

# UC Berkeley

## Controls and Information Technology

### Title

Machine learning approaches to predict thermal demands using skin temperatures:  
Steady-state conditions

### Permalink

<https://escholarship.org/uc/item/3qt1n6qv>

### Authors

Dai, Changzhi  
Zhang, Hui  
Arens, Edward  
[et al.](#)

### Publication Date

2017-03-01

### DOI

10.1016/j.buildenv.2016.12.005

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial-ShareAlike License, available at  
<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Peer reviewed



# Machine learning approaches to predict thermal demands using skin temperatures: Steady-state conditions



Changzhi Dai <sup>a,b,\*</sup>, Hui Zhang <sup>b</sup>, Edward Arens <sup>b</sup>, Zhiwei Lian <sup>a</sup>

<sup>a</sup> Department of Architecture, School of Naval Architecture, Ocean & Civil Engineering, Shanghai Jiao Tong University, China

<sup>b</sup> Center for the Built Environment, UC Berkeley, USA

## ARTICLE INFO

### Article history:

Received 7 September 2016

Received in revised form

20 November 2016

Accepted 5 December 2016

Available online 8 December 2016

### Keywords:

Thermal environment

Skin temperature

Support vector machine

Intelligent control

## ABSTRACT

Inefficient controlling strategies in heating and cooling systems have given rise to a large amount of energy waste and to widespread complaints about the thermal environment in buildings. An intelligent control method based on a support vector machine (SVM) classifier is proposed in this paper. Skin temperatures are the only inputs to the model and have shown attractive prediction power in recognizing steady state thermal demands. Data were accumulated from two studies to consider potential use for either individuals or a group of occupants. Using a single skin temperature correctly predicts 80% of thermal demands. Using combined skin temperatures from different body segments can improve the model to over 90% accuracy. Results show that three skin locations contained enough information for classification and more would cause the curse of dimensionality. Models using different skin temperatures were compared. Optimal parameters for each model were provided using grid search technique. Considering the overfitting possibility and the cases without learning processes, SVM classifiers with a linear kernel are preferred over Gaussian kernel ones.

© 2016 Elsevier Ltd. All rights reserved.

## 1. Introduction

Widespread in both residential and commercial buildings, heating, ventilation and air conditioning (HVAC) systems consume almost 20% of the world's energy. The benefits from this consumption are not as great as they should be. In a large study, only 11% of buildings met the basic criterion [1,2] that 80% or more of their occupants be satisfied with their thermal environment [3]. Behind the unsatisfactory cases may be ineffective thermal environment control strategies.

Most of the air conditioning systems have temperature/humidity controllers. In some cases such as residential buildings and private offices, occupants tune the set points according to their perception, without a sense of what temperatures could be comfortable or the energy costs associated with the temperatures selected. As a result, the set points are frequently revised and energy is wasted [4]. In other circumstances like conference halls, occupants have little access to the controllers. Temperature is preset based on standard recommended temperatures or on the operators' feelings about what causes the least thermal complaints,

which often results in overcooling of the space.

Over the past few decades, many thermal comfort controllers have been proposed [5–7]. The essential idea is to replace the occupants' feedback with thermal sensation prediction based on built-in comfort models or data-driven self-learning methods. The inputs are commonly physical environment parameters such as air temperature, humidity, air velocity and radiation temperature. Measuring these in occupied spaces presents a number of challenges. In addition, clothing insulation and occupant activity level are difficult-to-measure factors that greatly affect comfort model accuracy. To alleviate these problems, one potential approach is to control the thermal environment based on physiological parameters.

The dramatic progress of wearable devices has created technology ready for monitoring body parameters in daily life. Surface body temperature sensors could be attached to watches, clothes and so on. Fiber Bragg grating (FBG) based sensors have made it possible to monitor the skin temperature of different body parts under intelligent clothing [8]. The development of infrared camera technology also makes it possible to capture uncovered skin

\* Corresponding author. Shanghai Jiao Tong University, 800 Dongchuan Road, Minhang District, Shanghai, 200240, China.  
E-mail address: [dczdxt@163.com](mailto:dczdxt@163.com) (C. Dai).

temperatures remotely [9]. However, previous models linking sensation with physiological parameters such as by Fiala [10] and Zhang [11–14] are too complicated to be used in designing practical controllers.

To develop practical ways of using skin temperatures to control heating and cooling systems, Taniguchi et al. [15] proposed an equation to estimate car occupants' thermal sensation based on face skin temperatures. Wang et al. [16] conducted a lab study to explore the hypothesis proposed by Humphreys et al. [17] that finger temperature or an air and finger temperature combination was capable of predicting thermal sensations. They found that the temperature gradient between arm and hand could be a good indicator for cool sensation. To a certain extent, these studies have laid the groundwork for designing intelligent controller based on skin temperatures.

In this paper, we have combined a machine learning algorithm with local steady-state skin temperatures. Different SVM classifiers, model parameters and skin temperature combinations were tried to explore to what extent skin temperatures could go in predicting the thermal states. Exact ways of applying the prediction models to control heating and cooling system were proposed. We collected data from two studies to test the model performance for building areas with either one or more occupants. As the SVM approach needs a process of data learning, we discuss the learning sample sizes needed for certain classifiers to work well automatically, and also examine the performance of preset models for new occupants without training.

## 2. SVM controlling prototype

As shown in Fig. 1, the SVM approach is a data-driven model. During the learning procedure, data are collected in traditional ways in which occupants adjust the heating and cooling system according to their perception of the thermal environment. Thermal demands and corresponding local skin temperatures are input to the SVM classifiers to study decision boundaries. After that, SVM models predict the thermal demands of occupants based on real-time skin temperature measurements. The application scope of the controller is flexible, including but not limited to HVAC systems and more localized task-ambient conditioning (TAC) systems.

Imagine that one or more local skin temperatures  $T$  constitute an  $m$ -dimensional space, in which the sample data vectors  $(T, y)_i$  ( $i = 1, \dots, n$ ) are points with corresponding thermal demands  $y$  ( $y = -1, 0, 1$ ). The basic idea behind the SVM classifier here is to construct the optimal hyperplanes in the space that could differentiate the data vectors from others with different thermal demands. As the SVM classifier is basically a two-class method, a one-vs-one (OvO) strategy is adopted in this paper to reduce multiclass classification into a multiple binary problem [18]. Choosing a proper kernel function has always been the largest challenge in using SVM classifiers. In this paper, we compare the performance of linear and Gaussian kernels. Standardization of the dataset was implemented first. As there are two parameters  $C$  and  $\gamma$  for the Gaussian kernel

SVM classifier and one parameter  $C$  for the linear kernel SVM classifier, we used a “grid search” method on  $C$  and  $\gamma$  to find the proper parameters for the classifiers. Exponentially growing sequences of  $C$  ( $2^{-2}, 2^{-1}, 2^0, \dots, 2^{10}$ ) and  $\gamma$  ( $2^{-10}, 2^{-9}, 2^{-8}, \dots, 2^4$ ) were tried and the ones achieving the best cross-validation accuracy were picked. To grasp a good understanding of the SVM classifier, one can refer to the introductions of Hsu et al. [19] and Noble [20].

## 3. Experiment 1

This experiment considers the validation of SVM classifiers to learn and predict steady-state thermal demands for a group of occupants in a uniform environment. We place the focus on three basic issues: a. the selection of input features; b. the comparison of classifiers with different kernels; c. tuning parameters for the SVM models.

### 3.1. Methods

The data were accumulated from a series of tests carried out in the Controlled Environmental Chamber at UC Berkeley to correlate skin temperatures with whole-body sensations for a variety of warm to cool conditions. Sensation votes were obtained at the end of the periods used to acclimatize the subjects to the environmental conditions in the tests. 70 tests were conducted using 11 subjects, with 969 votes collected.

Subjects were first preconditioned to the day's test in a Jacuzzi bath for 15 min. After that, thermocouples were attached to collect local skin temperatures every 5 s. There were totally 28 body locations measured during the tests. Only 13 of them were used in this paper: the forehead, cheek, chest, back, abdomen and 8 extremity skin temperatures on left the side of the body: upper arm, forearm, hand, finger, thigh, shin, calf and foot (Fig. 2). Subjects wore a long-sleeve elastic leotard (0.32clo) and socks (0.02clo) with the thermocouples covered except those at head and hand locations. Whole-body thermal sensation was investigated repetitively by pop-up questionnaires on the computer at varying time steps of 1–3 min. Experimental details are graphically described in previous publications [21,22].

In this study, the first 10 votes in each test were abandoned in order to make sure that the data represent steady-state conditions. The total data set was here split into two subsets: a training set of 80% (774) and a holdout set of 20% (195). The distributions of total data and testing data are shown in Fig. 3.

The sensation scale is similar to the ASHRAE 7-point scale, adding “very hot” and “very cold” (9-point scale: 4- “very hot”, 3- “hot”, 2- “warm”, 1- “slightly warm”, 0- “neutral”, -1- “slightly cool”, -2- “cool”, -3- “cold”, -4- “very cold”). Statistical analysis was implemented in SPSS (IBM SPSS Statistics for Macintosh, Version 22.0). Data were classified into 5 groups based on the thermal sensation votes (TSV): heating demand ( $TSV < -1.5$ , cold); slight heating demand ( $-1.5 \leq TSV < -0.5$ , cool); neutral ( $-0.5 \leq TSV \leq 0.5$ ); slight cooling demand ( $0.5 < TSV \leq 1.5$ , warm);

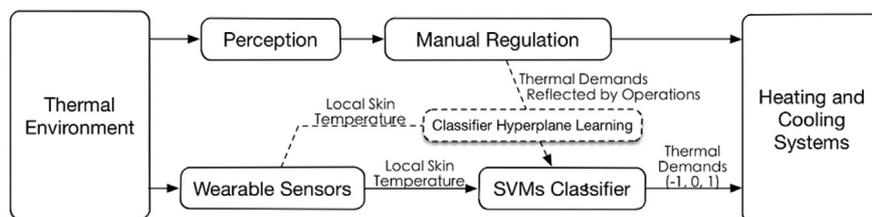


Fig. 1. The controlling concept of SVM classifier based on skin temperature.

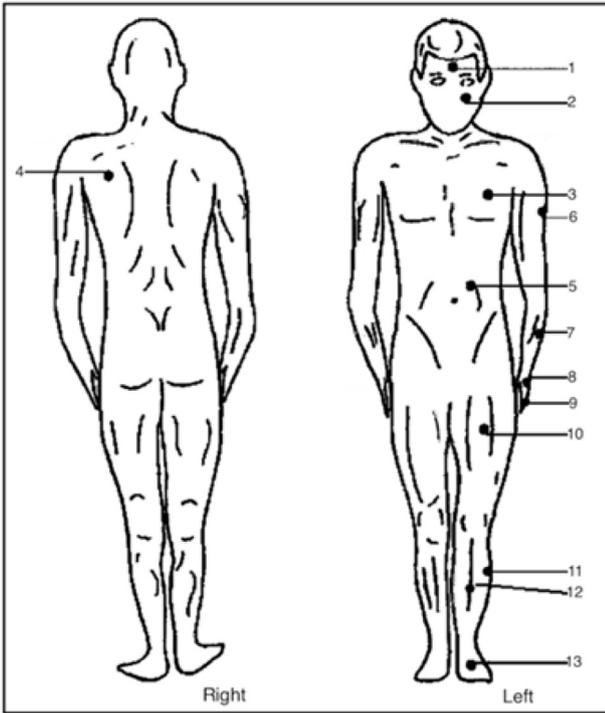


Fig. 2. Thirteen locations of skin temperature measured in the tests (referred from Ref. [21]).

cooling demand (TSV > 1.5, hot). Levene's test was performed first to assess the equality of variances. As the homogeneity of variances was violated, we used Welch's ANOVA to examine the differences of skin temperatures between groups and used Games Howell post hoc test to confirm where the differences occurred. Effect sizes of local skin temperature between any two adjoining thermal demands were computed with Cohen's d.

$$d = \frac{\bar{T}_1 - \bar{T}_2}{s}$$

$$s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

where T denotes the average local skin temperature, n denotes the sample size and s denotes variance in each group.

The SVM classifiers were trained and tested in MATLAB 2016a (The MathWorks, Inc., MA, 2016). Because we only cared about how to operate the heating and cooling systems, the data were separated into 3 groups (-1-“heating demand”, 0-“neutral”, 1-“cooling demand”) instead of 5 with the partition located at ±0.5 TSV. A tenfold cross-validation approach was used to tune parameters C and γ with the 774 training votes. After the optimal C and γ was found, classifier performance was evaluated based on the testing data set shown in Fig. 3(b).

### 3.2. Results and discussion

Generally, one should be acquainted with the data before selecting input features for machine learning models. The thirteen local skin temperature were measured in four parts of human body: two from head, three from trunk, four from upper limbs and the last four from lower limbs. As shown in Fig. 4 sequentially, skin temperatures fluctuate more in cool/cold conditions than those in warm/hot conditions, which can be reflected not only by the growth of standard deviations for each location respectively, but also by the variations of mean temperatures between different skin locations. These variations imply that, aside from the values of each local skin temperature itself, the combined ones could have more effective information to differentiate thermal demands of occupants.

All thirteen local skin temperatures are significantly influenced by the conditions (P < 0.001 for each one). Post-hoc test results are shown in Table 1. Larger temperature differences are observed between conditions in the limbs parts, which improves their suitability as input features, but the limbs also have larger standard deviations which attenuate this advantage. We consequently computed the effect sizes to evaluate the sensitivity of the thirteen potential indicators in reflecting thermal demands (Table 1). Different from the P value telling whether differences exist between groups, effect sizes quantify the differences [23]. Cohen's,

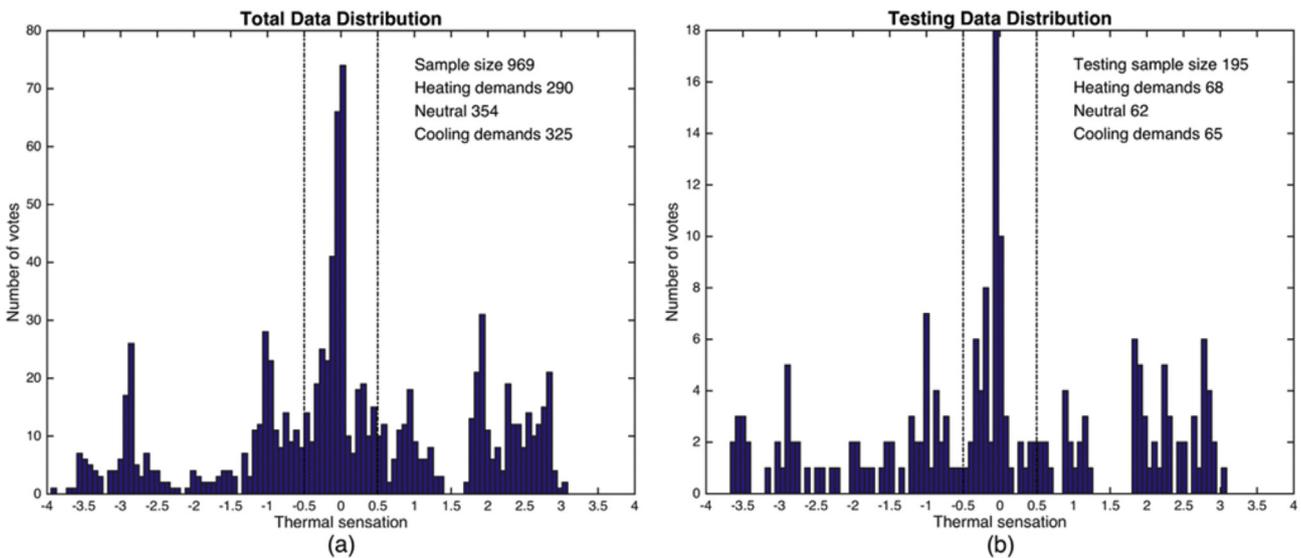


Fig. 3. Sample data distribution.

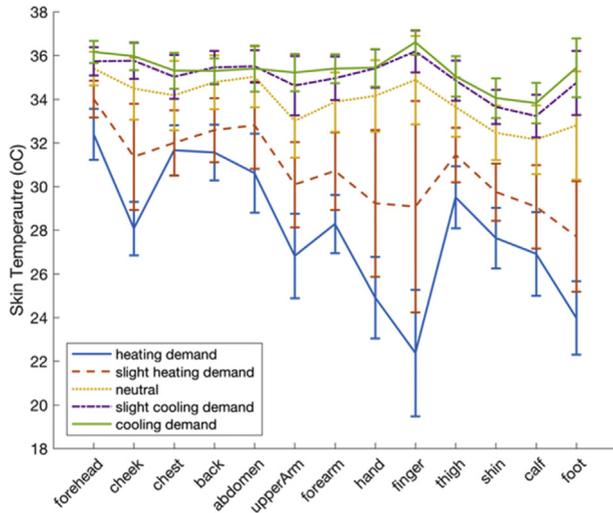


Fig. 4. Local skin temperatures under different thermal demand conditions.

Table 1  
Effect sizes between two adjoining conditions.

	-1 vs -2	0 vs -1	1 vs 0	2 vs 1
Forehead	<b>1.592*</b>	<b>1.761*</b>	<b>0.450*</b>	<b>0.771*</b>
Cheek	<b>1.684*</b>	<b>1.748*</b>	<b>0.969*</b>	0.299
Chest	0.250	<b>1.393*</b>	<b>0.578*</b>	0.318
Back	<b>0.743*</b>	<b>1.683*</b>	<b>0.587*</b>	-0.246
Abdomen	<b>1.151*</b>	<b>1.391*</b>	<b>0.373*</b>	-0.109
Upper Arm	<b>1.681*</b>	<b>1.650*</b>	<b>1.003*</b>	<b>0.580*</b>
forearm	<b>1.536*</b>	<b>2.117*</b>	<b>0.841*</b>	<b>0.565*</b>
Hand	<b>1.568*</b>	<b>2.145*</b>	<b>0.837*</b>	0.036
Finger	<b>1.666*</b>	<b>1.796*</b>	<b>0.716*</b>	<b>0.587*</b>
Thigh	<b>1.454*</b>	<b>1.651*</b>	<b>0.952*</b>	0.207
Shin	<b>1.562*</b>	<b>2.149*</b>	<b>1.025*</b>	<b>0.462*</b>
Calf	<b>1.124*</b>	<b>1.831*</b>	<b>0.733*</b>	<b>0.630*</b>
Foot	<b>1.725*</b>	<b>2.037*</b>	<b>0.851*</b>	<b>0.500*</b>

Note: [Bold]\* indicates that the post-hoc tests show significant difference at level  $p < 0.01$  when comparing the local skin temperatures between two conditions; -2 denotes 'heating demand'; -1 denotes 'slight heating demand'; 0 denotes 'no heating/cooling demand'; 1 denotes 'slight cooling demand'; 2 denotes 'cooling demand'.

d values of 0.2, 0.5 and 0.8 correspond to small, medium and large effects respectively [24]. As shown in Table 1, the largest large effect sizes exist between cool and neutral conditions. As a result, it seems

easier to differentiate heating-demand (-1, -2) conditions from neutral (0) than cooling-demand (1, 2) conditions from neutral. The seven locations showing large effect size (>0.8) in skin temperatures between neutral and warm conditions were the shin, upper arm, cheek, thigh, foot, forearm, hand in descending order.

Considering the practical measurement of skin temperature, it is desirable to minimize the number of input features. Table 2 shows how the SVM classifier works if just one location is used. The prediction accuracies for three thermal demands and an overall rate were computed. The shin skin temperature predicts the thermal demands best in both the Gaussian and linear kernel models, with the overall testing accuracy reaching 83.6% and 81.5% respectively. Optimal parameters were  $(2^2, 2^{-4})$  for the Gaussian SVM classifier and  $2^{-2}$  for the linear SVM classifier. Two classifiers both work well in predicting heating demands (around 90% accuracy), but do not perform as well in predicting cooling demands and neutral. Among the 13 locations, abdomen skin temperature had the worst prediction performance with overall rates of 56.4% and 54.4% respectively.

It has been claimed in a previous study that the upper limb skin temperature gradient and forehead-to-extremity gradient could indicate warm or hot sensations [16]. From the definition of hyperplane and the following transformation, we can see that the classifier using forehead and finger temperatures is theoretically equivalent to that trained with forehead temperature and forehead-finger temperature gradient. Consequently, Table 3 shows the performance of models with different input combinations considering these gradients. Generally, there is a slight improvement by adding an upper extremity skin temperature gradient or head-to-extremity gradient. The optimal SVM classifier is the Gaussian-kernel one with cheek and hand skin temperatures as its input features. The overall accuracy reaches 85.1% when the optimal parameters  $(2^2, 2^{-2})$  are adopted.

$$\omega \cdot \begin{bmatrix} T_1 \\ T_2 \end{bmatrix} + \omega_0 = \omega' \cdot \begin{bmatrix} T_1 \\ T_2 - T_1 \end{bmatrix} + \omega_0 = 0$$

Considering the possibility of remote skin temperature measurements using a thermal camera, it became interesting to realize how well uncovered skin temperatures could predict thermal demands (Table 4). The Gaussian kernel SVM performs best at the rate of 88.7% with 4 input features: skin temperatures of forehead, cheek, forearm and hand while the linear kernel SVM reaches its highest accuracy of 80% when the forehead is excluded. Note that more features would not guarantee better performance, which can

Table 2  
The performance of SVM classifiers using a single skin temperature to predict thermal demands.

Input features	C	γ	Testing accuracy	Accuracy for -1	Accuracy for 0	Accuracy for 1
Forehead	4   0.25	0.125   -	72.8%   58.5%	85.3%   82.4%	82.3%   77.4%	50.8%   15.4%
Cheek	8   0.25	0.0625   -	78.5%   74.9%	91.2%   73.5%	71.0%   85.5%	72.3%   66.2%
Chest	32   0.25	0.25   -	70.8%   76.9%	79.4%   88.2%	43.5%   54.8%	87.7%   86.2%
Back	32   0.5	0.25   -	76.9%   70.3%	91.2%   83.8%	77.4%   37.1%	61.5%   87.7%
Abdomen	16   2	0.5   -	56.4%   54.4%	66.2%   64.7%	69.4%   100.0%	33.8%   0.0%
Upper Arm	256   1	0.0625   -	79.5%   79.4%	72.1%   82.9%	80.6%   70.82%	86.2%   83.4%
forearm	256   0.25	0.0625   -	75.9%   80.5%	86.8%   89.7%	54.8%   69.4%	84.6%   81.5%
Hand	1   0.25	0.25   -	79.0%   77.4%	94.1%   80.9%	58.1%   67.7%	83.1%   83.1%
Finger	2   0.5	2   -	72.8%   70.8%	89.7%   89.7%	43.5%   32.3%	83.1%   87.7%
Thigh	0.5   0.25	0.25   -	75.4%   74.4%	89.7%   83.8%	54.8%   50.0%	80.0%   87.7%
<b>Shin</b>	4   0.25	0.0625   -	<b>83.6%   81.5%</b>	91.2%   89.7%	77.4%   80.6%	81.5%   73.8%
Calf	1   0.5	4   -	76.4%   76.9%	89.7%   62.9%	62.9%   76.9%	75.4%   76.9%
Foot	32   0.5	0.125   -	76.9%   69.3%	95.6%   80.7%	58.1%   63.0%	75.4%   66.2%

Note: "Gaussian SVM result | linear SVM result" in each cell. Column "C" and "γ" provide optimal penalty factors and kernel scale factors, respectively. "Testing accuracy" is the over-all prediction rate. Last three columns represent for the probability of detection for each thermal demand. -1 denotes heating-demand condition; 0 denotes neutral condition; 1 denotes cooling-demand condition. Models showing the highest prediction accuracy are highlighted in bold fonts.

**Table 3**  
The improvement by adding skin temperature gradients in SVM classifiers.

Input features	C	$\Upsilon$	Testing accuracy	Accuracy for -1	Accuracy for 0	Accuracy for 1
Forehead; Forearm	4   0.5	0.25   -	83.1%   <b>80.5%</b>	89.7%   89.7%	69.4%   69.4%	89.2%   81.5%
Forehead; Hand	0.25   8	0.25   -	83.1%   75.9%	92.6%   82.4%	74.2%   62.9%	81.5%   81.5%
Forehead; Finger	2   8	0.25   -	79.5%   72.8%	95.6%   89.7%	58.1%   41.9%	83.1%   84.6%
Cheek; Forearm	0.5   0.5	0.125   -	83.6%   80.0%	98.5%   89.7%	82.3%   74.2%	69.2%   75.4%
<b>Cheek; Hand</b>	4   2	0.25   -	<b>85.1%</b>   80.0%	95.6%   83.8%	77.4%   88.7%	81.5%   67.7%
Cheek; Finger	16   64	0.5   -	81.0%   75.9%	89.7%   88.2%	61.3%   46.8%	90.8%   90.8%
Forearm; Hand	1   2	0.125   -	80.0%   76.9%	89.7%   80.9%	61.3%   69.4%	87.7%   80.0%
Forearm; Finger	64   256	0.5   -	81.5%   74.9%	91.2%   88.2%	56.5%   51.6%	95.4%   83.1%
Hand; Finger	8   1	0.25   -	72.8%   74.4%	92.6%   80.9%	43.5%   53.2%	80.0%   87.7%

Note: Models showing the highest prediction accuracy are highlighted in bold fonts.

**Table 4**  
The performance of SVM classifiers considering uncovered skin temperatures.

Input features	#features	C	$\Upsilon$	Testing accuracy	Accuracy for -1	Accuracy for 0	Accuracy for 1
Forehead; Cheek	2	4   64	0.125   -	85.1%   77.9%	97.1%   83.8%	82.3%   83.9%	75.4%   66.2%
<b>Cheek; Forearm; Hand</b>	3	4   8	0.25   -	84.1%   <b>80.0%</b>	95.6%   91.2%	75.8%   72.6%	80.0%   75.4%
Forehead; Forearm; Hand	3	2   64	0.25   -	85.1%   76.4%	92.6%   85.3%	79.0%   62.9%	83.1%   80.0%
<b>Forehead; Cheek; Forearm; Hand</b>	4	1   64	0.25   -	<b>88.7%</b>   79.0%	95.6%   88.2%	80.6%   72.6%	89.2%   75.4%
Forehead; Cheek; Hand; Finger; Forearm	5	2   64	0.5   -	81.5%   73.8%	95.6%   83.8%	59.7%   53.2%	87.7%   83.1%

Note: Models showing the highest prediction accuracy are highlighted in bold fonts.

be explained by the curse of dimensionality: the available data may become sparse as the volume of the space increases.

Regardless of how to implement measurements in practice, we tried many other input features to look for the optimal skin temperature combinations in Table 5. Gaussian kernel SVM classifiers generally outperform linear kernel ones in terms of overall accuracy. Two-input Gaussian SVM could differentiate thermal demands with an accuracy of around 90%. The linear kernel SVM performs best with 87.2% accuracy when upper arm and shin skin temperatures are used. The performance of linear kernel SVM classifier does not improve obviously when one more feature is added in the two-input models. Three-input Gaussian SVM classifiers, however, reach the best rate of 94.4% when chest, upper arm and shin skin temperatures are considered. We tried two four-input classifiers and the results seem to show that more skin temperature inputs do not improve the model performance.

To graphically compare the linear kernel and Gaussian kernel SVM, we trained both models with shin and upper arm skin temperatures and presented their decision boundaries in Figs. 5 and 6.

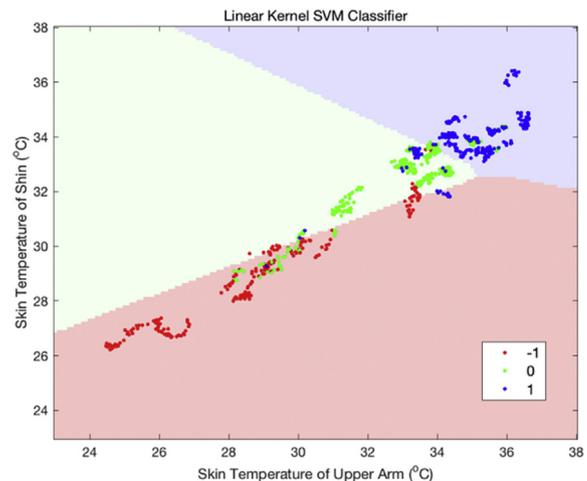


Fig. 5. SVM classifier with linear kernels.

**Table 5**  
Optimal input features considering skin temperatures from four anatomical segments.

Input features	#features	C	$\Upsilon$	Testing accuracy	Accuracy for -1	Accuracy for 0	Accuracy for 1
Cheek; Chest; Hand; Foot	4	2   16	0.25   -	91.8%   84.1%	95.6%   86.8%	91.9%   77.4%	87.7%   87.7%
Cheek; Back; Upper arm; Shin	4	2   64	0.25   -	93.8%   87.7%	98.5%   92.6%	93.5%   83.9%	89.2%   86.2%
Chest; Hand; Foot	3	2   0.25	0.25   -	92.3%   82.6%	98.5%   88.2%	90.3%   71.0%	87.7%   87.7%
Back; Hand; Foot	3	16   4	0.5   -	89.7%   83.1%	94.1%   91.2%	75.8%   67.7%	98.5%   89.2%
<b>Chest; Upper arm; Shin</b>	3	4   1	0.25   -	<b>94.4%</b>   87.2%	98.5%   91.2%	91.9%   83.9%	92.3%   86.2%
Back; Upper arm; Shin	3	0.5   2	0.125   -	93.8%   85.1%	98.5%   85.3%	93.5%   83.9%	89.2%   86.2%
Back; Shin; Hand	3	1   64	1   -	90.3%   83.1%	94.1%   91.2%	79.0%   83.9%	96.9%   73.8%
Cheek; Chest; Upper arm	3	0.25   0.25	0.125   -	87.7%   77.4%	89.7%   79.4%	88.7%   82.3%	84.6%   70.8%
Cheek; Chest; Foot	3	2   8	0.25   -	90.8%   85.1%	98.5%   89.7%	83.9%   77.4%	89.2%   87.7%
<b>Cheek; Upper arm; Shin</b>	3	4   32	0.25   -	90.8%   <b>88.2%</b>	98.5%   94.1%	82.3%   83.9%	90.8%   86.2%
Chest; Upper arm	2	32   64	0.25   -	88.2%   82.1%	97.1%   89.7%	79.0%   79.0%	87.7%   76.9%
<b>Upper arm; Shin</b>	2	1   256	0.125   -	90.3%   <b>87.2%</b>	98.5%   91.2%	79.0%   83.9%	92.3%   86.2%
Back; Shin	2	4   128	0.25   -	89.1%   82.6%	92.3%   89.7%	87.0%   83.9%	88.7%   73.8%
Chest; Shin	2	1   0.25	0.125   -	87.2%   80.5%	95.6%   89.7%	75.8%   75.8%	89.2%   75.4%
Cheek; Shin	2	2   0.5	0.125   -	88.7%   85.1%	95.6%   89.7%	85.5%   79.0%	84.6%   86.2%
<b>Cheek; Back</b>	2	1   4	0.125   -	<b>90.8%</b>   82.1%	98.5%   88.2%	82.3%   83.9%	90.8%   73.8%

Note: Models showing the highest prediction accuracy are highlighted in bold fonts.

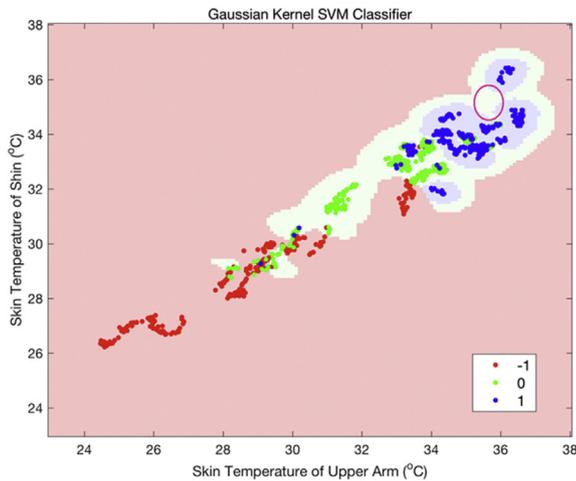


Fig. 6. SVM classifier with Gaussian kernels.

The areas with different background colors represent different combinations of shin and upper arm skin temperatures, and the associated thermal demands predicted by the models. Lines (Fig. 5) and curves (Fig. 6) that separate areas with different colors are learned through training data shown in scatter plots. In the Gaussian kernel model, cooling and heating demands are realistically separated by the neutral space. However, the model has higher complexity and is liable to overfitting especially in regions where the space has not been trained with data (no data) as shown in the red circle in Fig. 6. The SVM classifier with linear kernels is simpler. The hyperplanes can be interpreted by linear combinations of different local skin temperatures. The same skin temperature on the x-axis may correspond to three different thermal demands sometimes, and the skin temperature on the y-axis can help to distinguish the real sensation at those points. In summary, classifiers with the Gaussian kernel show slightly higher prediction accuracy, but require better training data set. The overfitting problem is more likely to happen for Gaussian kernel models than linear ones.

## 4. Experiment 2

Instead of a group of occupants, this experiment intends to validate the classification method for individual use. Varied clothing insulation is considered. Two additional issues in practice are discussed in this part: a. how many training data are needed before the classifier can work well automatically; b. the performance of a preset classifier for new occupants.

### 4.1. Methods

This experiment followed one male occupant in a private office in China for seven days from March to April. As the purpose was to accumulate physiological data in different thermal states, the environment was preset to a given temperature (Table 6) and was

regulated later depending on the occupant's responses. Temperature regulation was implemented by a common split air conditioner. Two kinds of clothing insulation were considered in the experiment. One clothing ensemble was long sleeve shirt, long pants, socks and sneakers (whole body insulation: 0.6 clo); the other one was the same, plus a down packable jacket (1.15 clo). The two clothing ensembles were determined considering a person putting on/taking off his/her coat, which is a general way of adjusting clothing in daily life. The insulation levels were estimated according to ISO standard [25]. As shown in Table 6, each clothing insulation level was studied in several tests with different preset indoor temperature. Apart from the 0.6 clo, 14 °C condition, a total of 13 tests were conducted. Air temperature and relative humidity (TR-72, Japan) were recorded at the working station and outdoors.

The subject entered the room after the thermal environment had been preconditioned to a stable temperature. We accepted 1 °C deviation from the target temperature during the presetting process. PyroButton (OPULUS Ltd, America, 0.2 °C precision) was attached at five locations to measure skin temperatures every 10 s: neck, upper back, left wrist, left hand, right shin. The occupant stayed in the room for 1 h without restricting his activity (primarily sedentary, sometimes strolling) and then reported his current thermal demands for the environment with “prefer a warmer condition” (-1), “no change is needed” (0) or “prefer a cooler condition” (1). Temperature was then adjusted by a two-degree step towards the neutral condition. The time for the office temperature to reach to the new setpoints varied but was less than 10 min. After each regulation, the occupant needed to stay in the condition for 30 min before he reported his thermal demands again. This process was repeated until the subject reached neutral. A total of 59 votes from 13 tests were accumulated for the validations of classification models.

As the data used to tune the model should not coincide with that used to evaluate the model's performance, we did not tune classifier parameters in this study. Instead, all the models were trained with the parameters reported from Experiment 1. The twentyfold cross validation method was used to evaluate the performance of the SVM classifiers. To figure out the required sample size needed to train good classifiers, learning curves were plotted using Python programming language and the scikit-learn package [26].

### 4.2. Results and discussion

Significant improvement of performance is observed in Table 7 when a combination of skin temperatures is considered in the classifier instead of only one skin temperature. Overall prediction accuracy could reach as high as 94% with a linear kernel classifier. The support vectors (SVs) are the data points used in the classifier to determine the final decision boundaries. The less number of support vectors means a smaller Vapnik-Chervonenkis (VC) dimension, which brings forth better model generalization. Consequently, the linear kernel SVM classifier using back, shin and hand skin temperatures performs best here.

The learning curves show the prediction accuracy as functions of training sample size (Fig. 7). The performance of the model to

Table 6  
Background information of the dataset.

Preset indoor T (°C)	Clothing insulation (clo)	Working station		Outdoors	
		T (°C)	RH (%)	T (°C)	RH (%)
14, 16, 18 20, 22, 24, 26	0.6 and 1.15	14.1–25.8	34–58	5.3–19.3	31–54

Note: T denotes air temperature, RH denotes relative humidity.

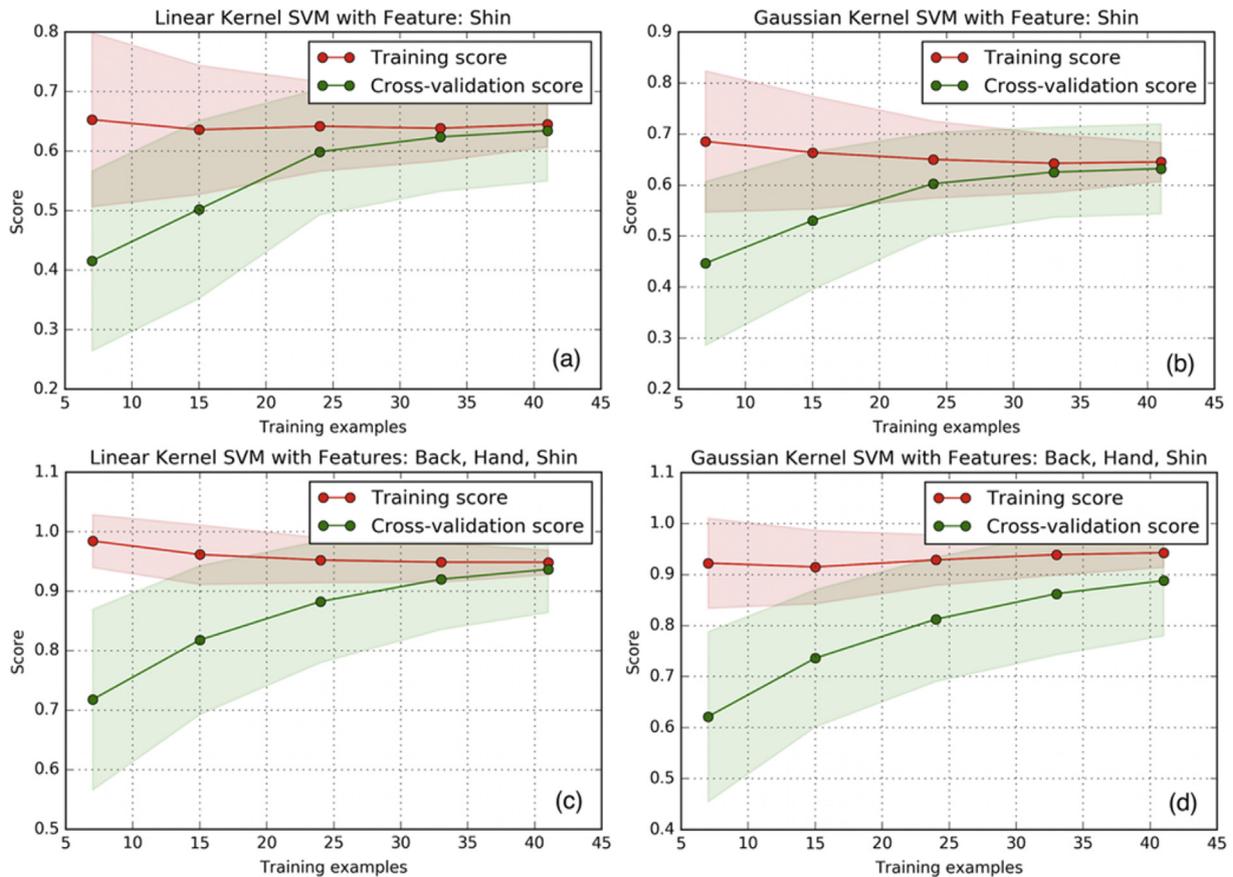
**Table 7**  
Twentyfold cross-validation results of two SVM classifiers.

Input features	Kernel	Accuracy	Number of SVs for each class		
			-1	0	1
Shin	Gaussian	63.2%	14	15	11
	Linear	62.7%	15	20	11
Back; Shin	Gaussian	87.3%	7	17	10
	Linear	87.2%	3	11	7
Back; Shin; Hand	Gaussian	92.6%	13	16	10
	Linear	94.1%	2	7	6

predict data in training dataset is illustrated by the training scores and the performance of predicting future data is illustrated by the cross-validation scores in the figures. These two scores will finally be equivalent when the training sample size is large enough, and the model performance will not be improved even though more data are trained. The dataset was randomly split 100 times to compute an average performance. Standard deviations are displayed in the shaded zones. Results demonstrate that the learning processes become shallow after 25 data points for the SVM classifier with only one input. To reach a rate of 80% prediction accuracy, 14 training data are needed for linear kernel SVM with three features, while 23 are required for the Gaussian kernel model. The three-inputs linear model would predict the thermal demands with 90% accuracy after 28 training samples. One should notice that these results are general recommendations instead of precise conclusions, as one would encounter different training data and predictions in practice.

In some cases, there may be no chances for a classifier to learn data. An example is in a conference room where occupants stay only temporarily. It is reasonable to preset a well-trained model in the controller. However, significant individual differences may exist in the room population, considering different skin temperature responses with the same thermal preference. As a result, one may be willing to know how well a preset SVM classifier could work for new occupants. We used the model trained in Experiment 2 to predict the data in Experiment 1. Prediction results (Output Class) were compared with true votes (Target Class) in Fig. 8. Each column in a confusion matrix represents the instances in a true group. Take the first column in Fig. 8(a) as an example: there are a total of 290 samples in which the occupants reported heating demands. 20 of them are correctly detected as shown in the green box, and 270 instances are misclassified as neutral. Grey boxes along the bottom show the probability of detection for each class and those in the right-hand column show the reliability of each prediction. Consequently, the model with the Gaussian kernel mistakes almost all the heating and cooling demand conditions as neutral. The overall prediction accuracy is only 29.2%. The linear kernel SVM classifier works much better in identifying heating (92.8%) and cooling demand (78.5%) conditions. However, the model misclassifies around 88% of neutral votes as cooling demand ones and the remaining 12% as heating demand ones. It predicts the thermal demands correctly at an overall rate of 54.2%. In this case, the comfort of occupants might be more continuously provided, but more energy would be wasted as the heating and cooling systems would continue running in neutral conditions.

Compared with the results of models using as input shin skin temperatures from Experiment 1, the performance using shin



**Fig. 7.** Accuracy learning curves for one person thermal demands prediction.

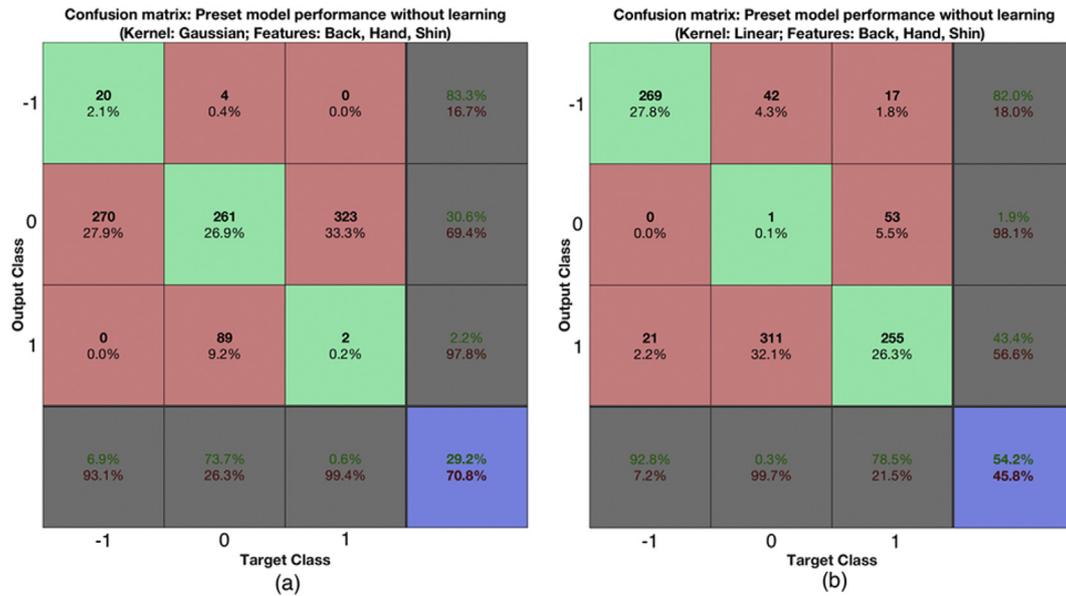


Fig. 8. The performance of preset model for new occupants.

temperatures from Experiment 2 was significantly lower. This is because upper-body clothing varied in Experiment 2 which was not reflected in the shin temperature. The whole-body clothing insulating was increased by adding a jacket while the local clothing insulation for the shin did not change. As a result, the occupant was likely to report the same demands with larger whole-body clothing insulation in a colder condition, while the shin skin temperature may have actually decreased. If this explanation stands up, the models using back skin temperature might perform better. We therefore trained the classifiers using back skin temperature. Twentyfold cross-validation results showed an accuracy of 83.1% for the linear kernel model and 81.4% for the Gaussian kernel model. Practically, varying clothing ensembles make it more difficult for single one skin temperature to predict thermal demands precisely. A combination of skin temperatures from different body segments helps to solve this situation.

As two different levels of clothing insulation were studied, and

the SVM classifiers with combined skin temperatures could still work well, it seems that using skin temperature instead of physical parameters like room temperature in machine learning algorithms is successful in minimizing the effect that varying clothing has on model accuracy. However, the use of skin temperature may produce contrary results when sweating happens. Consequently, like Humphreys et al. [17] proposed, a combination of physical and physiological parameters might be more powerful when the cases become complicated.

Skin temperature has been used to predict the thermal state of humans. We reviewed the relevant studies over the last ten years, and incorporated their modeling methods, applicable environments, and validation results in Table 8. As we can see, the statistical regression is the common modeling method used in these studies. The challenge with this is that a regression equation has to be preselected. The question is whether these equations are the proper ones to describe real-world thermal perception [34]. By

**Table 8**  
Recent studies (within 10 years) using skin temperature to predict whole-body thermal states.

Year	Authors	Model inputs	Description	Applicable environment and model performance
2007	Luo X et al. [27]	Core temperature, skin temperature and its change rate (10 body parts)	Use Fussy Neural Network (FNN) to predict local thermal sensation, and then to predict overall thermal sensation.	Uniform and non-uniform, steady state and transient. Not validated
2009	Nakayama et al. [28]	Left index finger skin temperature	Calculate the frequency and gradient of the skin temperature to evaluate thermal sensation.	Large mean square errors for overall sensation prediction; better for local thermal sensation estimation.
2010	Zhang H et al. [11–13]	Skin temperatures, the time derivatives of skin and core temperatures, MST (19 body segments)	Use regression methods to predict local sensation firstly, and then predict whole-body thermal sensation based on local sensations.	Uniform and non-uniform, steady state and transient. R <sup>2</sup> is over 0.9, residual SD is less than 0.75 (9-point scale)
2011	Liu W et al. [29]	Compare 26 types of MST calculation methods and recommend the 10-site one	Linear regression: $MSV = a \cdot MST + b$	Transient: R <sup>2</sup> is over 0.9. Not validated
2013	Takada et al. [30]	MST and its time differential (7-point method)	A nonlinear multiple regression model based on normalized MST and its time differential	Homogeneous, steady and non-steady state: R <sup>2</sup> = 0.873
2014	Liu H et al. [31]	Skin temperature (8-point method) and environmental parameters	Calculate the heat loss from skin temperature and build linear models of thermal sensation based on heat loss.	Transient: R <sup>2</sup> is over 0.9. Not validated
2016	Ghahramani et al. [32]	Facial skin temperatures (4 points)	Linearly combine the skin temperatures on different facial points to define the thermal uncomfortable zones	Steady and non-steady state: 95% confidence level
2016	Sim et al. [33]	Wrist (3 points) and finger skin temperatures	Stepwise regression models based on MST, temperature gradient, temperature time differential and average power of frequency band.	Stable and transient. Mean RMSE: $1.06 \pm 0.29$ for personalized model.

Note: MST indicates mean skin temperature. MSV means the mean sensation vote. R<sup>2</sup> is the regression determination coefficient. RMSE is root mean square error. SD means standard deviation.

following certain optimal rules, machine learning algorithms can help solve this difficulty and make the best prediction for the dataset. Their iterative aspect is a great advantage to be used in a control system to adapt independently when exposed to new data. Another point we can grasp from Table 8 is that more researchers now pay attention to practical and simple measurements to predict thermal states [26,30,31]. In this paper, we tried to systematically evaluate the performance of models using different skin temperature points and their combinations. The results have shown that in homogenous, steady-state conditions, as few as two proper locations of skin temperatures are sufficient to predict the thermal demands with high accuracy.

## 5. Conclusions

In this paper, we applied SVM classifiers with linear and Gaussian kernels to predict occupant's thermal demands in steady states. Skin temperatures were used as the only inputs of the models. A controlling prototype was proposed to intelligently control heating and cooling systems based on these machine learning models.

The attractive prediction power showed that skin temperatures are effective inputs to identify the thermal demands of occupants in steady state. A combination of the skin temperatures outperforms single one skin temperature in the model. Generally, three local skin temperatures from different body segments have contained enough information for the classification. More skin locations are not necessary and may cause the curse of dimensionality.

A single skin temperature could guarantee an accuracy over 80% at most when the clothing was controlled. The prediction accuracy would decrease significantly if the thermal states were maintained by regulating clothing insulation for other body segments in different conditions. To deal with this problem, using skin temperatures from different segments is an effective way.

With respect to the locations of skin temperature, we recommend shin and upper arm for models with 2 inputs. The model can be improved by adding one more skin temperature at chest. Considering only uncovered locations, we recommend cheek and hand. Some locations could be replaced with others in the same segment like back for chest, forehead for cheek. The performance would be influenced in a small way. Different combinations and the corresponding model parameters could refer to Tables 2–5. However, we insist that a grid search technique be used for new classifiers to tune the model parameters.

We prefer the SVM classifier with a linear kernel to that with a Gaussian kernel in general. The former is simpler and less likely to cause overfitting problem. The preset model with a linear kernel outperforms that with a Gaussian kernel for new occupants. The SVM classifier can work well with a small amount of training data. 28 samples are needed for the three-inputs linear kernel model to achieve 90% prediction accuracy. To achieve 80% accuracy level, 14 training data were required.

Skin temperatures have more power in predicting heating demands than cooling demands because of two possible reasons. One is that skin temperatures in different body segments vary much more in cool than in warm environments, especially for the body's extremities. The other reason is that for each location, the difference of skin temperatures between neutral and cool conditions is much bigger than it is between neutral and warm conditions, a phenomenon demonstrated by the effect sizes.

This paper is the first step of a continuing study to build innovative controllers for heating and cooling systems. The next step is to build the prediction models for human demands in transient states. Both the wearable device and the intelligent house are becoming common concepts. We hope that this study can help to

set up a foundation for developing thermal wearable devices and new ways of controlling buildings.

## Acknowledgements

This work is financially supported by the National Science Foundation of the United States [grant number 20161719] and the key project of National Natural Science Foundation of China [grant number 51238005].

## References

- [1] ASHRAE, Standard 55-2013: Thermal Environmental Conditions for Human Occupancy, American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE), Atlanta, 2013.
- [2] CEN15251: Indoor Environmental Input Parameters for Design and Assessment of Energy Performance of Buildings Addressing Indoor Air Quality, Thermal Environment, Lighting and Acoustics, European Committee for Standardization, Brussels, 2007.
- [3] C. Huizenga, S. Abbaszadeh, L. Zagreus, E. Arens, Air quality and thermal comfort in office Buildings: results of a large indoor environmental quality survey, *Proc. Heal. Build. III* (2006) 393–397. <http://escholarship.org/uc/item/7897g2f8?jsessionid=CEA1E13173D8003D5F74BD638E71785C>.
- [4] Z. Mao, F.L. Wang, T. Gao, Y.C. Dai, Q.C. Zhao, Y. Zhao, B. Sun, J. Guo, F. Zhang, Research of the room occupant complaining behavior pattern for the indoor environmental control, *Adv. Mater. Res.* 374–377 (2011) 1064–1067. <http://dx.doi.org/10.4028/www.scientific.net/AMR.374-377.1064>.
- [5] C.C. Federspiel, H. Asada, User-adaptable comfort control for HVAC systems, *Am. Control Conf.* 1992 (116) (1992) 2312–2319. <http://dx.doi.org/10.1115/1.2899242>.
- [6] M. Hamdi, G. Lachiver, F. Michaud, A new predictive thermal sensation index of human response, *Energy Build.* 29 (1999) 167–178. [http://dx.doi.org/10.1016/S0378-7788\(98\)00054-1](http://dx.doi.org/10.1016/S0378-7788(98)00054-1).
- [7] W. Liu, Z. Lian, B. Zhao, A neural network evaluation model for individual thermal comfort, *Energy Build.* 39 (2007) 1115–1122. <http://dx.doi.org/10.1016/j.enbuild.2006.12.005>.
- [8] H. Li, H. Yang, E. Li, Z. Liu, K. Wei, Wearable sensors in intelligent clothing for measuring human body temperature based on optical fiber Bragg grating, *Opt. Express* 20 (2012) 11740–11752. <http://dx.doi.org/10.1364/OE.20.011740>.
- [9] B.F. Jones, P. Plassmann, Digital infrared thermal imaging of human skin, *IEEE Eng. Med. Biol. Mag.* 21 (2002) 41–48.
- [10] D. Fiala, Dynamic Simulation of Human Heat Transfer and Thermal Comfort, Ph. D. Thesis, Institute of Energy and Sustainable Development, De Montfort University, Leicester, 1998. [http://www.utci.org/cost/publications/ICEE\\_2007\\_-\\_Abstract-UTCI-Fiala.et.al.pdf](http://www.utci.org/cost/publications/ICEE_2007_-_Abstract-UTCI-Fiala.et.al.pdf).
- [11] H. Zhang, E. Arens, C. Huizenga, T. Han, Thermal sensation and comfort models for non-uniform and transient environments: Part I: local sensation of individual body parts, *Build. Environ.* 45 (2010) 380–388. <http://dx.doi.org/10.1016/j.buildenv.2009.06.018>.
- [12] H. Zhang, E. Arens, C. Huizenga, T. Han, Thermal sensation and comfort models for non-uniform and transient environments, part II: local comfort of individual body parts, *Build. Environ.* 45 (2010) 389–398. <http://dx.doi.org/10.1016/j.buildenv.2009.06.018>.
- [13] H. Zhang, E.A. Arens, C. Huizenga, T. Han, Thermal sensation and comfort models for non-uniform and transient environments, part III: whole-body sensation and comfort, *Build. Environ.* 45 (2010) 399–410. <http://dx.doi.org/10.1016/j.buildenv.2009.06.020>.
- [14] H. Zhang, Human Thermal Sensation and Comfort in Transient and Non-uniform Thermal Environments, Ph.D. Thesis, Architecture, University of California, Berkeley, 2003.
- [15] Y. Taniguchi, H. Aoki, K. Fujikake, Study on car air conditioning system controlled by car occupants' skin temperatures—Part 1: research on a method of quantitative evaluation of car occupants' thermal sensations by skin temperatures, *SAE Tech. Pap. Ser.* 920169 (1992).
- [16] D. Wang, H. Zhang, E. Arens, C. Huizenga, Observations of upper-extremity skin temperature and corresponding overall-body thermal sensations and comfort, *Build. Environ.* 42 (2007) 3933–3943. <http://dx.doi.org/10.1016/j.buildenv.2006.06.035>.
- [17] M.A. Humphreys, K.J. McCartney, J. Nicol, I.A. Raja, An analysis of some observations of the finger temperature and thermal comfort of office workers, in: *Proc. 8th Int. Conf. Indoor Air Qual. Clim. (Indoor Air 99)*, Edinburgh, 1999.
- [18] M. Galar, A. Fernández, E. Barrenechea, H. Bustince, F. Herrera, An overview of ensemble methods for binary classifiers in multi-class problems: experimental study on one-vs-one and one-vs-all schemes, *Pattern Recognit.* 44 (2011) 1761–1776. <http://dx.doi.org/10.1016/j.patcog.2011.01.017>.
- [19] C.-W. Hsu, C.-C. Chang, C.-J. Lin, A practical guide to support vector classification, *Tech. Rep.* 101 (2003) 1396–1400. <http://dx.doi.org/10.1177/02632760022050997>.
- [20] W.S. Noble, What is a support vector machine? *Nat. Biotechnol.* 24 (2006) 1565–1567. <http://dx.doi.org/10.1038/nbt1206-1565>.
- [21] C. Huizenga, H. Zhang, E. Arens, D. Wang, Skin and core temperature response

- to partial- and whole-body heating and cooling, *J. Therm. Biol.* 29 (2004) 549–558, <http://dx.doi.org/10.1016/j.jtherbio.2004.08.024>.
- [22] E. Arens, H. Zhang, C. Huizenga, Partial- and whole-body thermal sensation and comfort - Part I: uniform environmental conditions, *J. Therm. Biol.* (2006) 53–59, <http://dx.doi.org/10.1016/j.jtherbio.2005.11.028>.
- [23] G.M. Sullivan, R. Feinn, Using effect size - or Why the P Value is not enough, *J. Grad. Med. Educ.* 4 (2012) 279–282, <http://dx.doi.org/10.4300/JGME-D-12-00156.1>.
- [24] M.E. Rice, G.T. Harris, Comparing effect sizes in follow-up studies: ROC area, Cohen's d, and r, *Law Hum. Behav.* 29 (2005) 615–620, <http://dx.doi.org/10.1007/s10979-005-6832-7>.
- [25] ISO-9920–2004: Ergonomics of the thermal environment - Estimation of the thermal insulation and evaporative resistance of a clothing ensemble.
- [26] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É. Duchesnay, scikit-learn: machine learning in Python, *J. Mach. Learn. Res.* 12 (2012) 2825–2830, <http://dx.doi.org/10.1007/s13398-014-0173-7.2>.
- [27] X. Luo, W. Hou, Y. Li, Z. Wang, A fuzzy neural network model for predicting clothing thermal comfort, *Comput. Math. Appl.* 53 (2007) 1840–1846, <http://dx.doi.org/10.1016/j.camwa.2006.10.035>.
- [28] K. Nakayama, T. Suzuki, K. Kameyama, Estimation of thermal sensation using human peripheral skin temperature, *Conf. Proc. IEEE Int. Conf. Syst. Man. Cybern.* (2009) 2872–2877, <http://dx.doi.org/10.1109/ICSMC.2009.5346126>.
- [29] W. Liu, Z. Lian, Q. Deng, Y. Liu, Evaluation of calculation methods of mean skin temperature for use in thermal comfort study, *Build. Environ.* 46 (2011) 478–488, <http://dx.doi.org/10.1016/j.buildenv.2010.08.011>.
- [30] S. Takada, S. Matsumoto, T. Matsushita, Prediction of whole-body thermal sensation in the non-steady state based on skin temperature, *Build. Environ.* 68 (2013) 123–133, <http://dx.doi.org/10.1016/j.buildenv.2013.06.004>.
- [31] H. Liu, J. Liao, D. Yang, X. Du, P. Hu, Y. Yang, B. Li, The response of human thermal perception and skin temperature to step-change transient thermal environments, *Build. Environ.* 73 (2014) 232–238, <http://dx.doi.org/10.1016/j.buildenv.2013.12.007>.
- [32] A. Ghahramani, G. Castro, B. Becerik-Gerber, X. Yu, Infrared thermography of human face for monitoring thermoregulation performance and estimating personal thermal comfort, *Build. Environ.* 109 (2016) 1–11, <http://dx.doi.org/10.1016/j.buildenv.2016.09.005>.
- [33] S.Y. Sim, M.J. Koh, K.M. Joo, S. Noh, S. Park, Y.H. Kim, K.S. Park, Estimation of thermal sensation based on wrist skin temperatures, *Sensors* 16 (2016), <http://dx.doi.org/10.3390/s16040420>.
- [34] G.S. Brager, R.J. de Dear, Thermal adaptation in the built environment: a literature review, *Energy Build.* 27 (1998) 83–96, [http://dx.doi.org/10.1016/S0378-7788\(97\)00053-4](http://dx.doi.org/10.1016/S0378-7788(97)00053-4).