

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/349931719>

Sensing physiological and environmental quantities to measure human thermal comfort through Machine Learning techniques

Article in IEEE Sensors Journal · March 2021

DOI: 10.1109/JSEN.2021.3064707

CITATIONS

3

READS

79

7 authors, including:



Nicole Morresi

Università Politecnica delle Marche

6 PUBLICATIONS 35 CITATIONS

[SEE PROFILE](#)



Sara Casaccia

Università Politecnica delle Marche

50 PUBLICATIONS 270 CITATIONS

[SEE PROFILE](#)



M. Arnesano

Università Telematica "E-Campus"

46 PUBLICATIONS 381 CITATIONS

[SEE PROFILE](#)



Amaia Uriarte

Tecnalia

17 PUBLICATIONS 123 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



P2Endure [View project](#)

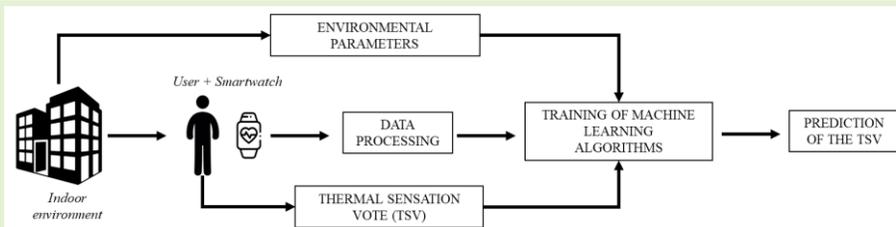


AAL eWare [View project](#)

Sensing physiological and environmental quantities to measure human thermal comfort through Machine Learning techniques

N. Morresi, S. Casaccia, M. Sorcinelli, M. Arnesano, A. Uriarte, J. I. Torrens-Galdiz and G. M. Revel

Abstract— This paper presents the results from the experimental application of smartwatch sensors to predict occupants' thermal comfort under varying environmental conditions. The goal is to investigate the measurement accuracy of smartwatches when used as thermal comfort sensors to be integrated into Heating, Ventilation and Air Conditioning (HVAC) control loops. Ten participants were exposed to various environmental conditions as well as warm - induced and cold-induced discomfort tests and 13 participants were exposed to a transient-condition while a network of sensors and a smartwatch collected both environmental parameters and heart rate variability (HRV). HRV features were used as input to Machine Learning (ML) classification algorithms to establish whether a user was in discomfort, providing an average accuracy of 92.2 %. ML and Deep Learning regression algorithms were trained to predict the thermal sensation vote (TSV) in a transient environment and the results show that the aggregation of environmental and physiological quantities provide a better TSV prediction in terms of Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE), 1.2 and 20% respectively, than just the HRV features used for the prediction. In conclusion, this experiment supports the assumption that physiological quantities related to thermal comfort can improve TSV prediction when combined with environmental quantities.



Index Terms— Thermal comfort; environmental control; human perception; thermal sensation vote; wearable sensors; heart rate variability

I. INTRODUCTION

THE preservation of a comfortable thermal environment in buildings is necessary to guarantee a pleasant experience from occupants' point of view. Indoor environmental quality (IEQ) has a massive impact on occupants' daily life in terms of health, comfort and well-being, since it is reported that an average up to 90% of the total time of an individual is spent in indoor environments [1]. An inadequate thermal environment produces several negative effects on occupants, including, for example, lower work productivity, and consequently thermal satisfaction helps to improve occupants' performance [2], [3]. Thermal comfort is a topic of interest in many types of facilities like educational, office and residential buildings. These places require a great amount of energy to guarantee occupants a satisfactory thermal environment. Moreover, indoor thermal discomfort also affects human health and can be particularly critical in case of susceptible subjects [4]–[6]. 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. A 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. The index currently used for the evaluation of thermal comfort conditions in mechanically controlled environments is the Predicted Mean Vote (PMV). In addition, PMV is deployed for evaluating the energetic performances related to indoor environmental quantities, as reported in European Standards EN 16798-1, EN 16798-2 [7], [8]. PMV is an index whose aim is to forecast the mean value of thermal sensation votes of a group of occupants. The PMV model involves two human-related quantities, i.e. users' clothing insulation and metabolic rate, combined with environmental quantities, i.e. temperature, air velocity, mean radiant temperature and relative humidity. Recent studies are highlighting several limitations of the PMV model when deployed in real conditions. Firstly, the model was designed to predict the average comfort of a large group of users, therefore it exhibits low predictive accuracy if applied to a small group of users; secondly, even if there are Standards in the field of the Ergonomics of the Thermal Environment that specify how the quantities for the assessment of PMV can be measured, it can happen that in real-world settings it is very difficult to assess the exact value of the input variables [9]–[13]. Literature provides studies that indicate how to perform the

measurement of such parameters to reduce measurement uncertainty of PMV [14]. However, the complexity in the measurement of PMV derives from the fact that when environmental monitoring is done in environments experienced daily by the occupants, these quantities may vary and can be difficult to have a real-time and punctual control.

Metabolic rate can change during the day, and one possible solution could be overcome by the usage of smartwatches or smart bands that can measure the level of activity of the occupant; moreover, in real-life context, the clothing insulation may not be constant in time, leading to inaccuracies in the measurement of PMV.

In addition, it can occur that, things being equal, different occupants may have different subjective responses [15]–[17]. So far, the scientific community has investigated thermal comfort starting from the single individual and its perception of the environment, since thermal comfort is deeply related to behavioral, physiological and psychological factors and, as a result, it differs from one individual to another [13]. Human thermal comfort is controlled 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 [14],[15]. Since this process is managed by involuntary mechanisms that take place in the hypothalamus, recent studies are focusing on estimating thermal comfort of individuals through the analysis of specific human physiological signals.

A. Sensors for the measurement of physiological signal related to thermal comfort

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. In fact, 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 to 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 [21]–[24].

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 [25]. The user's thermal sensation can be correlated to indices obtained from EEG. Given this assumption, there are researches that are developing systems based on the brain-computer interface (BCI) for the control of conditioning system to obtain optimal thermal comfort conditions [26]. However, EEG sensors are

expensive and are not suitable for daily monitoring, since the user manages to wear them for up to three hours [27].

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 [28]–[31]. They are also less expensive than EDA, EEG and ECG technologies. For this reason, in this paper a smartwatch was used to evaluate and measure human thermal comfort, since this particular device better reflects the necessity of having a non-intrusive and cost-effective sensor to measure both HR and HRV with suitable accuracy [32]. This paper aims to present a methodology that employs smartwatches for evaluating the possibility of measuring human thermal comfort, expressed by the TSV, in response to a variation of the environmental conditions.

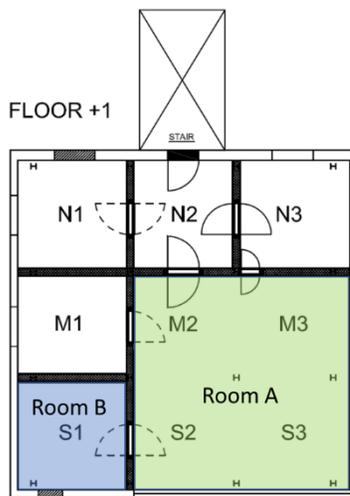
B. Measurement of Heart rate variability for thermal comfort

HRV is obtained by measuring the difference in time between two consecutive heartbeats of the electrocardiogram (ECG) signal. However, wearable technologies today have made it possible to retrieve HRV from less invasive equipment. In fact, wearable technologies can measure the photoplethysmography signal (PPG) in which a heartbeat is individuated when a systolic peak occurs. The difference in time between two consecutive systolic peaks provides HRV traces [33].

This paragraph contains a description of the most recent works related to the measurement of HRV in relation to different environmental conditions. Generally, HRV is analyzed in terms of time-domain and frequency-domain parameters after using specific processing of data [34], [35]. 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 [36]. 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



(a)



(b)

Fig. 1. (a) External appearance of the Kubik facility. (b) Planimetry of the floor in the facility where the tests were performed.

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, [37] recruited 6 participants to explore the relationship between HRV features and different environmental conditions; also [38] in their work, predicts the user's thermal comfort states using HRV signal collected from

a smartband, using a sample size of 6 participant. However, the sample size in thermal comfort studies can be also be reduced, as it is reported in [39] and [40], which tested their solution on a sample size of 4 and 1, respectively.

There are thermal comfort studies that consider age-related differences: higher temperatures are preferred by elderly people, while lower temperatures are more suitable for the younger population [16]. Gender-related differences are another important aspect that is taken into consideration in the thermal comfort field: in fact, women are reported to be more susceptible to temperature fluctuations and are generally more dissatisfied than men in relation to thermal environment. Considering the work conducted in [41], it has also been found that race-related differences do not play a significant role in the perception of thermal comfort.

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 [30]. It has also been demonstrated that LF/HF is influenced by the psychological state of the user: in [42] it is reported that LF/HF rises when the user is in a bored state, while it decreases when the user is in a joyful state.

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 [37], [38], [43]. To collect HRV data, different sensors are used in current literature: ECG is still the predominant methodology in many researches [30], [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 [35], [40], [44]. In this perspective, HRV-related works are developed through the use of commercially available smartwatches, which are less expensive and provide



Fig. 2. Picture of room B used to create a discomfort condition for the participants.

TABLE I

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		

reliable accuracy [45].

In addition, most of these works have highlighted that Machine Learning (ML) approaches produce relevant results when physiological and environmental parameters are merged to forecast occupants' thermal sensation. In this work [46], 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 [47] 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

TABLE II

TECHNICAL INFORMATION OF THE SENSORS EMPLOYED TO COLLECT ENVIRONMENTAL QUANTITIES INSIDE TEST ROOM B

Sensor	Manufacturer	Model	Accuracy
Air Temperature	Thermo Sensor GmbH	PT100 (4 wired) 2113-1-074	± 0.1 °C
Relative humidity	Ahlborn	FHAD46C41A	±2% of reading
Air velocity	Ahlborn	FVAD05TOK300	±1% of reading
Globe temperature	Ahlborn	FPA805GTS	± 0.1 °C

[48] uses ML regression models that produce an RMSE value up to 0.04±0.01. HRV indices and ML algorithms are also at the basis of the work presented in [49], which demonstrated the possibility of estimating the thermal sensation vote of each individual starting from HRV features extracted from an ECG signal with an accuracy up to 82%.

To the best of the authors' knowledge, in literature there is a lack of studies that employ Deep Learning algorithms for thermal comfort prediction starting from HRV signals. However, there are works that show various fields of application of HRV used to forecast some user-related physiological or psychological states. [50], [51] 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. [52] uses a novel unified deep learning framework for sleep-wake classification with two heterogeneous sensors which includes acceleration and HRV.

In [52] 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 [53] 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%. [54] also uses HRV as input to train a CNN for the prediction of sleep scoring.

This work was developed as part of a research that wants to address the possibility of including physiological parameters in the prediction of human thermal comfort. The results of [49] have laid the foundations for the development of the research activity here presented and the tuning of the methodology according to its purpose. The experiment conducted in [49] pointed out that it is possible to use physiological parameters such as LF/HF to estimate the TSV of the user. It is important to point out that the experiment was conducted in a highly controlled environment, showing that physiological quantities can lead to the prediction of TSV with an accuracy of up to 82%, and that the personal characteristics of each subject (such as age, gender, birthplace, BMI) do not

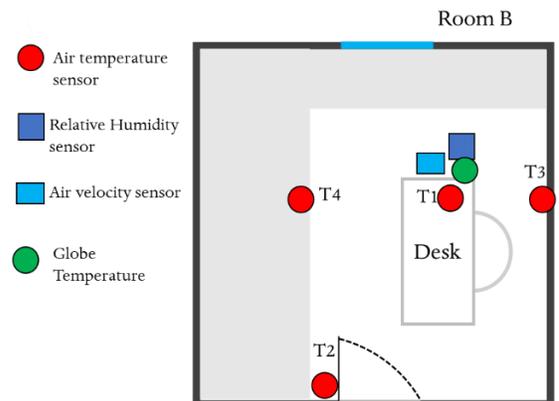


Fig. 3. Measurement set-up in Room B including furniture and sensor location.

lead to statistically significant results.

This work shows that in less controlled condition closer to reality, while the participant performs light office activities, physiological parameters provide lower performance in the estimation of the TSV, in contrast to the previous work, but adding environmental parameters can lead to an improvement in the prediction of the TSV. The personal characteristics of the user (e.g. gender, BMI) are not included in this research since, [49] has shown that the analysis of environmental variables and personal characteristics in the prediction of user thermal comfort, does not lead to statistically significant results.

In contrast with the previous work, the activity here presented shows that in a less controlled environment, closer to reality, while participants carry out light office activities, physiological parameters provide lower performance in the estimation of TSV, but the addition of environmental parameters can lead to an improvement in its prediction.

The experimental setup for the acquisition of physiological quantities was reduced to a minimally-invasive smartwatch that collects HRV data and participants are free to perform light office activities. By reducing the level of discomfort caused by external stimuli different from thermal change, the HRV signals produced are influenced only by environmental discomfort. In the previous work, the sample 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. In addition, 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 proof to provide better performance. In a transient environment, it is preferable to use ML regression algorithms rather than classification algorithms, since TSV is manipulated differently.

Given the above, this work 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 particular, the methodology adopted in this work 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.

II. MATERIALS AND METHODS

This section describes the experimental procedure developed to analyze human thermal perception in response to different external temperatures. The test took place in KUBIK (Bilbao, Spain), Fig 1 (a), 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 [55]. 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 Fig 1 (b).

The tests consisted of three trials: 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, Fig 2, has three outdoor exposed elements, i.e. rooftop, west and south façades.

A. 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 1. 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 1), 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} .

All the participants gave written informed consent to use their personal data and were duly informed about the goal of the research.

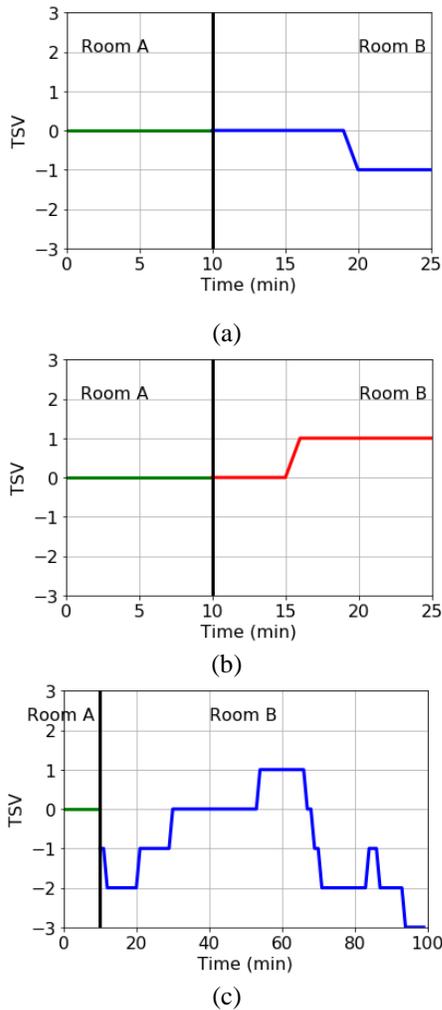


Fig. 4. (a) TSV trend during Experiment 1 for one participant. (b) TSV trend during Experiment 2 for one participant. (c) TSV trend during Experiment 3 for one participant.

B. Experimental set-up

Table 2 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 Fig 3.

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 (T2, T3, T4) were mounted on the perimetral walls of the room, at 1.4 m from the ground (Fig 3). Thermocouple T1 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 (T1) was considered [56].

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

TABLE III
DESCRIPTION OF THE THREE EXPERIMENTS CONDUCTED AND THE AVERAGE VALUE OF THE OUTDOOR PARAMETERS

	Room	Duration	Indoor Temperature		Outdoor Environment	
			Temperature			
Exp. 1	A	10 min	Neutral (19 - 21 °C)		Temperature	7.2 °C
	B	15 min	Cold (15°C)		Wind	5 MPH
Exp. 2	A	10 min	Neutral (19 - 21 °C)		Dew Point	5.8 °C
	B	15 min	Cold (26°C)		Temperature	11.5 °C
Exp. 3	A	10 min	Neutral (19 - 21 °C)		Wind	13 MPH
	B	5 min	15°C		Dew Point	5.4 °C
		5 min	up to 26°C		Wind	10 MPH
		5 min	26 °C		Dew Point	5.8 °C
			down to 15°C			

every 5 seconds. All the data collected were sent to an acquisition board that saves data locally.

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.

C. 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. Fig 4 shows an example of the TSV collected in each experiment for a single participant.

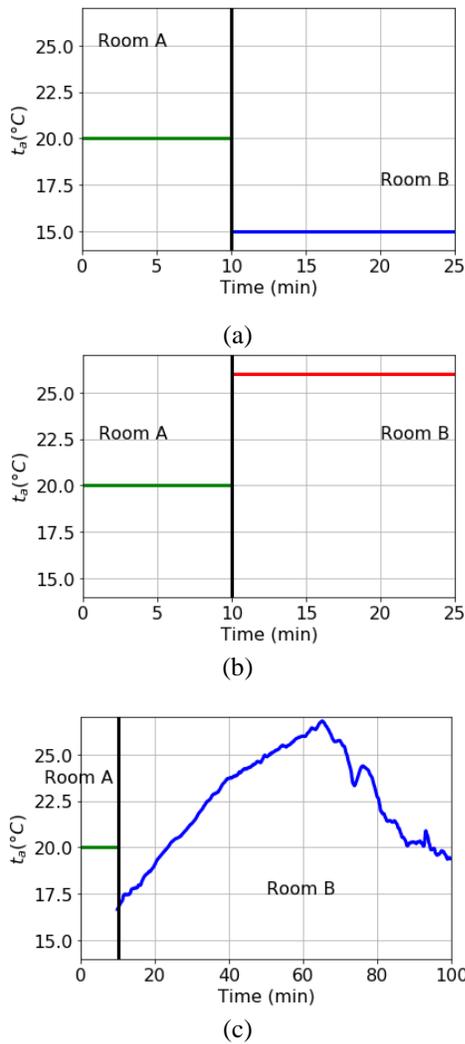


Fig. 5. (a) Trend of t_a during Experiment 1. (b) Trend of t_a during Experiment 2. (c) Trend of t_a during Experiment 3.

D. 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 3):

- 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,

TABLE IV

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

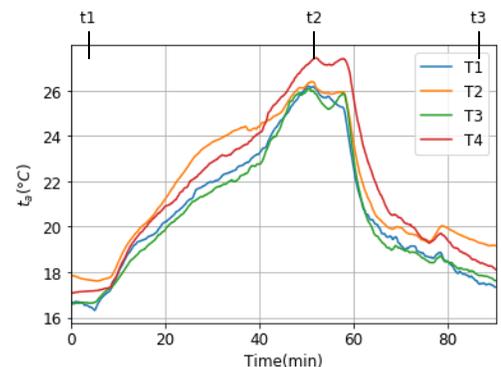


Fig. 6. Air temperature profiles during Experiment 3 and instant of time considered for the evaluation of the homogeneity of the thermal environment.

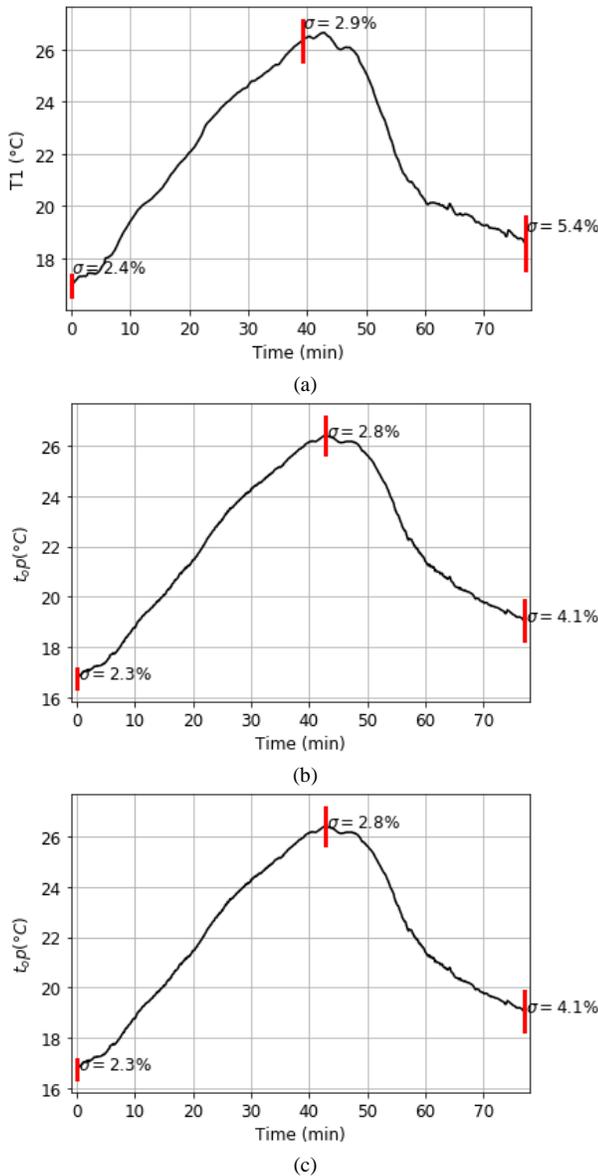


Fig. 7. Average profile for each environmental parameter, with the maximum deviation computed among all participants. (a) Air temperature in point T1. (b) Operative temperature. (c) Mean radiant temperature.

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.

To assess the thermal environment, the authors followed the specifications and methods of ISO 7726 Standard. 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

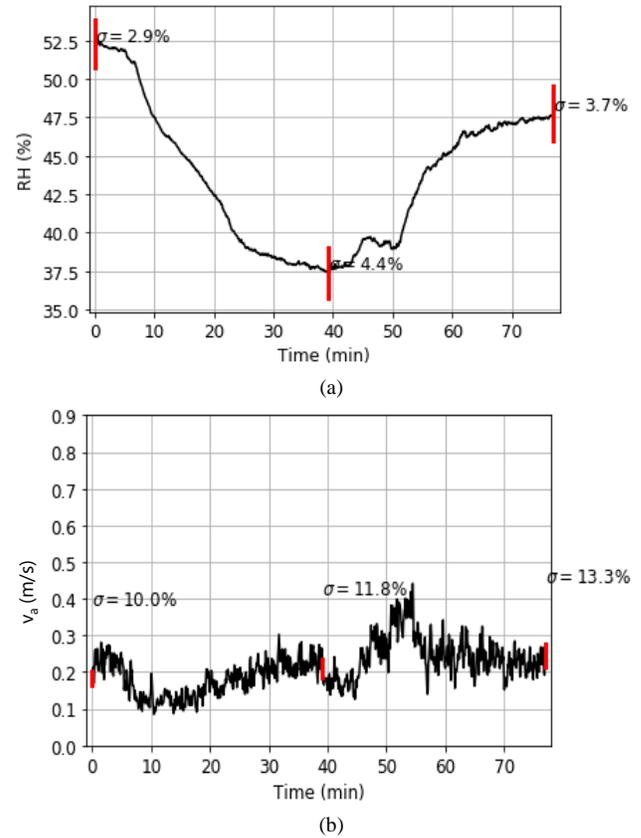


Fig. 8. Average profile for each environmental parameter, with the maximum deviation computed among all participants. (a) Relative humidity. (b) Air velocity.

Table 4. Fig 6 reports the trends of the air temperatures measured using the configuration of Fig 3.

Each participant in Experiment 3 was exposed to the same environmental conditions, as shown in Fig 7 and Fig 8, which display the average trend of each parameter during Experiment 3 and the maximum standard deviations in three different instants of time. Fig 7 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.

III. DATA PROCESSING

This paragraph contains a detailed explanation of how each signal 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.

A. Physiological signal analysis

The HRV trace was extracted from the smartwatch and the very first step of the procedure was the removal of outliers from the signal. In this work, HRV was derived from the photoplethysmographic signal (PPG) sensor installed on the smartwatch. This type of technology 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 [45] the authors measured the accuracy in the monitoring of HRV through a Samsung Galaxy smartwatch compared with a multi-parametric belt, Bioharness 3.0, and the results showed an accuracy in measuring HRV of 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 [21], [35].

B. 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 [36]. 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 [57]. According to literature [58], 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. 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 5.

C. 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.

D. 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

TABLE V
DESCRIPTION OF THE HRV FEATURES EXTRACTED.

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

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:

$$A = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

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

E. 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 ($f(HRV_t)$, $f(HRV_f)$ and $f(HRV_{nl})$); 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 [49].

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 layers [59]. 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 [60]. 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 [61]. 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 [62]. 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 [63], [64]. 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 [37]. 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.

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

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

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.

IV. RESULTS

A. 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

TABLE VI
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	90.8
RF	90.6	79	96.8	100	88.6	98.2	100	91.6	98.2	91.7
ETC	90.5	72	94.5	100	86	96.5	100	92.5	98.2	92.2
Mean										92.2

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 authors try 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 6, 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. 6 out of the 10 participants were included, since they provided a different TSV between Room A and Room B. As shown in Table 7, 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.

B. 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

TABLE VII
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

TABLE VIII
DESCRIPTION OF THE DATASETS BUILT FOR THE ANALYSIS.

Features	Dataset					
	F_H			F_{H+E}		
	$f(HRV_i)$	$f(HRV_j)$	$f(HRV_{nl})$			
RMSSD	LF	SD1	RMSSD	LF	SD1	t_a
MEAN	HF	SD2	MEAN	HF	SD2	RH
MEDIAN	LF/HF	SD1*SD2	MEDIAN	LF/HF	SD1*SD2	v_a
	HF/LF			HF/LF		

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.

C. Features and dataset description

The algorithms were first trained and tested by using as input only the physiological features previously described in Table 5. 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 F_H .

F_H was subsequently combined with the three

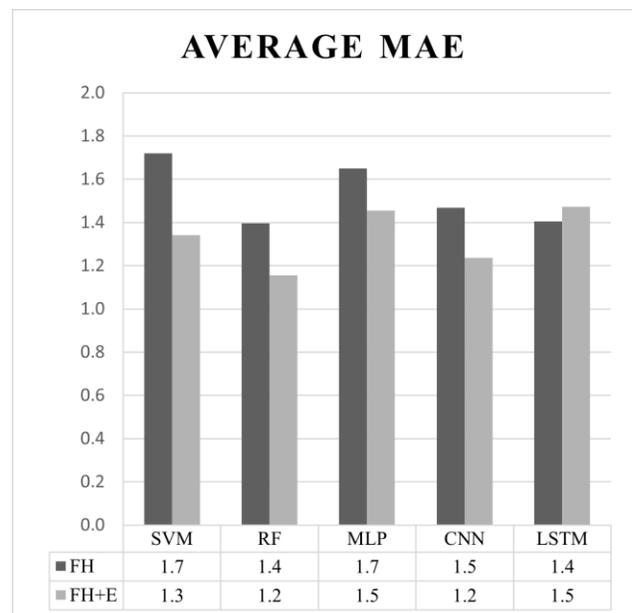


Fig. 9. Average MAE computed by averaging the MAE obtained from the prediction of each algorithm on the single user, left out of the training.

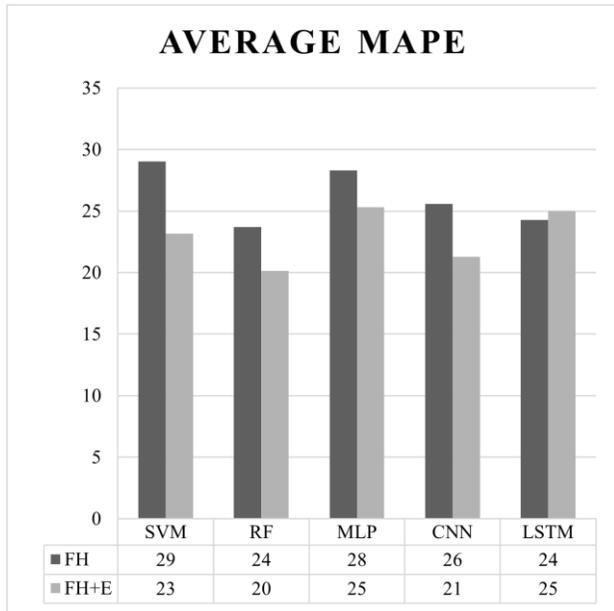


Fig. 10. Average MAPE computed by averaging the MAPE obtained from the prediction of each algorithm on the single user, left out of the tra

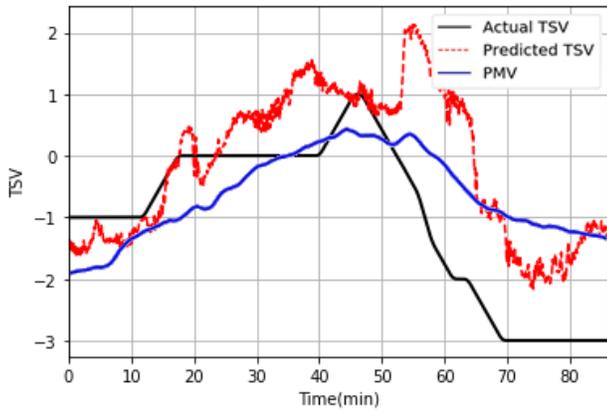


Fig. 11. Real TSV and PMV against the predicted TSV obtained from the testing on one user, adopting the LOSO approach.

environmental parameters (t_a , v_a , RH), thus creating F_{H+E} , which is showed in Table 8.

The performance obtained by testing each algorithm on the participant that was not previously included in the training is shown in Fig 9 and Fig 10. 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 F_{H+E} , both MAE and MAPE decreased compared to the training made only with physiological HRV features (F_H), except for the average MAE computed from the LSTM algorithm. 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 F_H dataset was used. On the other hand, CNN and RF provided comparable performances when using the F_{H+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.

Fig 11 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.

D. Uncertainty analysis

An uncertainty analysis on the results of the LOSO approach obtained with Experiment 3 was performed to evaluate the impact of the uncertainty in the measurement of the HRV collected with the smartwatch on the prediction of the TSV. The model was trained to predict TSV, and it was therefore interesting to evaluate the impact of any source of uncertainty added to the raw HRV data on the model trained. The accuracy of the RF model in predicting TSV depends on the accuracy of HRV, which is the only physiological variable collected. The uncertainty analysis was therefore performed to quantitatively evaluate that output uncertainty (i.e. TSV) was distributed between different sources of variation in the model inputs (i.e. HRV).

The methodology applied is the Monte Carlo method: the basic principle of this methodology is to perturb the calculation model by adding a distribution to the input variable, which, in this precise context, is the HRV collected from the smartwatch, and evaluate the model sensitivity due to the perturbation on the resulting output variation [41]. The output variation obtained in response to the perturbed input variable was assessed in relation to a reference condition. Therefore, the final aim was to investigate, through an uncertainty analysis, the impact of a perturbed HRV signal on the prediction of the TSV using a trained ML model. The perturbation assigned to the input HRV was ± 4 ms, as reported in Table 9 [45]. More in detail, the reference value was obtained by an initial processing of the input HRV belonging to all the subjects that performed Experiment 3. The model trained provided the reference values of the output. The test dataset and the corresponding TSV predicted using the trained ML model are used as reference values. The procedure is repeated by generating 10^6 perturbed HRV signals and applying the trained model to predict the perturbed TSV. Perturbation is assigned by summing each reference HRV value with a random value extracted from a uniform distribution. The simulation used the input uncertainty

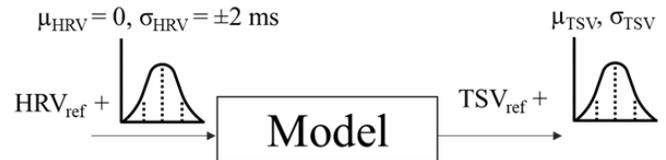


Fig. 12. Schematic representation of the methodology adopted for computing the uncertainty of the TSV due to the uncertainty of the input HRV.

TABLE IX

INPUT AND OUTPUT OF THE UNCERTAINTY ANALYSIS.

Input variable	Distribution	Uncertainty Input	Uncertainty output (TSV unit)
HRV	Uniform ($\mu=0$, $\sigma=2$ ms)	± 4 ms	± 0.14

specified in Table 9 and generated a random sample of possible combinations of the input variable. Fig 12 describes the approach adopted for computing the uncertainty associated to the prediction of TSV in response to a perturbed HRV. The result of the Monte Carlo analysis shows that the uncertainty of the smartwatch in the measurement of HRV, which is ± 4 ms, results in an output uncertainty which turned out to be ± 0.14 , expressed in TSV units.

V. DISCUSSION AND LIMITATION

The experiments conducted in [49] are considered the starting point used to develop the procedure conducted in this work. In particular, [49] demonstrated that when the user does not perform any activity and environmental parameters (t_a , RH and v) are more controlled, HRV features can provide accurate performance in the estimation of TSV. However, in this work, 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 [49]. 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, [49] 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, the authors of this work 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.

The evaluation of the performance of the algorithms

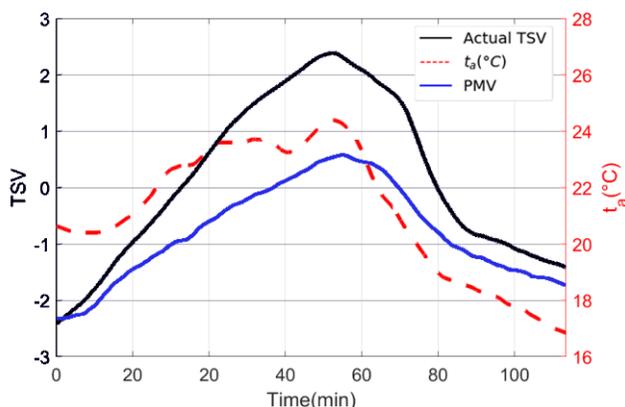


Fig. 13. Average trend of TSV and PMV against air temperature during Experiment 3, for all the users.

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 [65].

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 F_H 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 [49], 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. Another point of discussion that arises from the results obtained from this study is that Experiment 3 gives the possibility to examine any hysteretic behavior of TSV and HRV: in particular, Experiment 3 was built so that when the test started the temperature was set at 15°C , then risen up to 26°C and then set back to 15°C , therefore, it was also possible to observe if the TSV returned back to the initial condition. Fig 13 represents the mean behavior of the TSV and the temperature among all participants in Experiment 3. It is possible to see that the average TSV among all participants does go back to the initial condition, since the TSV in the first part of the test starts at -2 , goes up to almost 3 and then, in the cooling phase, returns back to the initial condition. On the other hand, the HRV does not directly show the same behavior as the TSV. To examine this aspect in depth, further ad-hoc experimental campaigns should be carried out to study the presence of hysteresis in physiological data, such as HRV, under diverse environmental conditions. Figure 13 also shows the trend of the average PMV which follows that of the TSV, although there are some numerical discrepancies.

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, the authors have 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.

VI. CONCLUSION

The aim of this paper is to address the possibility of measuring the TSV of participants exposed to different environmental conditions starting from physiological parameters related to HRV. A baseline analysis conducted through Experiment 1 (Cold-induced discomfort) and Experiment 2 (Warm-induced discomfort) showed that when participants are exposed to a comfort condition and then suddenly to a discomfort condition, the physiological response, represented by LF/HF, can be different from participant to participant. ML classification algorithms can distinguish between comfort and discomfort with an average accuracy of 92.2%, which suggests that frequency-domain quantities of HRV can be used as indicators to distinguish whether a user is thermally comfortable or in discomfort.

When it comes to a transient environment, as in Experiment 3, the problem becomes a regression issue in the prediction of TSV. Results show how physiological quantities in transient conditions can be used to estimate TSV with mean MAE and MAPE values that reach up to 1.4 and 24%, respectively. However, the aggregation of environmental and physiological features improves the performances of the ML regression algorithm, thus improving the mean MAE and MAPE by 1.2 and 24%. These results demonstrate that the addition of physiological features can help to improve the measurement of users' TSV. In addition, an uncertainty analysis regarding the impact of the uncertainty of the smartwatch in the prediction of TSV has pointed out that the uncertainty associated to the output is ± 0.14 , expressed as TSV units.

The work presented has opened the possibility to include physiological parameters in the prediction of thermal comfort, to develop more tailored solutions for users in non-controlled environments, which are often similar to real-life conditions. However, this methodology could also be applied under controlled thermal conditions (e.g., climatic chamber) on a higher number of participants, to create generalized models. Wearable devices can become a powerful tool to acquire more complex and subjective features related to thermal comfort that can be included in thermal comfort modeling, and consequently help in the management of the HVAC set-point that regulates air temperature in buildings.

ACKNOWLEDGMENT

The authors wish to thank the company Tecnia (<https://www.tecnialia.com/es/>) for providing the KUBIK facility and the sensors for the acquisition of the environmental parameters. The activity presented in this work was carried out within the RenoZEB project, funded by the European Union's Horizon 2020 research and innovation programme under Grant Agreement No.768718. The authors also wish to thank the RenoZEB partners for the fruitful discussions and collaboration.

REFERENCES

- [1] E. Lee, 'Indoor environmental quality (IEQ) of LEED-certified home: Importance-performance analysis (IPA)', *Building and Environment*, vol. 149, pp. 571–581, Feb. 2019, doi: 10.1016/j.buildenv.2018.12.038.
- [2] Y. Geng, W. Ji, B. Lin, and Y. Zhu, 'The impact of thermal environment on occupant IEQ perception and productivity', *Building and Environment*, vol. 121, pp. 158–167, Aug. 2017, doi: 10.1016/j.buildenv.2017.05.022.
- [3] S. Tanabe, Y. Iwahashi, S. Tsushima, and N. Nishihara, 'Thermal comfort and productivity in offices under mandatory electricity savings after the Great East Japan earthquake', *Architectural Science Review*, vol. 56, no. 1, pp. 4–13, Feb. 2013, doi: 10.1080/00038628.2012.744296.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] 'CEN EN Standard 16798-2. Energy Performance of Buildings—Ventilation for Buildings - Part 2: Interpretation of the Requirements in EN 16798-1—Indoor Environmental Input Parameters for Design and Assessment of Energy Performance of Buildings Addressing Indoor Air Quality, Thermal Environment, Lighting and Acoustics—Module M1-6; European Committee for Standardization: Brussels, Belgium, 2019.'
- [8] 'European Committee for Standardization. EN 16798-1:2019 Energy performance of buildings - Ventilation for buildings - Part 1: Indoor environmental input parameters for design and assessment of energy performance of buildings addressing indoor air quality, thermal environment, lighting and acous. 2019.'
- [9] 'ISO. ISO Standard 7726. Ergonomics of the Thermal Environment—Instruments for Measuring Physical Quantities; International Organization for Standardization: Geneva, Switzerland, 1998.'
- [10] 'ISO. ISO Standard 8996. Ergonomics of the Thermal Environment—Determination of Metabolic Rate; International Organization for Standardization: Geneva, Switzerland, 2004.'
- [11] 'ISO 7730: Ergonomics of the Thermal Environment – Analytical Determination and Interpretation of Thermal Comfort Using Calculation of the PMV and PPD Indices and Local Thermal Comfort Criteria Genève (2005).'
- [12] '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.'
- [13] Malchaire, Jacques, Francesca Romana d'AMBROSIO ALFANO, and Boris Igor Palella. "Evaluation of the metabolic rate based on the recording of the heart rate." *Industrial health* 55.3 (2017): 219-232.
- [14] F. R. d'Ambrosio Alfano, B. W. Olesen, B. I. Palella, D. Pepe, and G. Riccio, 'Fifty Years of PMV Model: Reliability, Implementation and Design of Software for Its Calculation', *Atmosphere*, vol. 11, no. 1, Art. no. 1, Jan. 2020, doi: 10.3390/atmos11010049.
- [15] S. Torresin, G. Pernigotto, F. Cappelletti, and A. Gasparella, 'Combined effects of environmental factors on human perception and objective performance: A review of experimental laboratory works', *Indoor Air*, vol. 28, no. 4, pp. 525–538, 2018, doi: <https://doi.org/10.1111/ina.12457>.
- [16] Schweiker, Marcel, et al. "Drivers of diversity in human thermal perception—A review for holistic comfort models." *Temperature* 5.4 (2018): 308-342.
- [17] Z. Wang et al., 'Individual difference in thermal comfort: A literature review', *Building and Environment*, vol. 138, pp. 181–193, Jun. 2018, doi: 10.1016/j.buildenv.2018.04.040.
- [18] Angilletta Jr, Michael J., et al. "The neuroscience of adaptive thermoregulation." *Neuroscience letters* 692 (2019): 127-136.
- [19] Wilson, Thad E., and Kristen Metzler-Wilson. "Autonomic Thermoregulation." *Oxford Research Encyclopedia of Neuroscience*. 2018.
- [20] Zangróniz, Roberto, et al. "Electrodermal activity sensor for classification of calm/distress condition." *Sensors* 17.10 (2017): 2324.
- [21] 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.

[22] Matalucci, Berardo, et al. "An experimental design framework for the personalization of indoor microclimates through feedback loops between responsive thermal systems and occupant biometrics." *International Journal of Architectural Computing* 15.1 (2017): 54–69.

[23] Pietroni, Filippo, et al. "Identification of Users' Well-Being Related to External Stimuli: A Preliminary Investigation." *Convegno Nazionale Sensori*. Springer, Cham, 2018.

[24] S. Casaccia, E. J. Sirevaag, E. J. Richter, J. A. O'Sullivan, L. Scalise, and J. W. Rohrbach, '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.

[25] 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.

[26] 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.

[27] 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.

[28] 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.

[29] 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.

[30] 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.

[31] 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.

[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] A. Alqaraawi, A. Alwosheel, and A. Alasaad, 'Heart rate variability estimation in photoplethysmography signals using Bayesian learning approach', *Health Technol Lett*, vol. 3, no. 2, pp. 136–142, Jun. 2016, doi: 10.1049/hlt.2016.0006.

[34] 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.

[35] 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.

[36] Catai, Aparecida Maria, et al. "Heart rate variability: are you using it properly? Standardisation checklist of procedures." *Brazilian journal of physical therapy* 24.2 (2020): 91–102.

[37] 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.

[38] 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.

[39] 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.

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

[41] Q. J. Kwong, N. M. Adam, and B. B. Sahari, 'Thermal comfort assessment and potential for energy efficiency enhancement in modern tropical buildings: A review', *Energy and Buildings*, vol. 68, pp. 547–557, Jan. 2014, doi: 10.1016/j.enbuild.2013.09.034.

[42] 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.

[43] 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.

[44] 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.

[45] 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).

[46] K. N. Nkurikiyeyezu, Y. Suzuki, and G. F. Lopez, 'Heart rate variability as a predictive biomarker of thermal comfort', *J Ambient Intell Human Comput*, vol. 9, no. 5, pp. 1465–1477, Oct. 2018, doi: 10.1007/s12652-017-0567-4.

[47] 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.

[48] 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.

[49] 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.

[50] 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.

[51] 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.

[52] 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/aa5a9.

[53] 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.

[54] 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.

[55] 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.

[56] 'UNI EN ISO 7726:2002'. http://store.uni.com/catalogo/uni-en-iso-7726-2002?josso_back_to=http://store.uni.com/josso-security-check.php&josso_cmd=login_optional&josso_partnerapp_host=store.uni.com (accessed Jan. 25, 2021).

[57] Nkurikiyeyezu, Kizito, et al. "Thermal Comfort and Stress Recognition in Office Environment." *HEALTHINF*. 2019.

[58] 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.

[59] 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.

[60] Z. Car, S. Baressi Šegota, N. Anđelić, I. Lorencin, and V. Mrzljak, 'Modeling the Spread of COVID-19 Infection Using a Multilayer Perceptron', *Computational and Mathematical Methods in Medicine*, May 29, 2020. <https://www.hindawi.com/journals/cmmm/2020/5714714/> (accessed Feb. 03, 2021).

[61] D. P. Kingma and J. Ba, 'Adam: A Method for Stochastic Optimization', arXiv:1412.6980 [cs], Jan. 2017, Accessed: Feb. 03, 2021. [Online]. Available: <http://arxiv.org/abs/1412.6980>.

[62] Z. Chen et al., 'A Novel Ensemble Deep Learning Approach for Sleep-Wake Detection Using Heart Rate Variability and Acceleration', *IEEE Transactions on Emerging Topics in Computational Intelligence*, pp. 1–10, 2020, doi: 10.1109/TETCI.2020.2996943.

[63] S. Oh, J.-Y. Lee, and D. K. Kim, 'The Design of CNN Architectures for Optimal Six Basic Emotion Classification Using Multiple Physiological Signals', *Sensors*, vol. 20, no. 3, Art. no. 3, Jan. 2020, doi: 10.3390/s20030866.

[64] R. Banerjee, A. Ghose, and K. M. Mandana, 'A Hybrid CNN-LSTM Architecture for Detection of Coronary Artery Disease from ECG', in *2020 International Joint Conference on Neural Networks (IJCNN)*, Jul. 2020, pp. 1–8, doi: 10.1109/IJCNN48605.2020.9207044.

[65] 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.



Nicole Morresi was born in Macerata, Italy, in 1993. She received the master's (cum laude) degree in biomedical engineering from the Università Politecnica delle Marche, Italy, in 2018. She is currently pursuing the Ph.D. degree in mechanical and thermal measurement. Her research is focused on comfort measurement and user well-being.



J. Ignacio Torrens-Galdiz is an R&D Engineer with more than 10 years of international experience in building performance and energy efficiency research projects. He joined the Building Technologies Division of TECNALIA in 2018 after working various European academic research groups and research institution. Author of more than 15 journal and conference papers, he has participated in different roles in many national and international research projects



Sara Casaccia was born in Jesi, Italy, in 1987. She received her Master degree (cum laude) in Biomedical Engineering in 2011 at Università Politecnica delle Marche (Italy) and her Ph.D in Mechanical Engineering in 2015 in the same institution. She attended the department of Electrical and System Engineering of the Washington University in Saint Louis (Missouri-USA) for a period of 7 months during her Ph.D. She is a Postdoc Research Fellow at the

Department of Industrial Engineering and Mathematical Sciences (DIISM) of Università Politecnica delle Marche (UNIVPM). Her research focuses on sensors and measurement techniques for supporting people in the life environments (e.g. comfort, wellbeing), data processing to extract complex information (e.g. using AI), sensors for biomedical applications.



Gian Marco Revel was born in Ancona, Italy, in 1970. He received the master's degree in mechanical engineering from the Università di Ancona, Italy, in 1995, and the Ph.D. degree in measurement science in 1998. He is Full Professor of Mechanical and Thermal Measurement with Università Politecnica delle Marche. He is the author of more than 180 publications and responsible for several national and international research projects.



Matteo Sorcinelli was born in Fano, Italy, in 1991. He received the master's degree in biomedical engineering from Università Politecnica delle Marche, Italy, in 2018. He has worked as research fellow from 2019 to 2020, focusing his activity on sensors and measurements for improving the life of older people.



Marco Arnesano was born in Campi Salentina, Italy, in 1982. He received the master's (cum laude) degree in Mechanical Engineering and the Ph.D. degree in Measurement Science from Università Politecnica delle Marche, in 2009 and 2013, respectively. He is Associate Professor of Mechanical and Thermal Measurements and Coordinator of Industrial Engineering at Università e-Campus (Italy).



Amaia Uriarte is Senior Researcher in the field of low energy building. She joined TECNALIA in 2001 as a researcher in emissions to the atmosphere related to energy and industrial processes, from which she evolved into the field of design and energy renovation of buildings in 2008. She is focused on the integration of innovative technologies for demand reduction, renewable generation and energy management in buildings.

She contributes in 15 scientific and conference papers. She belongs an Expert Reviewer of National Greenhouse Gas Inventories by the UNFCCC for the Industrial and Energy Sector.