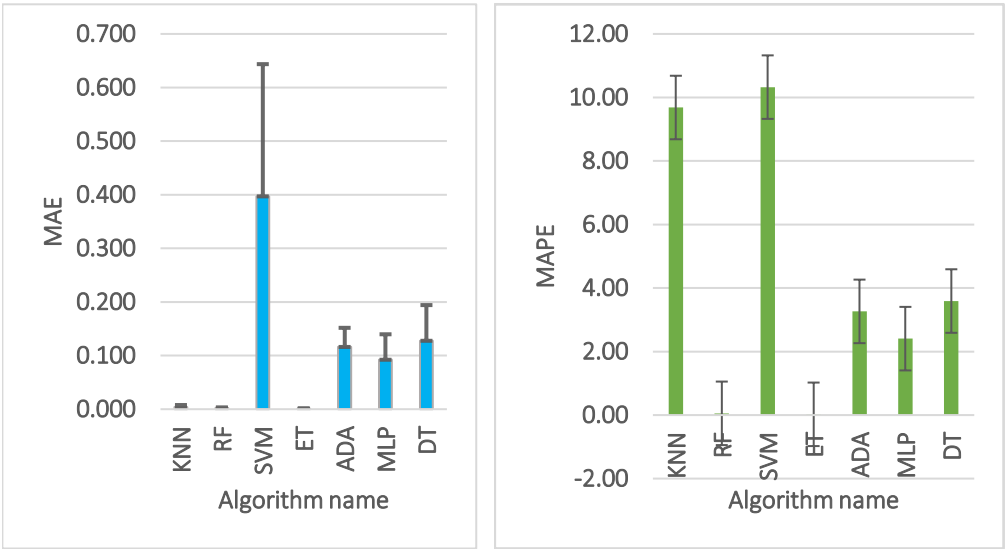
The combination of the results obtained in Figure 36 and Table 22, tells us the common features to consider in the final dataset that will be used for the measurement of TSV. As explained previously, the predictor importance is computed by means of the Gini Importance of each feature: higher values of the Gini index are associated with higher importance of the feature in the prediction of the TSV.

The Pearson correlation coefficients captures the linearity between the environmental parameters and the TSV; therefore, features such as $\mathbf{t_M}$, $\mathbf{t_L}$, $\mathbf{t_{a1}}$, $\mathbf{t_H}$, $\mathbf{t_{r1}}$ can be included in the dataset; however, the predictor importance methodology highlights relationship between the TSV and some physiological features, that should be considered in the dataset; the most contributing features that contribute to the TSV are: t_M , t_L, t_{a1}, t_H, t_{r1}, LF/HF_n, MEDIAN, TSKIN, MEAN,RH. The threshold of the feature importance was set of 0.3.

Moreover, since t_M , t_L, t_{a1}, t_H, t_{r1} are correlated between them with Pearson values higher than 96%, it was chosen to keep only one variable of them, which is t_M that is reported to be the parameter that has the greatest importance, according to Figure 36.

4.4.3 Thermal comfort prediction using AI

In this experiment, the aim was to measure the TSV of each participant who experienced different environmental conditions, by using a dataset made of a small number of features (t_M, t_{skin}, LF/HF_n, MEDIAN, TSKIN, MEAN, RH) . Two approaches were used: the first one consists on creating a PCM, one for each participants, by training and testing the algorithms on data coming from the same subject. Results, illustrated in Figure 37 (a-b), showed that the 3 models provide best average MAE (0.004±0.004, 0.003±0.001 and 0.001±0.0003, for KNN, RF and ET, respectively), while best MAPE was obtained by RF and ET (0.05%±0.034% and 0.02%±0.09%, respectively); thus RF and ET are the models that provide higher performance when predicting TSV by using PCM. These results may arise some concerns due to possible overfitting of the data, even though the 30% of the data were excluded from the training.

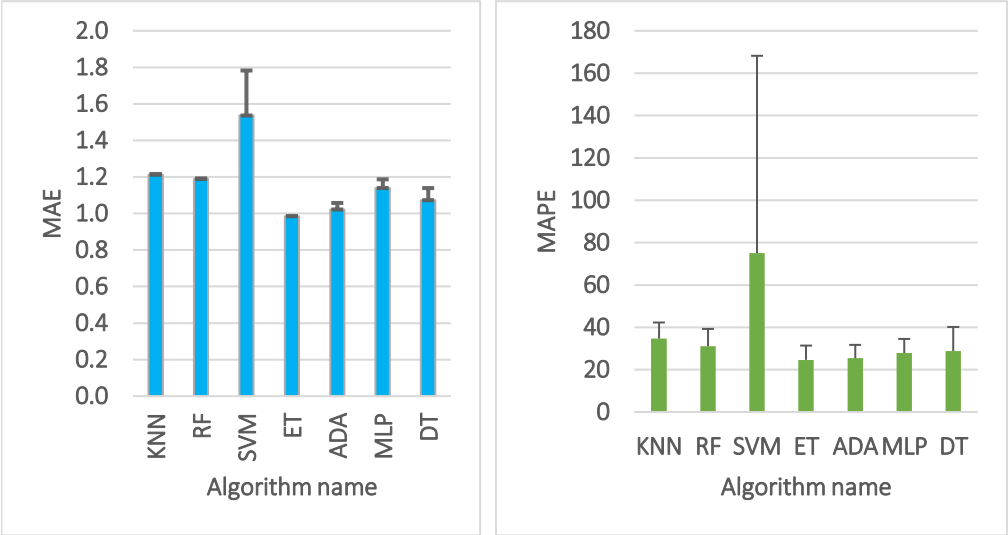


(a) (b)

Figure 37. (a) average MAE computed among MAE obtained from each PCM, one for each participant. (b) average MAPE computed among MAPE obtained from each PCM, one for each participant.

For this reason, the LOSO approach was proposed, which uses a dataset made of data coming from all the participants, except for one which is used as test dataset. Results are illustrated in Figure 38: with this configuration, the best performance in terms of MAE is reached by ET(1±0.2) and the best MAPE was obtained by ET(25%±7%) and ADA(25%±6%). In the first place, it can be noted that the performances of the LOSO approach have decreased, compared to

the results obtained with the PCM. The first reason certainly concerns the different responses of each participant to environmental conditions.



(a) (b)

Figure 38. (a) average MAE computed among MAE obtained using the LOSO approach. (b) average MAPE obtained using the LOSO approach.

4.4.4 Discussion about the integration of skin temperature

The purpose of this research activity is to evaluate whether the addition of the skin temperature within the thermal comfort measurement chain can provide accurate measurements of the TSV in real-life settings. First, it can be noted that the skin temperature acquired through a minimally invasive sensor has a medium-high correlation with the TSV, which expresses thermal comfort. It is worth commenting on the results obtained by applying the linear correlation through the Pearson coefficient and the predictor importance technique; features selection in regression tasks, is a very complex procedure, especially when the features that made up the dataset are not correlated, such as in this case [91]. However, Pearson coefficient (R) is used to find dependencies between the data, and it is a technique that best adapts when the input variables are made of continuous values. This is the explanation behind the high R between t_{skin} and the TSV, and low R between TSV and LF/HF_n. However, this result is not compatible with the output provided by predictor importance method, in which LF/HF_n is reported to have higher importance than t_{skin}; these two quantities are mainly characterized by discrete values and therefore they should not be compared by means of R. Therefore, predictor importance is used in this context as a support tool to extrapolate relationship among the input features and the output.

The conditions of the test, allow the participant to carry out light-office activities, and this introduces additional noise to the HRV signal, and consequently this noise propagates to the

HRV features computation. The research presented in Section 3.2 of this work has shown that the combination of environmental parameters and HRV features can provide predictions of thermal comfort with a MAPE of 20%; however, in this research the user was instructed to limit the movements of the wrist as much as possible, to guarantee an uncertainty of the smartwatch of ±4 ms, as reported in the literature. The study in Section 3.3 on the other hand demonstrated when the user is on the move the uncertainty associated with TSV can increase. For this reason, the introduction of the skin temperature into the measurement set-up was useful and effective to prevent that HRV features impacts negatively on the TSV computation.

Chapter 5.

Conclusion

The overall aim of the current PhD thesis was to develop and validate a methodology to measure the thermal comfort of occupants in indoor environment and address challenges that society is now facing that regards the construction of a satisfactory thermal environment by means of a personalized human-centric approach.

The first study (Section 3.1) is the preliminary research, performed in a highly controlled climatic chamber. The main results that came out from the environmental data analysis, as well as physiological and psychological ones simultaneously monitored during specific tests conducted in winter and summer, can provide with some degree of accuracy the measurement of human thermal comfort expressed in terms of TSV. The analysis of the ECG signal, of 29 out of the 62 performed tests was considered. Extracted features are the ones most related to thermal comfort variation according to the literature, i.e., HRV and LF/HF ratio. ML algorithms have been applied to use the LF/HF and other indices to predict thermal sensation vote. The results confirm that LF/HF alone cannot provide high accuracy on the predictions but by using more HRV indices it is possible to predict human thermal comfort with an accuracy of 82%.

In the second research (Section 3.2) participants were allowed to carry out some light office activities, it was possible to demonstrate how HRV features are only able to discern between comfort and a discomfort condition of the participant and provide lower performance in predicting the TSV in transient conditions compared to the study in the controlled environment. For this reason, the effect of the combination of HRV features with environmental quantities was examined to test whether a combination of the two could help in the TSV's prediction. The evaluation of the performance of the algorithms suggests that features extracted from environmental parameters led to improved results compared to the use of HRV features alone. This result is compliant with literature, which suggests that environmental parameters can be

helpful in the prediction of thermal comfort and that physiological measurements can be used as support features to develop personalized models. Experiment 3 aims to simulate a real-life application in which users are able to perform light office activities while a smartwatch collects their HRV. By using the dataset made by physiological quantities to predict TSV, the MAE and MAPE values cannot provide satisfactory values of accuracy, which suggests that in real life conditions, different from the controlled environment created in Section 3.1, more complex mechanisms are present in the management of HRV. However, a step-forward in the prediction of TSV was made by adding environmental quantities to the HRV features, since the value of MAE and MAPE improved. The research proposed was performed in order to explore whether it is possible to build ML models by collecting physiological data from a smartwatch, providing a methodology that could predict the TSV of the participants. The LOSO approach used for testing the accuracy of the ML algorithms aims at creating a generalized model that does not take into account gender-related differences but considers only TSV as the ground truth of the model. Although the results of the LOSO approach highlight that a punctual, real-time prediction of TSV was not always accurate for the current test, it is important to underline that the results obtained from ML algorithms can be used to predict the overall comfort of a user. This information can point out that the overall thermal comfort is evaluated over a long-term period and it is not necessary to estimate a point-by-point trend.

The Monte Carlo analysis presented in Section 3.2.7 presents an approach to assess the measurement uncertainty of human thermal comfort, expressed in terms of TSV, by using a method that comprises a heterogeneous set of data, made by physiological and environmental quantities, and AI models. The objective is therefore to quantify the measurement uncertainty of the TSV, while the user is performing light-office activities, by using GUM guidelines for applying MCM. A preliminary analysis was conducted to assess the impact of the measurement uncertainty of instrument used to collect HRV, which is a commercial smartwatch. MCM was applied to compute the uncertainty associated with the features extracted from HRV, which will be later fed into a RF and CNN model. Results have shown that among 13 participants, there is uncertainty values in the measurement of features that ranges from ±0.01% to ±0.7%, suggesting that among different users the uncertainty can be generalized. Then, MCM was applied by perturbing a set of parameters (*HRV,* t_a*, RH and* v_a) to compute the uncertainty in measurement of the TSV, using RF model and a CNN. RF has highest uncertainty due to the t_a uncertainty (U = 1%), while CNN have highest uncertainty when *RH* is perturbed (U = 10.5%). On the other hand, the sensitivity analysis that expresses the relationship between the TSV and the input parameters, highlights that v_a is the parameter that causes the greatest variation on the TSV.

Finally, the research presented in Section 3.3 was developed to obtain an improvement in the measurement of TSV, through the inclusion of the skin temperature, that literature has extensively reported being positively associated with thermal comfort. Participants were exposed to changes in the air temperature of the test room, in summer, while they were free to carry out office activities, without constraints in the movements of the arm in which the sensors were installed. This can both negatively affect the HRV measurement which may be subject to motion artifacts, but at the same time it is ensured that the only perturbation of the signal is due to the external conditions of the room and not to emotional factors. Skin temperature inclusion affects positively the computation of the TSV, compensating the uncertainty associated with the

movement of the wrist, since participants were not subjected to particular restrictions of the hand's movements. Results have shown that PCM provides good performance in terms of MAE (0.004±0.004, 0.003±0.001 and 0.001±0.0003, for KNN, RF and ET, respectively), and MAPE obtained by RF and ET (0.05%±0.034% and 0.02%±0.09%, respectively). However, a more generalized approach such LOSO approach that uses a dataset made of data coming from all the participants, except for one which is used as test provides lower performances in terms of MAE (1±0.2 for the ET) and MAPE of (25%±7% and 25%±6%) for ET and ADA, respectively. The proposed set of experiments showed a new measurement procedure to measure the TSV of users in a real and controlled environment, through a vast network of both environmental and physiological sensors.

REFERENCES

[1] B. Yang *et al.*, 'Non-invasive (non-contact) measurements of human thermal physiology signals and thermal comfort/discomfort poses -A review', *Energy and Buildings*, vol. 224, p. 110261, Oct. 2020, doi: 10.1016/j.enbuild.2020.110261.

[2] O. Kaynakli and M. Kilic, 'Investigation of indoor thermal comfort under transient conditions', *Building and Environment*, vol. 40, no. 2, pp. 165–174, Feb. 2005, doi: 10.1016/j.buildenv.2004.05.010.

[3] C. Shan, J. Hu, J. Wu, A. Zhang, G. Ding, and L. X. Xu, 'Towards non-intrusive and high accuracy prediction of personal thermal comfort using a few sensitive physiological parameters', *Energy and Buildings*, vol. 207, p. 109594, Jan. 2020, doi: 10.1016/j.enbuild.2019.109594.

[4] N. Jain *et al.*, 'Building Performance Evaluation of a New Hospital Building in the UK: Balancing Indoor Environmental Quality and Energy Performance', *Atmosphere*, vol. 12, no. 1, Art. no. 1, Jan. 2021, doi: 10.3390/atmos12010115.

[5] H. Shin, M. Kang, S.-H. Mun, Y. Kwak, and J.-H. Huh, 'A study on changes in occupants' thermal sensation owing to CO_2 concentration using PMV and TSV', *Building and Environment*, vol. 187, p. 107413, Jan. 2021, doi: 10.1016/j.buildenv.2020.107413.

[6] A. Kaushik, M. Arif, P. Tumula, and O. J. Ebohon, 'Effect of thermal comfort on occupant productivity in office buildings: Response surface analysis', *Building and Environment*, vol. 180, p. 107021, Aug. 2020, doi: 10.1016/j.buildenv.2020.107021.

[7] M. Tarantini, G. Pernigotto, and A. Gasparella, 'A Co-Citation Analysis on Thermal Comfort and Productivity Aspects in Production and Office Buildings', *Buildings*, vol. 7, no. 2, Art. no. 2, Jun. 2017, doi: 10.3390/buildings7020036.

[8] D. Ormandy and V. Ezratty, 'Thermal discomfort and health: protecting the susceptible from excess cold and excess heat in housing', *Advances in Building Energy Research*, vol. 10, no. 1, pp. 84–98, Jan. 2016, doi: 10.1080/17512549.2015.1014845.

[9] Y. Jiang, Z. Luo, Z. Wang, and B. Lin, 'Review of thermal comfort infused with the latest big data and modeling progresses in public health', *Building and Environment*, vol. 164, p. 106336, Oct. 2019, doi: 10.1016/j.buildenv.2019.106336.

[10] S. Tham, R. Thompson, O. Landeg, K. A. Murray, and T. Waite, 'Indoor temperature and health: a global systematic review', *Public Health*, vol. 179, pp. 9–17, Feb. 2020, doi: 10.1016/j.puhe.2019.09.005.

[11] N. Ma, D. Aviv, H. Guo, and W. W. Braham, 'Measuring the right factors: A review of variables and models for thermal comfort and indoor air quality', *Renewable and Sustainable Energy Reviews*, vol. 135, p. 110436, Jan. 2021, doi: 10.1016/j.rser.2020.110436.

[12] R. Becker and M. Paciuk, 'Thermal comfort in residential buildings – Failure to predict by Standard model', *Building and Environment*, vol. 44, no. 5, pp. 948–960, May 2009, doi: 10.1016/j.buildenv.2008.06.011.

[13] F. Salamone *et al.*, 'Application of IoT and Machine Learning techniques for the assessment of thermal comfort perception.', *Energy Procedia*, vol. 148, pp. 798–805, Aug. 2018, doi: 10.1016/j.egypro.2018.08.130.

[14] 'An experimental design framework for the personalization of indoor microclimates through feedback loops between responsive thermal systems and occupant biometrics - Berardo Matalucci, Kenton Phillips, Alicia A Walf, Anna Dyson, Joshua Draper, 2017'. https://journals.sagepub.com/doi/full/10.1177/14778077117691601 (accessed May 28, 2020).

[15] 'Identification of Users' Well-Being Related to External Stimuli: A Preliminary Investigation | SpringerLink'. https://link.springer.com/chapter/10.1007/978-3-030-04324-7_69 (accessed Jul. 01, 2020).

[16] S. Casaccia, E. J. Sirevaag, E. J. Richter, J. A. O'Sullivan, L. Scalise, and J. W. Rohrbaugh, 'Features of the non-contact carotid pressure waveform: Cardiac and vascular dynamics during rebreathing', *Review of Scientific Instruments*, vol. 87, no. 10, p. 102501, Oct. 2016, doi: 10.1063/1.4964624.

[17] M. Wu, H. Li, and H. Qi, 'Using electroencephalogram to continuously discriminate feelings of personal thermal comfort between uncomfortably hot and comfortable environments', *Indoor Air*, vol. 30, no. 3, pp. 534–543, May 2020, doi: 10.1111/ina.12644.

[18] J. R. Lim, G. H. Baek, and E. S. Jeon, 'Analysis of the Correlation between Thermal Sensations and Brain Waves via EEG Measurements', vol. 13, no. 8, p. 7, 2018.

[19] M. Wu and H. Qi, 'Using passive BCI to online control the air conditioner for obtaining the individual specific thermal comfort *', in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Jul. 2019, pp. 3139–3142. doi: 10.1109/EMBC.2019.8856497.

[20] F. Pietroni, S. Casaccia, G. M. Revel, and L. Scalise, 'Methodologies for continuous activity classification of user through wearable devices: Feasibility and preliminary investigation', in *2016 IEEE Sensors Applications Symposium (SAS)*, Apr. 2016, pp. 1–6. doi: 10.1109/SAS.2016.7479867.

[21] S. Y. Sim *et al.*, 'Estimation of Thermal Sensation Based on Wrist Skin Temperatures', *Sensors*, vol. 16, no. 4, Art. no. 4, Apr. 2016, doi: 10.3390/s16040420.

[22] Y. Fang, Y. Lim, S. E. Ooi, C. Zhou, and Y. Tan, 'Study of Human Thermal Comfort for Cyber–

Physical Human Centric System in Smart Homes', *Sensors*, vol. 20, no. 2, Art. no. 2, Jan. 2020, doi: 10.3390/s20020372.

[23] S. Casaccia, F. Pietroni, A. Calvaresi, G. M. Revel, and L. Scalise, 'Smart Monitoring of User's Health at Home: Performance Evaluation and Signal Processing of a Wearable Sensor for the Measurement of Heart Rate and Breathing Rate:', in *Proceedings of the 9th International Joint Conference on Biomedical Engineering Systems and Technologies*, Rome, Italy, 2016, pp. 175–182. doi: 10.5220/0005694901750182.

[24] Y. S. Can *et al.*, 'Real-Life Stress Level Monitoring Using Smart Bands in the Light of Contextual Information', *IEEE Sensors Journal*, vol. 20, no. 15, pp. 8721–8730, Aug. 2020, doi: 10.1109/JSEN.2020.2984644.

[25] F. Salamone *et al.*, 'Integrated Method for Personal Thermal Comfort Assessment and Optimization through Users' Feedback, IoT and Machine Learning: A Case Study †', *Sensors*, vol. 18, no. 5, Art. no. 5, May 2018, doi: 10.3390/s18051602.

[26] 'Heart Rate Variability: Are You Using It Properly? Standardisation Checklist of Procedures - PubMed'. https://pubmed.ncbi.nlm.nih.gov/30852243/ (accessed May 28, 2020).

[27] H. Zhu, H. Wang, D. Li, Z. Xiao, H. Su, and X. Kuang, 'Evaluation of the Human Thermal Comfort under Simulated Weightlessness: an Experimental Study Based on the Power Spectrum Analysis of the Heart Rate Variability', *Microgravity Sci. Technol.*, vol. 31, no. 1, pp. 73–83, Feb. 2019, doi: 10.1007/s12217-018-9669-7.

[28] H. Zhu, H. Wang, Z. Liu, D. Li, G. Kou, and C. Li, 'Experimental study on the human thermal comfort based on the heart rate variability (HRV) analysis under different environments', *Sci. Total Environ.*, vol. 616–617, pp. 1124–1133, Mar. 2018, doi: 10.1016/j.scitotenv.2017.10.208.

[29] N. Pivac, S. Nižetić, V. Zanki, and A. M. Papadopoulos, 'Application of wearable sensory devices in predicting occupant's thermal comfort in office buildings during the cooling season', *IOP Conf. Ser.: Earth Environ. Sci.*, vol. 410, p. 012092, Jan. 2020, doi: 10.1088/1755-1315/410/1/012092.

[30] K. Nkurikiyeyezu and G. Lopez, 'Toward a real-time and physiologically controlled thermal comfort provision in office buildings', p. 10.

[31] W. Liu, Z. Lian, and Y. Liu, 'Heart rate variability at different thermal comfort levels', *Eur J Appl Physiol*, vol. 103, no. 3, pp. 361–366, Jun. 2008, doi: 10.1007/s00421-008-0718-6.

[32] F. Akbar, G. Mark, I. Pavlidis, and R. Gutierrez-Osuna, 'An Empirical Study Comparing Unobtrusive Physiological Sensors for Stress Detection in Computer Work', *Sensors*, vol. 19, no. 17, Art. no. 17, Jan. 2019, doi: 10.3390/s19173766.

[33] J. Lee and Y. Ham, 'Physiological sensing-driven personal thermal comfort modelling in consideration of human activity variations', *Building Research & Information*, vol. 0, no. 0, pp. 1–13, Nov. 2020, doi: 10.1080/09613218.2020.1840328.

[34] N. Morresi, S. Casaccia, M. Sorcinelli, M. Arnesano, and G. M. Revel, 'Analysing performances of Heart Rate Variability measurement through a smartwatch', *2020 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*.

[35] N. Morresi, S. Casaccia, and G. M. Revel, 'Metrological characterization and signal processing of a wearable sensor for the measurement of heart rate variability', in *2021 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, Jun. 2021, pp. 1–6. doi: 10.1109/MeMeA52024.2021.9478713.

[36] T. Chaudhuri, D. Zhai, Y. C. Soh, H. Li, L. Xie, and X. Ou, 'Convolutional Neural Network and Kernel Methods for Occupant Thermal State Detection using Wearable Technology', in *2018 International Joint Conference on Neural Networks (IJCNN)*, Jul. 2018, pp. 1–8. doi: 10.1109/IJCNN.2018.8489069.

[37] W. Li, J. Zhang, T. Zhao, and J. Ren, 'Experimental study of an indoor temperature fuzzy control method for thermal comfort and energy saving using wristband device', *Building and Environment*, vol. 187, p. 107432, Jan. 2021, doi: 10.1016/j.buildenv.2020.107432.

[38] P. Wei, Y. Liu, H. Kang, C. Yang, and X. Jiang, 'A Low-Cost and Scalable Personalized Thermal Comfort Estimation System in Indoor Environments', in *Proceedings of the First International Workshop on Cyber-Physical-Human System Design and Implementation*, New York, NY, USA, May 2021, pp. 1–6. doi: 10.1145/3458648.3460006.

[39] 'Machine learning method for real-time non-invasive prediction of individual thermal preference in transient conditions - ScienceDirect'. https://www.sciencedirect.com/science/article/pii/S0360132318307121 (accessed Nov. 07, 2021).

[40] F. Kobiela, R. Shen, M. Schweiker, and R. Dürichen, 'Personal thermal perception models using skin temperatures and HR/HRV features: comparison of smartwatch and professional measurement devices', in *Proceedings of the 23rd International Symposium on Wearable Computers*, London, United Kingdom, Sep. 2019, pp. 96–105. doi: 10.1145/3341163.3347737.

[41] K. Nkurikiyeyezu, A. Yokokubo, and G. Lopez, 'Affect-aware thermal comfort provision in intelligent buildings', in *2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*, Sep. 2019, pp. 331–336. doi: 10.1109/ACIIW.2019.8925184.

[42] L. V. Coutts, D. Plans, A. W. Brown, and J. Collomosse, 'Deep learning with wearable based heart rate variability for prediction of mental and general health', *Journal of Biomedical Informatics*, vol. 112, p. 103610, Dec. 2020, doi: 10.1016/j.jbi.2020.103610.

[43] W. Taylor, S. A. Shah, K. Dashtipour, A. Zahid, Q. H. Abbasi, and M. A. Imran, 'An Intelligent Non-Invasive Real-Time Human Activity Recognition System for Next-Generation Healthcare', *Sensors*, vol. 20, no. 9, Art. no. 9, Jan. 2020, doi: 10.3390/s20092653.

[44] J. Malik, Y.-L. Lo, and H. Wu, 'Sleep-wake classification via quantifying heart rate variability by convolutional neural network', *Physiol. Meas.*, vol. 39, no. 8, p. 085004, Aug. 2018, doi: 10.1088/1361-6579/aad5a9.

[45] G. Giannakakis, E. Trivizakis, M. Tsiknakis, and K. Marias, 'A novel multi-kernel 1D convolutional neural network for stress recognition from ECG', in *2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*, Sep. 2019, pp. 1–4. doi: 10.1109/ACIIW.2019.8925020.

[46] S. Haghayegh, S. Khoshnevis, M. H. Smolensky, K. R. Diller, and R. J. Castriotta, 'Deep Neural Network Sleep Scoring Using Combined Motion and Heart Rate Variability Data', *Sensors*, vol. 21, no. 1, Art. no. 1, Jan. 2021, doi: 10.3390/s21010025.

[47] R. F. Rupp, N. G. Vásquez, and R. Lamberts, 'A review of human thermal comfort in the built environment', *Energy and Buildings*, vol. 105, pp. 178–205, Oct. 2015, doi: 10.1016/j.enbuild.2015.07.047.

[48] S. Van Craenendonck, L. Lauriks, C. Vuye, and J. Kampen, 'A review of human thermal comfort experiments in controlled and semi-controlled environments', *Renewable and Sustainable Energy Reviews*, vol. 82, pp. 3365–3378, Feb. 2018, doi: 10.1016/j.rser.2017.10.053.

[49] 'ISO. ISO Standard 7726. Ergonomics of the Thermal Environment—Instruments for Measuring Physical Quantities; International Organization for Standardization: Geneva, Switzerland, 1998'.

[50] 'UNI EN ISO 7730:2006'. http://store.uni.com/catalogo/uni-en-iso-7730-2006 (accessed May 28, 2020).

[51] 'Bioharness™ Multivariable Monitoring Device: Part. I: Validity'. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3737934/ (accessed Apr. 08, 2021).

[52] 'Ekanayake, Hiran. "P300 and Emotiv EPOC: Does Emotiv EPOC capture real EEG?." Web publication http://neurofeedback. visaduma. info/emotivresearch. htm 133 (2010).'.

[53] 'BITalino: A Novel Hardware Framework for Physiological Computing':, in *Proceedings of the International Conference on Physiological Computing Systems*, Lisbon, Portugal, 2014, pp. 246–253. doi: 10.5220/0004727802460253.

[54] 'ISO 10551:1995–Ergonomics of the Thermal Environment - Assessment of the Influence of the Thermal Environment Using Subjective Judgement Scales'.

[55] Z. Qavidel Fard, Z. S. Zomorodian, and S. S. Korsavi, 'Application of machine learning in thermal comfort studies: A review of methods, performance and challenges', *Energy and Buildings*, vol. 256, p. 111771, Feb. 2022, doi: 10.1016/j.enbuild.2021.111771.

[56] 'ISO 17772-1:2017-Energy Performance of Buildings – Indoor Environmental Quality – Part 1: Indoor Environmental Input Parameters for the Design and Assessment of Energy Performance of Buildings." .'.

[57] 'ASHRAE-ANSI-ASHRAE Standard 62.1-2016 Ventilation for Acceptable Indoor Air Quality; ASHRAE: Atlanta, GA, USA, 2016'.

[58] J. A. Gomes, S. L. Winters, D. Stewart, A. Targonski, and P. Barreca, 'Optimal bandpass filters for time-domain analysis of the signal-averaged electrocardiogram', *The American Journal of Cardiology*, vol. 60, no. 16, pp. 1290–1298, Dec. 1987, doi: 10.1016/0002-9149(87)90610-2.

[59] 'A Real-Time QRS Detection Algorithm | IEEE Journals & Magazine | IEEE Xplore'. https://ieeexplore.ieee.org/abstract/document/4122029?casa_token=OcYZKlyTOHoAAAAA:go14bYuHT2j_eqazVHgQLPYswTttFacaiQOFyKuSg7ZCR-DNM9ju1kK7uKH9Go-n1o2oNjE (accessed Nov. 05, 2021).

[60] T. F. of the E. S. of C. the N. A. S. of P. Electrophysiology, 'Heart Rate Variability', *Circulation*, vol. 93, no. 5, pp. 1043–1065, Mar. 1996, doi: 10.1161/01.CIR.93.5.1043.

[61] R. Garay *et al.*, 'Energy Efficiency Achievements in 5 Years Through Experimental Research in KUBIK', *Energy Procedia*, vol. 78, pp. 865–870, Nov. 2015, doi: 10.1016/j.egypro.2015.11.009.

[62] 'ISO. ISO Standard 9920. Ergonomics of the Thermal Environment—Estimation of the Thermal Insulation and Evaporative Resistance of a Clothing Ensemble; International Organization for Standardization: Geneva, Switzerland, 2009.'

[63] '(PDF) Thermal Comfort and Stress Recognition in Office Environment'. https://www.researchgate.net/publication/330754493_Thermal_Comfort_and_Stress_Recognition_in_Office_Environment (accessed May 28, 2020).

[64] K. Nkurikiyeyezu, A. Yokokubo, and G. Lopez, 'Importance of Individual Differences in Physiological-Based Stress Recognition Models', in *2019 15th International Conference on Intelligent Environments (IE)*, Jun. 2019, pp. 37–43. doi: 10.1109/IE.2019.00006.

[65] I. Pigliautile, S. Casaccia, N. Morresi, M. Arnesano, A. L. Pisello, and G. M. Revel, 'Assessing occupants' personal attributes in

relation to human perception of environmental comfort: Measurement procedure and data analysis', *Building and Environment*, vol. 177, p. 106901, Jun. 2020, doi: 10.1016/j.buildenv.2020.106901.

[66] A. Jovic, K. Brkic, and G. Krstacic, 'Detection of congestive heart failure from short-term heart rate variability segments using hybrid feature selection approach', *Biomedical Signal Processing and Control*, vol. 53, p. 101583, Aug. 2019, doi: 10.1016/j.bspc.2019.101583.

[67] A. Ferrero and S. Salicone, 'A Comparison Between the Probabilistic and Possibilistic Approaches: The Importance of a Correct Metrological Information', *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 3, pp. 607–620, Mar. 2018, doi: 10.1109/TIM.2017.2779346.

[68] A. Fornaser, M. D. Cecco, P. Bosetti, T. Mizumoto, and K. Yasumoto, 'Sigma-zrandom forest, classification and confidence', *Meas. Sci. Technol.*, vol. 30, no. 2, p. 025002, Dec. 2018, doi: 10.1088/1361-6501/aaf466.

[69] M. Khanafer and S. Shirmohammadi, 'Applied AI in instrumentation and measurement: The deep learning revolution', *IEEE Instrumentation Measurement Magazine*, vol. 23, no. 6, pp. 10–17, Sep. 2020, doi: 10.1109/MIM.2020.9200875.

[70] S. Casaccia, G. M. Revel, G. Cosoli, and L. Scalise, 'Assessment of Domestic Well-Being: From Perception to Measurement', *IEEE Instrumentation Measurement Magazine*, vol. 24, no. 6, pp. 58–67, Sep. 2021, doi: 10.1109/MIM.2021.9513641.

[71] J. Singh, L. A. Kumaraswamidhas, N. Bura, and N. Dilawar Sharma, 'A Monte Carlo simulation investigation on the effect of the probability distribution of input quantities on the effective area of a pressure balance and its uncertainty', *Measurement*, vol. 172, p. 108853, Feb. 2021, doi: 10.1016/j.measurement.2020.108853.

[72] 'Evaluation of measurement data—guide for the expression of uncertainty in measurement, JCGM 100: 2008'.

[73] A. Chen and C. Chen, 'Comparison of GUM and Monte Carlo methods for evaluating measurement uncertainty of perspiration measurement systems', *Measurement*, vol. 87, pp. 27–37, Jun. 2016, doi: 10.1016/j.measurement.2016.03.007.

[74] V. Isaiev and O. Velychko, 'Metrological characterisation of current transformers calibration unit for accurate measurement', *ACTA IMEKO*, vol. 10, pp. 6–13, Jun. 2021, doi: 10.21014/acta_imeko.v10i2.918.

[75] F. Ezedine, J.-M. Linares, W. M. W. Muhamad, and J.-M. Sprauel, 'Identification of most influential factors in a Virtual Reality tracking system using hybrid method', *ACTA IMEKO*, vol. 2, no. 2, Art. no. 2, Jan. 2014, doi: 10.21014/acta_imeko.v2i2.136.

[76] T. T. Nguyen, A. Amthor, and C. Ament, 'Algorithm for a high precision contactless measurement system', *ACTA IMEKO*, vol. 2, no. 2, Art. no. 2, Jan. 2014, doi: 10.21014/acta_imeko.v2i2.82.

[77] 'Sensitivity Analysis for Importance Assessment - Saltelli - 2002 - Risk Analysis - Wiley Online Library'. https://onlinelibrary.wiley.com/doi/full/10.1111/0272-4332.00040 (accessed Jun. 15, 2021).

[78] O. Sima and M.-C. Lépy, 'Application of GUM Supplement 1 to uncertainty of Monte Carlo computed efficiency in gamma-ray spectrometry', *Applied Radiation and Isotopes*, vol. 109, pp. 493–499, Mar. 2016, doi: 10.1016/j.apradiso.2015.11.097.

[79] A. Chen and C. Chen, 'Comparison of GUM and Monte Carlo methods for evaluating measurement uncertainty of perspiration measurement systems', *Measurement*, vol. 87, pp. 27–37, Jun. 2016, doi: 10.1016/j.measurement.2016.03.007.

[80] N. Morresi *et al.*, 'Sensing physiological and environmental quantities to measure human thermal comfort through Machine Learning techniques', *IEEE Sensors Journal*, pp. 1–1, 2021, doi: 10.1109/JSEN.2021.3064707.

[81] G. M. Revel, E. Sabbatini, and M. Arnesano, 'Development and experimental evaluation of a thermography measurement system for real-time monitoring of comfort and heat rate exchange in the built environment', *Meas. Sci. Technol.*, vol. 23, no. 3, p. 035005, Feb. 2012, doi: 10.1088/0957-0233/23/3/035005.

[82] B. Salehi, A. H. Ghanbaran, and M. Maerefat, 'Intelligent models to predict the indoor thermal sensation and thermal demand in steady state based on occupants' skin temperature', *Building and Environment*, vol. 169, p. 106579, Feb. 2020, doi: 10.1016/j.buildenv.2019.106579.

[83] S. Lee, P. Karava, A. Tzempelikos, and I. Bilionis, 'A smart and less intrusive feedback request algorithm towards human-centered HVAC operation', *Building and Environment*, vol. 184, p. 107190, Oct. 2020, doi: 10.1016/j.buildenv.2020.107190.

[84] W. Liu, X. Tian, D. Yang, and Y. Deng, 'Evaluation of individual thermal sensation at raised indoor temperatures based on skin temperature', *Building and Environment*, vol. 188, p. 107486, Jan. 2021, doi: 10.1016/j.buildenv.2020.107486.

[85] S. Liu, S. Schiavon, H. P. Das, M. Jin, and C. J. Spanos, 'Personal thermal comfort models with wearable sensors', *Building and Environment*, vol. 162, p. 106281, Sep. 2019, doi: 10.1016/j.buildenv.2019.106281.

[86] H. Wang and L. Liu, 'Experimental investigation about effect of emotion state on people's thermal comfort', *Energy and Buildings*, vol. 211, p. 109789, Mar. 2020, doi: 10.1016/j.enbuild.2020.109789.

[87] K. Okamoto-Mizuno and K. Tsuzuki, 'Effects of season on sleep and skin temperature in the elderly', *Int J Biometeorol*, vol. 54, no. 4, pp. 401–409, Jul. 2010, doi: 10.1007/s00484-009-0291-7.

[88] K. M. Dalmeida and G. L. Masala, 'HRV Features as Viable Physiological Markers for Stress Detection Using Wearable Devices', *Sensors*, vol. 21, no. 8, Art. no. 8, Jan. 2021, doi: 10.3390/s21082873.

[89] 'Physiological response to typical temperature step-changes in winter of China - ScienceDirect'. https://www.sciencedirect.com/science/article/pii/S037877881631948X?casa_token=sQNwo72dFkkAAAAA:5IOzPv2mQOTSToWgFERdMsxgt344KpEkQxbR9QHlwmHh8yf0D1CZQlwRFMT3b3f2OlwBKWnr (accessed Nov. 05, 2021).

[90] A. Aryal and B. Becerik-Gerber, 'Thermal comfort modeling when personalized comfort systems are in use: Comparison of sensing and learning methods', *Building and Environment*, vol. 185, p. 107316, Nov. 2020, doi: 10.1016/j.buildenv.2020.107316.

[91] M. Baak, R. Koopman, H. Snoek, and S. Klous, 'A new correlation coefficient between categorical, ordinal and interval variables with Pearson characteristics', *Computational Statistics & Data Analysis*, vol. 152, p. 107043, Dec. 2020, doi: 10.1016/j.csda.2020.107043.