Build Simul (2024) 17: 839–855
https://doi.org/10.1007/s12273-024-1114-9

Predicting the clothing insulation through machine learning algorithms: A comparative analysis and a practical approach

Pablo Aparicio-Ruiz, Elena Barbadilla-Martín (✉), José Guadix, Jesús Muñuzuri

Grupo de Ingeniería de Organización, Escuela Técnica Superior de Ingeniería, Universidad de Sevilla, Camino de los Descubrimientos S/N, 41092, Seville, Spain

Research Article

Abstract

Since indoor clothing insulation is a key element in thermal comfort models, the aim of the present study is proposing an approach for predicting it, which could assist the occupants of a building in terms of recommendations regarding their ensemble. For that, a systematic analysis of input variables is exposed, and 13 regression and 12 classification machine learning algorithms were developed and compared. The results are based on data from 3352 questionnaires and 21 input variables from a field study in mixed-mode office buildings in Spain. Outdoor temperature at 6 a.m., indoor air temperature, indoor relative humidity, comfort temperature and gender were the most relevant features for predicting clothing insulation. When comparing machine learning algorithms, decision tree-based algorithms with Boosting techniques achieved the best performance. The proposed model provides an efficient method for forecasting the clothing insulation level and its application would entail optimising thermal comfort and energy efficiency.

Keywords

clothing insulation simulation
adaptive thermal comfort
behavioural adaptive actions
machine learning

Article History

Received: 04 October 2023
Revised: 09 January 2024
Accepted: 30 January 2024

1 Introduction

The clothing insulation level, standardized by clo, is used in thermal comfort models to represent the insulation of complete clothing ensembles when investigating thermal comfort situations and considering different metabolic rates, indoor and outdoor conditions, as well as it can also be applied when evaluating the thermal stress (ISO 2009a).

The clo unit was introduced in 1941 by Gagge et al. (1941) to define the relative resistance to heat transfer by clothing. Since clothing insulation level is a variable related to thermal stress and comfort, it is a feature that influences the design process of buildings, and even ISO 7730 (ISO 2006) and EN 16798-1 (EN 2020) define standardized values for it.

Currently, the ISO 7730 (ISO 2006) is used in the determination and interpretation of thermal comfort through the Predicted Mean Vote Model and the Predicted Percentage of thermally Dissatisfied occupants, where the clothing insulation level is an input parameter. EN 16798–1 (EN 2020) and ASHRAE Standard 55 (ANSI/ASHRAE 2020) include the adaptive approach, where the occupants are freely adapt to the environment through adaptation processes, which were classified by de Dear and Brager (1998) as physiological, psychological and behavioural, considering in the latter clothing changes. Additionally, ISO 11079 (ISO 2009b) considers the determination and interpretation of cold stress and defines the variable required clothing insulation (IREQ).

Therefore, the clothing insulation level is a factor that is present, both in the static thermal comfort approach, and in the adaptive thermal comfort approach. In the first one, the clothing insulation level is a quantitative factor to model occupants' thermal comfort and in the adaptive comfort model, it is used to explain occupants' behaviour and adaptive actions. Moreover, it has also been studied from the perspective of the heat and mass transfer from human skin to the environment, proposing different models based on the thermophysiological response (Joshi et al. 2019; Joshi et al. 2022).

The importance of the clothing level lies, not only in its relationship with thermal comfort, but it is also a decisive factor in individualized thermal comfort models and in energy consumption. Despite its relevance, the prediction of the

Advances in Modeling and Simulation Tools

E-mail: ebarbadilla@us.es

clothing insulation level is a challenging task, since it depends on multiple factors that are interdependent and can significantly vary between individuals and situations considering environmental factors (temperature, relative humidity or air velocity), gender, cultural components or other personal factors. The same clothing insulation level may not provide the same level of thermal comfort in different environmental conditions, therefore, it is a challenging task to develop a predictive model that accounts for all these factors and their interactions. Typically, a consistent clo level (0.5 clo in summer, 1.0 clo in winter) is usually considered on thermal comfort indices, aligning with the guidelines ISO 7730 (ISO 2006) or EN 16798-1 (EN 2020), which might not represent the reality of human clothing behaviour and could lead to overheating and overcooling situations. So, there is a need to improve knowledge and understanding of human behaviour in terms of an optimal selection of their ensemble and the possibility of applying predictive models.

Therefore, the clothing insulation level, would be helpful in the implementation of adaptive thermal comfort, e.g., by means of user assistance systems, that would improve the user's comfort and the energy efficiency of a building. Recommendations regarding clothing insulation level would be useful, for example, in terms of the dress code, that is, it could lead to relaxation of the clothing conventions of the workers in some companies according to the machine learning (ML) model.

Previous studies (Humphreys et al. 2013) have also highlighted the importance of between-day clothing changes, showing that predictions regarding clothing insulation levels would lead to improve adaptive actions of the occupants by adjusting their clothing to the indoor environment, before entering the building or once inside. Moreover, in the work carried out by Cui and Watanabe (2014), the "Cool Biz" and "Warm Biz" initiative is exposed as a potential measure to conditioning systems load reduction. Therefore, a model that would predict how building occupants change clothes would greatly improve conditioning systems efficiency and energy savings. Additionally, the possibility of developing a system for recommending an optimal clothing ensemble is feasible thanks to studies such as the ones carried out by Li and Chen (2021) and Liu et al. (2022a).

For that, the present work focuses on the study of the indoor clothing insulation levels following the research objectives explained below:

- Analysing different machine learning techniques and carrying out a comparative analysis between them, regarding their performance for predicting clothing level.
- Proposing an approach to apply machine learning techniques for predicting the clothing insulation of the users of a building, which could also be useful for giving them advice in order to avoid the overuse of the heating, ventilating and air conditioning (HVAC) systems.
- Identifying the most relevant variables when predicting clothing insulation based on a case study and proposing new ones.
- Broaden the knowledge based on indoor clothing level, considering a new type of climate and building conditioning.

The paper is structured as follows: Section 1 shows the literature review regarding the approaches and variables considered in other studies. Section 2 describes the case study in which the analysis is based on. Section 3 exposes the machine learning algorithms developed for predicting clothing insulation level, as well as the relevant parameters considered. Section 4 discusses the relevance of the input variables and the performance of the ML algorithms. Finally, Section 5 presents the main conclusions based on the results obtained.

For the reasons exposed previously, clothing insulation level has been an object of study in the literature. In 1970, the work carried out by Humphreys (1970) showed a relationship between the clothing insulation level and the comfort temperature. And Nicol and Roaf (1996) exposed a connection between clothing insulation level and indoor and outdoor air temperature in 1996.

Since then, various studies have focused on the analysis of the level of clothing. Table 1 shows a summary of the references that have been carried out in recent years with this objective. The type of building, their space conditioning approach, the location in where the studies were developed, as well as the technique used to analyse the clothing behaviour, are specified.

Previous studies have discussed the clothing insulation level from different approaches, in terms of buildings and locations considered, input variables taken into account and techniques for it.

With the rise of the Internet of Things and machine learning, more research builds on these principles to face new challenges (Joshi et al. 2019). In the field of knowledge of thermal comfort, alternative machine learning approaches have been proposed in recent years to delve into the study of thermal comfort and the variables that are related to it.

In this context, some studies in the literature propose a machine learning framework to estimate the clothing insulation using vision recognition technology (Zang et al. 2019; Choi et al. 2021; Medina et al. 2022) or even analysing the material properties of different garments (Antweiler et al. 2022). However, the majority of these studies predominantly centre around thermal comfort through the thermal sensation vote, the thermal preference vote, the thermal comfort vote or the indoor comfort temperature itself (Deng and Chen 2018; Chai et al. 2020; Wang et al. 2020; Liu et al. 2022b; Yu et al. 2022).

Table 1 Previous studies regarding clothing behaviour

Reference	Year	Building	Conditioning system[1]	Location	Technique
Morgan and de Dear 2003	2003	Office building	AC	Australia	Regression models
De Carli et al. 2007	2007	Office, residential	NV, AC	28 locations all over the world	Linear regression models
de Carvalho et al. 2013	2013	University classrooms	NV	Portugal	Multiple linear regression
Schiavon and Lee 2013	2013	Office	AC, FR	Australia, United States	Multiple linear mixed models
Jiao et al. 2017	2017	Residential	FR	China	Multiple linear regression model
Liu et al. 2018	2018	University	NV	China	Logistic functions
Ngarambe et al. 2019	2019	University classrooms	FC	Korea	Linear model and deep neural network
Wang et al. 2019a	2019	Office, school, residential	NV	Locations considered is ASHRAE Database	Statistical models: logistic, power, linear and logarithm transformed linear models
Ming et al. 2020	2020	Office building	MM	China	Linear regression model
Rupp et al. 2021b	2021	Office building	NV, AC, MM	Locations considered in ASHRAE Database	Multi regression analysis
Nakagawa and Nakaya 2021	2021	University classrooms	NV, AC	Japan	Weighted regression analysis
Rupp et al. 2021a	2021	Office building	MM	Brazil	Simple and multiple logistic models
Su et al. 2022	2022	All types	AC	China	Multivariable linear mixed models
Zhao et al. 2022	2022	Rural resident buildings	AC	China	Five-parameter logistic model, Multi-linear regression
Duhirwe et al. 2022	2022	Classroom, multifamily housing, office, senior centre, others	NV, AC, MM	Locations in the ASHRAE Database	Convolutional Neural Network
Zhuang et al. 2022	2022	University classrooms	NV	China	Linear regression model

[1]Conditioning system: naturally ventilated (NV), air-conditioned (AC), mixed-mode (MM), fan foil units (FC).

Nevertheless, there are fewer studies that use complex ML techniques for the analysis of the clothing insulation level and its connection with other variables, like indoor variables, outdoor variables or personal features (Table 2): Ngarambe et al. (2019) proposed the prediction of mean daily clothing insulation level based on values from a field study with neural networks and Duhirwe et al. (2022) developed a convolutional neural network in order to predict the occupants' clothing insulation. Therefore, this research aims to bridge this gap by introducing an analytical approach that includes 13 regression and 12 classification machine learning algorithms and by delving into the study of variables related to clothing insulation, proposing new factors not previously considered. All of that, for a new combination of type of weather and building conditioning not existing in other studies.

It can be observed that most investigations carried out in terms of analysing clothing behaviour focus on statistical models considering linear regression model, multiple regression models or more complex statistical functions like logistic or logarithm. Nevertheless, the application of machine learning techniques would usually lead to better accuracy when predicting a variable under study, as well as to discover the complex relationships between output variables and input parameters (Duhirwe et al. 2022; Su et al. 2022).

The whole previous studies (listed in Table 1) highlight the importance of the clothing insulation level and how it influences the comfort level in a building, which could affect the optimal indoor comfort temperature in several degrees centigrade. And although some researches consider machine learning techniques in order to analyse the clothing insulation, it is necessary to further investigate the impact that different variables have on its prediction, as well as to consider different type of buildings, climatology, and other machine learning algorithms.

As for the variables that are related to clothing insulation level, Table 2 lists the key variables for understanding occupants' clothing behaviour that the references showed in Table 1 considered.

Numerous references consider indoor temperature and outdoor temperature variables as main parameters linked to clothing insulation. However, indoor and outdoor features do not describe the whole human behaviour regarding clothing adaptation (Duhirwe et al. 2022; Zhao et al. 2022) therefore, it is also important to delve into the factors that are potentially linked to it.

Regarding the machine learning algorithms, neural

Table 2 Key variables related to clothing behaviour

Reference	Input variable
Morgan and de Dear 2003	Indoor temperature, Outdoor temperature, Outdoor maximum temperature
De Carli et al. 2007	Indoor air temperature, Outdoor temperature at 6 a.m., Outdoor mean daily temperature, Outdoor mean temperature, Outdoor weighted value on the temperature over last 4 days
de Carvalho et al. 2013	Maximum outside temperature during that day, Maximum outside temperature during previous days, Outdoor running mean temperature, Effective temperature inside (ET)
Schiavon and Lee 2013	20 variables from ASHRAE Database, highlighting Outdoor air temperature, Indoor operative temperature, Indoor relative humidity
Jiao et al. 2017	Age, Indoor air temperature, Indoor relative humidity
Liu et al. 2018	Outdoor running mean temperature
Ngarambe et al. 2019	Outdoor temperature at 6 a.m., Outdoor dew point temperature at 6 a.m., Gender, Season, Mode of transport
Wang et al. 2019a	Climate, Season, Building type, Indoor temperature, Outdoor temperature
Ming et al. 2020	Outdoor temperature
Rupp et al. 2021b	Indoor air temperature, Season, Building ventilation type
Nakagawa and Nakaya 2021	Indoor operative temperature
Rupp et al. 2021a	Thermal sensation vote, Previous clothing insulation, Indoor operative temperature
Su et al. 2022	Indoor air temperature, Climate zone, Building type
Zhao et al. 2022	Age, Indoor operative temperature, 7-days running mean outdoor temperature, daily outdoor temperature
Duhirwe et al. 2022	Personal factors, Building factors, Indoor and outdoor environmental factors, Thermal comfort parameters, Climate
Zhuang et al. 2022	Indoor operative temperature

networks have been considered (Table 1), but no other ML techniques have been contemplated nor there is a detailed comparative analysis between such techniques that deeply examines their development (Ngarambe et al. 2019; Duhirwe et al. 2022).

It is also important to underline that the models previously obtained may be not applicable to other type of climate, conditioning approach or type of building that the ones considered in the specific studies (Wang et al. 2019a; Su et al. 2022). It is necessary to further investigate in terms of new contexts in order to reinforce the base of knowledge of adaptive actions regarding clothing behaviour (Rupp et al. 2021a).

Therefore, the present study focuses on the relevant variables regarding clothing insulation and on predicting the individual clothing insulation level when users felt comfort. In particular, the comfort clothing insulation level refers to the level of clothing in which users feel comfortable when no additional adaptive actions, such as turning on the heating, ventilating and air conditioning systems would be necessary (Rupp et al. 2021b; Su et al. 2022). A systematic analysis of the variables that influence the clothing insulation level and a comparison between 13 regression ML algorithms and 12 classification ML algorithms were carried out.

The investigation is focused on mixed-mode office buildings, that are not usually aim of study in the literature, compared to NV or AC buildings, and on the southwestern area of Spain, where no studies regarding the clothing insulation have been identified. Since the paper focuses on the clothing insulation level, as well as the comfort clothing ensemble for a MM building, the below analysis and results will be shown considering the whole data (defined as the MM mode) and only the data when the HVAC system was not in use (defined as the FR mode) respectively.

The analysis and the results would also be useful when implementing it into a heating, ventilating and air-conditioning system of a building, so that it could assist the occupants when making decisions regarding their comfortable clothing ensembles. This advice would help the occupants to make adaptive decisions that prioritize the adaptation of clothing over the application of active actions like turning on the HVAC system for achieving thermal comfort.

2 Case study

The present work is based on data collected from a field study exposed in Barbadilla-Martín et al. (2017). Its main characteristics are explained below.

2.1 Sample characteristics and climatology

The field study was carried out in Seville, Spain, which is categorised within temperate climates (Csa), according to the Köppen-Geiger classification (Kottek et al. 2006). Its climatology is distinguished by mild winters and very hot summers, ranging the average daily temperature in the winter season from 10 °C to 15 °C (October–February). In the summer season (May–September), such variable usually rises up to 30 °C, with maximum values about 40 °C.

Eleven rooms from three office buildings were selected for the field study, two of them were mixed-mode buildings, combining natural ventilation and user-controlled backup conditioning systems, and the third building was fully conditioned.

The number of participants was 54, aged between 20 and 60 years. Regarding the age distribution, most participants were between 20 and 40 years old (30% of the occupants were between 20 and 30 years old, 43% between 30 and 40 years old). Nine of the participants were women. They were all performing the same sedentary office activity, which metabolic rate is 1.2 met according to ISO 7730 (ISO 2006).

2.2 Data collection

During the investigation, environmental variables were monitored, both indoors and outdoors, and the participants were surveyed during a whole year, in the morning and in the afternoon, as well as the thermal comfort of the occupants through surveys and additional information related to them.

We measured indoor air temperature (T_{air}), globe temperature (T_{g}), relative humidity (RH), air velocity (V_{air}) at 15-minute intervals, by installing devices near the participants in each of the room considered in the field study. Recommendations of ISO 7726 standard (ISO 2002) regarding the specifications of monitoring equipment and their location were fulfilled. The model of each measuring device is shown in Barbadilla-Martín et al. (2017)-Table 3 as for the location and height of them.

The following outdoor variables were also monitored during the field study: outdoor air temperature ($T_{\text{air ext}}$), outdoor temperature at 6 a.m. ($T_{\text{out 6 am}}$), dew point at 6 a.m. (DWPT_{6am}), globe temperature ($T_{\text{g ext}}$), relative humidity (RH_{ext}), relative humidity at 6 a.m. (RH_{6am}) and barometric pressure (atm_{ext}). Based on Nicol and Humphreys (2010), we calculated the outdoor running mean temperature with an alpha equal to 0.8 ($T_{\text{rm 0.8}}$). Similarly, we also estimated the outdoor running mean relative humidity ($\text{RH}_{\text{rm 0.8}}$), outdoor temperature at 6 a.m. ($T_{\text{6amrm 0.8}}$), dew point at 6 a.m. ($\text{DWPT}_{\text{6am0.8}}$) and relative humidity at 6 a.m. ($\text{RH}_{\text{6amrm 0.8}}$).

We also surveyed the participants in order to collect information regarding their assessment of the indoor environment, as well as their clothing. The participants were asked to indicate their thermal sensation vote (five-points scale: warm, slightly warm, neutral, slightly cool, cool) and their thermal preference vote (five-point scale: much cooler, a bit cooler, no change, a bit warmer, much warmer). Aspects regarding to the age or the gender of the participants were also included in the questionnaire and we calculated indoor temperature comfort based on the Griffith method (T_{comfort}).

Additionally, a clothing questionnaire form included a list of individual garments in order the participants selected their clothing level. The clothing questionnaire contained aspects related to the shirts or blouses, trousers or dress/skirts, sweaters and jackets, socks and shoes the participants were wearing (ISO 2009a). The whole questionnaires were provided to the participants through a web page, which automatically registered the date it was fulfilled (month, hour and season).

In total, 3352 questionnaires were collected, but 2883 questionnaires after processing and eliminating data that contained missing values or errors were considered in the present study. Of them, 1208 questionnaires were collected when the HVAC was not in use.

2.3 Clothing insulation level

The insulation for each participant's ensemble was calculated by adding the insulation of their individual garments collected in the clothing questionnaire. Such method is proposed in ISO 9920 (ISO 2009a) and is widely applied in field studies (de Carvalho et al. 2013; Liu et al. 2018; Ngarambe et al. 2019).

Figure 1 depicts the distribution of the intrinsic clothing insulation (Icl) as a combination of the individual garments collected in questionnaires during the sampling period. In particular, the figure presents data clothing ensembles for comfort situations (based on the thermal sensation votes between −1 and 1) for both, the building operating in mixed-mode (MM) and the data when the HVAC system was not in used (FR), showing the frequency or the number of times certain value was collected. It can be observed that the dataset is proportional and relatively balanced.

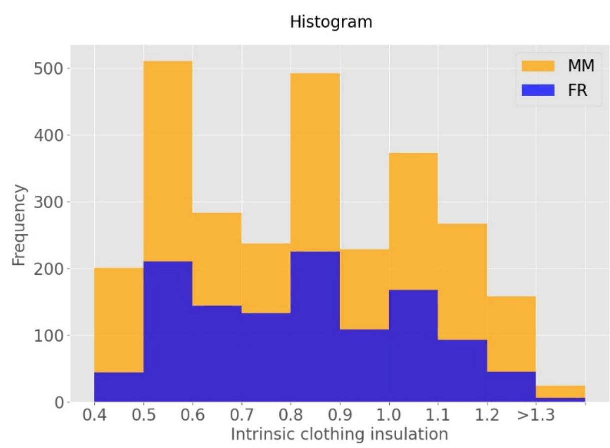


Fig. 1 Distribution of the intrinsic clothing insulation level

According to ISO 9920 standard (ISO 2009a), the most representative clothing insulation values can be identified around 0.5 and 1, concerning the summer and winter period respectively. However, other combinations of individual garments could also be determined as optimal values for the clothing insulation, depending on the climate or the type of building, which agrees with previous studies in the literature (de Carvalho et al. 2013; Jiao et al. 2017; Ngarambe et al. 2019; Wang et al. 2019a).

As stated by the ASHRAE and SCATs database (Arens et al. 2010), in general, for a certain indoor temperature, the clothing insulation changes about 0.8 clo unit, ranging from 0.4 to 1.2 clo.

3 Methodology

In the present section, the methodology for analysing the relevance of input variables regarding clothing insulation levels and the development of machine learning models for predictions are shown. The process defined in Figure 2 comprises the phases explained below.

Firstly, (1.) it is necessary to formulate the problem, therefore a study database is required. The collected data are pre-processed (in this step, the treatment of the data, the data reading and the categorical values defined for each variable are justified). Secondly, (2.) in the dataset, a study of the main characteristics is carried out, in order to define an efficient and structured database. The data are processed (3.) and divided, in order to avoid overfitting or underfitting, and two dataset are used: the training dataset and the validation dataset. The training data train the model in (4.) and in (5.) such model is evaluated with the evaluation dataset, including a scored index to select the one which has the best performance. A protocol for evaluating the problem is established (in this step, the criteria for choosing the appropriate model and its adjustment to achieve the best performance are defined). After carrying out the whole phases, the selected machine learning model could be used to predict new data.

Since data analysis involves missing values and inconsistencies in the data, it is important to note the relevance of the pre-processing (1. and 2.) in order to eliminate duplications and redundant information in the sample data for reliable results. During this process, prior to the training of the different models, the input for machine learning algorithms should be analysed, but the raw data need a previous process until it could be considered correct for a consistent dataset without errors. For that, first of all, missing value, anomalous and atypical values should be identified and corrected or eliminated.

To deal with missing values, two general ways are usually considered (Fan et al. 2021). The first one is to simply

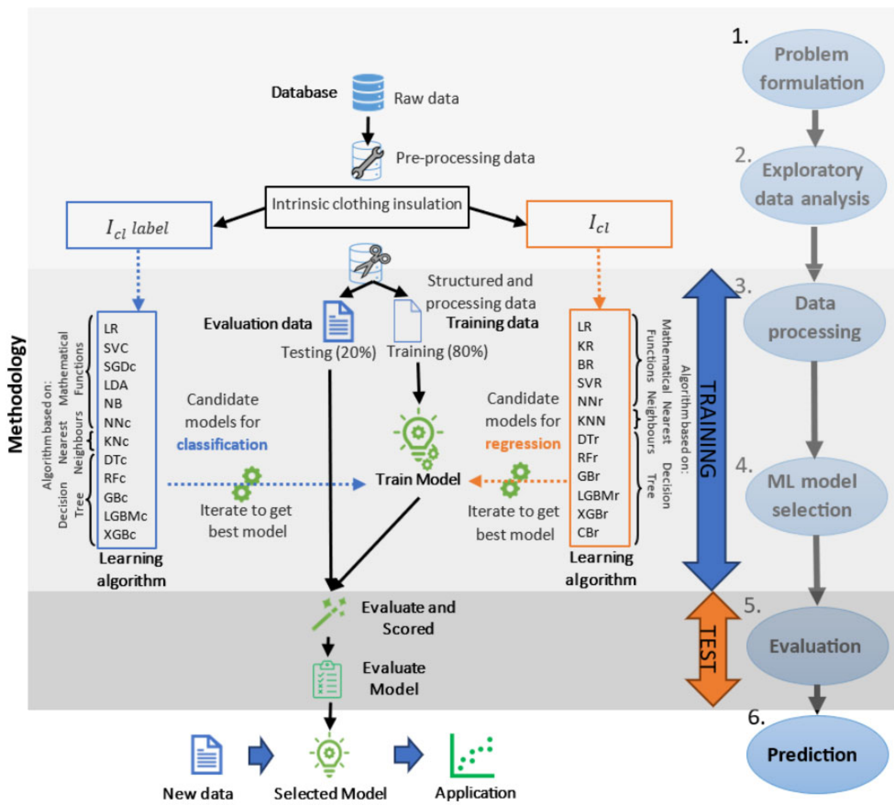


Fig. 2 Flowchart of the proposed methodology

discard samples with missing values, while the second one is to fill in the missing values.

Regarding missing values, techniques such as using a global constant to fill in for missing values, using a mean or median to fill the missing value, and using forward fill or backward fill method are commonly used.

As for noisy data, data smoothing techniques such as binning, clustering, and regression can be applied (Fan et al. 2021).

In the present study, we opted for not considering missing values or errors for the following analysis, by eliminating them from the original dataset.

The hierarchical structure we considered in the present study for categorising the algorithms is shown in Figure 2.

Firstly, algorithms based on mathematical functions contain regression models and predict a quantity, in regression problems, and label or category in classification problems. The logistic regression (LR) is a supervised ML model in regression and classification that identifies a linear connection among dependent and independent variables (Bishop 2006). Based on the classification, kernel ridge (KR) and bayesian ridge (BR) (Murphy 2012) also include weights in the models, they have a tendency to penalize the extremes of the weights. Support vector machine (SVM) can be applied to both, regression (SVR) and classification (SVC), they use different mathematical functions (kernel function) (Vapnik 1999; Tang et al. 2022). Derived from the SVM, the stochastic gradient descent (SGD) is an iterative method for optimizing by selecting a random data point, which would help to escape local minima and speeds up convergence (Mutlu and Acı 2022). Both, lineal discriminant analysis (LDA) and naïve bayes (NB), provide probabilities for class predictions (James et al. 2023). The mathematical functions involved in neural networks are fundamental to their ability to solve classification (NNc) and regression (NNr) problems. Artificial neurons, designed to mimic the behaviour of biological neurons, are organised into layers with activation functions, in general, those functions are linear function (Rocha et al. 2007).

Secondly, nearest neighbours algorithms. This is a clustering model of several categories known. In this category, the K-nearest neighbours (Song et al. 2017) algorithm is the most commonly used algorithm in regression (Song et al. 2017) and classification (Zhang et al. 2017)

Third category contains decision tree-based algorithms. A simple decision tree for regression (DTr) and classification (DTc) do not usually perform well enough, so the ensemble methods can be used to aggregate more trees. In this category, the random forest (RF) algorithm based on individual decision trees is the most common in regression (RFr) and classification (RFc) in the ensemble method named bagging consisting of building trees with bootstrapped samples with aggregating them (Breiman 2001; Cutler et al. 2012). Another ensemble method is boosting techniques, where gradient boost regression (GBr) and classification (GBc) (Friedman 2001) decision trees are method that combines weak learners, the individual decision trees, to obtain a stronger learner. All the trees are sequentially connected and each of them try to reduce the error of the previous tree. Due to the type of connection, this boosting technique is usually slower than others, but also more accurate. The light-gradient boosting machine regression (LGBMr) and classification (LGBMc) classifier (LGBMc) is a recent variant of previous algorithms based on trees (Ke et al. 2017). In this case, the tree grows vertically (grows tree leaf-wise) while another algorithm grows horizontally (grows level-wise) (Thai 2022). The algorithm will select the leaf with max "delta loss". When growing the same leaf, leaf-wise performs better than the level-wise algorithm. In general, it is not suitable for small datasets, due to overfitting (less 10k data). The extreme gradient boosting classifier (XGBc) and regression (XGBr) (Chen and Guestrin 2016) is a supervised predictive algorithm based on the principle of boosting. The idea behind boosting is to generate different "weak" models and improve their results in order to obtain a "stronger" model with better results. Finally, CatBoost (CBr) (Prokhorenkova et al. 2018), which name comes from "Category" and "Boosting" and builds symmetric trees, unlike XGB and LGBM. Leaves from the previous tree are split considering the following restriction: the feature-split pair for the lowest loss is considered and used for the whole nodes in a level.

Likewise, the adopted sampling method is specified. For regression methods, a fundamental step in the pre-processing of the data for improving the performance of the models is to use a method which scales the input data in order to be compared at the same scale. The method we considered in the present study is StandardScaler.

In the classification models, the output categorical data are handled, which requires a transformation from qualitative variables into quantitative ones. In this case, a set of labels has been defined for the clothing insulation level: Low (<0.6 clo), Medium-Low (0.6 clo–0.8 clo), Medium-High (0.8 clo–1.0 clo) and High (>1.0 clo). Figure 3 shows the distribution of the intrinsic clothing insulation level collected and the data distribution, based on the previous categorization.

Regarding the decision of input characteristics to predict the optimal comfort clothing insulation level, it is also important to highlight the pre-processing phase in order to avoid the excess of data.

An alternative, is the principal component analysis (PCA) technique, which is used to analyse large datasets (Friedman 2001) and improves finding relationships

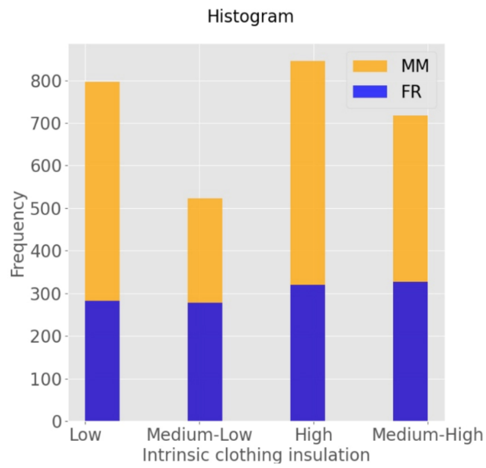


Fig. 3 Clothing insulation labels for classification methods

between predictor-response through a conversion and an adjustment of the input features (Kuhn and Johnson 2019). The reduction in the number of features analysed based on PCA seeks to find a small number of main components, which summarizes a large portion of the data variation (Aït-Sahalia and Xiu 2019) and simplifies the complexity of multi-dimensional spaces, while preserving their information.

Regarding the sampling method, we considered the raw dataset (Luo et al. 2020).

The "K-fold" cross-validation method is used in the present study for evaluating the resolution of the models. We used the 20-fold cross-validation, which has been proved to be a suitable value in the existing literature (Luo et al. 2020). Regarding the training and testing proportion, 80%–20% ratio is widely used in previous references (Gholamy et al. 2018; Wang et al. 2019b; Chai et al. 2020; Luo et al. 2020). Therefore, we considered the 80% of the dataset as training proportion.

For machine learning regression models, we used as evaluation metric the R2 in order to compare the result of our study with previous investigations, due to the fact that, most of them, analyse the clothing behaviour from a statistical point of view and R2 is the variable usually provided.

For machine learning classification models, accuracy and f1-score evaluation metrics were considered.

The accuracy rate of the model prediction is calculated by Equation (1):

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

where TP = true positives, TN = true negatives, FP = false positives and FN = false negatives.

The precision and the recall are shown in Equation (2) and Equation (3):

$$\text{Precision} = \frac{TP}{\text{all positive result}} = \frac{TP}{TP + FP} \tag{2}$$

$$\text{Recall} = \frac{TP}{\text{all correct case}} = \frac{TP}{TP + FN} \tag{3}$$

The balance between the precision and the recall is shown by f1-score (Sokolova et al.2006), defined as follows (Equation (4)):

$$\text{f1 - score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

4 Results and discussion

Firstly, the present section deals with an analysis of the relevance of the input variables and their relationship with the intrinsic clothing insulation level. Secondly, the performance of the 13 regression and 12 classification ML algorithms considered is also exposed and compared in order to define a final rank with them.

4.1 Exploratory data analysis: the relevance of input variables

The objective of the study is predicting the clothing insulation level based on other known data, so the investigation of the relationship between the input variables and the clothing insulation level and the relevance of such features is an important challenge.

As a result, and based on the references existing in the literature, a first dataset was obtained with twenty-one variables that could potentially be decisive in determining the clothing insulation level (Table 2), in which a comprehensive analysis was carried out for investigating their relevance.

The twenty-one variables were categorised into General information, Personal information, Comfort information, Indoor environment and Outdoor environment. General category comprises information regarding the month and the time the questionnaires were collected as well as the season (spring, summer, autumn or winter). Personal information refers to the gender of the participant. Comfort information includes the comfort temperature (T_{comfort}), calculated based on the Griffith method, and taking into account the indoor air temperature as well as the thermal sensation vote. Indoor environment comprises variables monitored during the field study and involves indoor air temperature (T_{air}), indoor relative humidity (HR), air velocity (V_{air}) and the temperature of a globe thermometer (T_{g}). Outdoor environment comprises variables monitored during the sample period and involves outdoor air temperature

($T_{\text{air ext}}$), outdoor relative humidity (HR_{ext}), the outdoor temperature of a globe thermometer ($T_{\text{g ext}}$), atmospheric pressure (Atm_{ext}), outdoor temperature at 6 a.m. ($T_{\text{out 6am}}$), dew point at 6 a.m. (DWPT_{6am}) and outdoor relative humidity at 6 a.m. (RH_{6am}). Additionally, we estimated the running mean temperature ($T_{\text{rm 0.8}}$) depending on the running mean temperature of the previous day and the average temperature of the previous day and with an alpha equal to 0.8 (Nicol and Humphreys 2010). Similarly, we figured the running mean relative humidity ($\text{RH}_{\text{rm 0.8}}$), the outdoor running mean temperature at 6 a.m. ($T_{\text{6amrm 0.8}}$), the running mean dew point ($\text{DWPT}_{\text{6am 0.8}}$) and the running mean relative humidity at 6 a.m. ($\text{RH}_{\text{6amrm 0.8}}$) in order to study feasible correlations with comfortable clothing insulation level and all of them with an alpha equal to 0.8.

Initially, the correlation of the variables with the clothing insulation level was studied and shown in Table 3. After analysing the set of variables, it could be inferred that some of them could be eliminated based on the higher correlation between input variables themselves, such as: hour or the globe temperature. Other variables require a further analysis of their importance for the model, as they will be studied with the feature importance technique.

It is observed a negative correlation between comfort clothing insulation and indoor temperature (Jiao et al. 2017; Ngarambe et al. 2019; Wang et al. 2019a; Nakagawa and Nakaya 2021; Rupp et al. 2021a; Thapa 2022; Zhuang et al. 2022), as well as with outdoor temperature and outdoor dew point temperature, which has been previously observed in previous studies (De Carli et al. 2007; Jiao et al. 2017; Liu et al. 2018; Ngarambe et al. 2019; Wang et al. 2019a; Ming et al. 2020). Indoor relative humidity and indoor air velocity seem to be slightly correlated with clothing behaviour (Schiavon and Lee 2013; Zhao et al. 2022), existing an inverse proportion with outdoor relative humidity.

Regarding gender, there is not an agreement in the current literature. Some papers point out that gender is not a relevant variable when selection the ensemble (De Carli et al. 2007; Jiao et al. 2017), however, investigations based on a machine learning approach highlight its contribution for predicting clothing insulation (Ngarambe et al. 2019; Duhirwe et al. 2022). In the present work, although the gender is not the most relevant factor, seems to be connected with clothing behaviour.

Outdoor variables are the factors with the strongest association with clothing insulation, higher than indoor parameters (Ngarambe et al. 2019). In particular, parameters related to outdoor temperature at 6 a.m. and dew point temperature at 6 a.m. seem to be the ones that most influence when people deciding their clothes (De Carli et al. 2007; Schiavon and Lee 2013; Ngarambe et al. 2019), followed by air temperature and running mean temperature (de Carvalho et al. 2013; Zhao et al. 2022). In the present study, it is also observed that the variables outdoor running mean temperature at 6 a.m. ($T_{\text{6amrm 0.8}}$) and the running mean dew point ($\text{DWPT}_{\text{6am 0.8}}$) seem to be better predictors of the clothing insulation level than just outdoor temperature at 6 a.m. and dew point temperature at 6 a.m.

Table 3 Relevance of input variables

Category	Variable	Corr. with clothing insulation level		Important feature		Permutation importance	
		MM	FR	MM	FR	MM	FR
General information	Month	−0.21	−0.27	5	2	15	2
	Hour	0.01	−0.10	13	13	6	7
	Seasons	−0.45	−0.18	1	1	1	1
Personal information	Gender	−0.34	−0.25	17	17	21	21
Comfort information	T_{comfort}	−0.30	−0.29	18	19	16	15
Indoor environment	T_{air}	−0.45	−0.61	21	21	19	20
	RH	−0.04	0.19	20	20	18	16
	V_{air}	−0.08	−0.13	9	14	17	17
	T_{g}	0.06	−0.11	12	10	12	8
Outdoor information	$T_{\text{air ext}}$	−0.63	−0.56	10	7	10	13
	RH_{ext}	0.39	0.31	14	12	7	10
	$T_{\text{g ext}}$	−0.63	−0.56	7	8	8	14
	atm_{ext}	0.39	0.22	16	15	11	12
	$T_{\text{rm 0.8}}$	−0.63	−0.60	15	16	14	18
	$\text{RH}_{\text{rm 0.8}}$	0.60	0.48	11	9	13	9
	$T_{\text{out 6 am}}$	−0.65	−0.55	3	5	3	5
	$T_{\text{6amrm 0.8}}$	−0.70	−0.63	19	18	20	19
	DWPT_{6am}	−0.54	−0.43	4	3	4	3
	$\text{DWPT}_{\text{6am0.8}}$	−0.64	−0.57	6	11	5	11
	RH_{6am}	0.35	0.23	2	4	2	4
	$\text{RH}_{\text{6amrm 0.8}}$	0.51	0.40	8	6	9	6

In the process of predicting clothing insulation, it is necessary an analysis to select the appropriate features for the ML models. For that, the feature importance technique was developed for the interpretability of the results. The feature importance assigns a score to each predictor based on its ability to improve predictions (Figure 4), so it is helpful for understanding what drives the model, but it does not show how that characteristic relates to the models' prediction. It is important to conduct this study of characteristics because it provides a measure of how much influences a specific predictor variable (a characteristic) on the accuracy of the models' prediction.

For analysing the relevance of input variables, it is necessary to consider: Which of the features does the model think are most important? And which characteristics have the greatest impact on predictions? These concepts are known

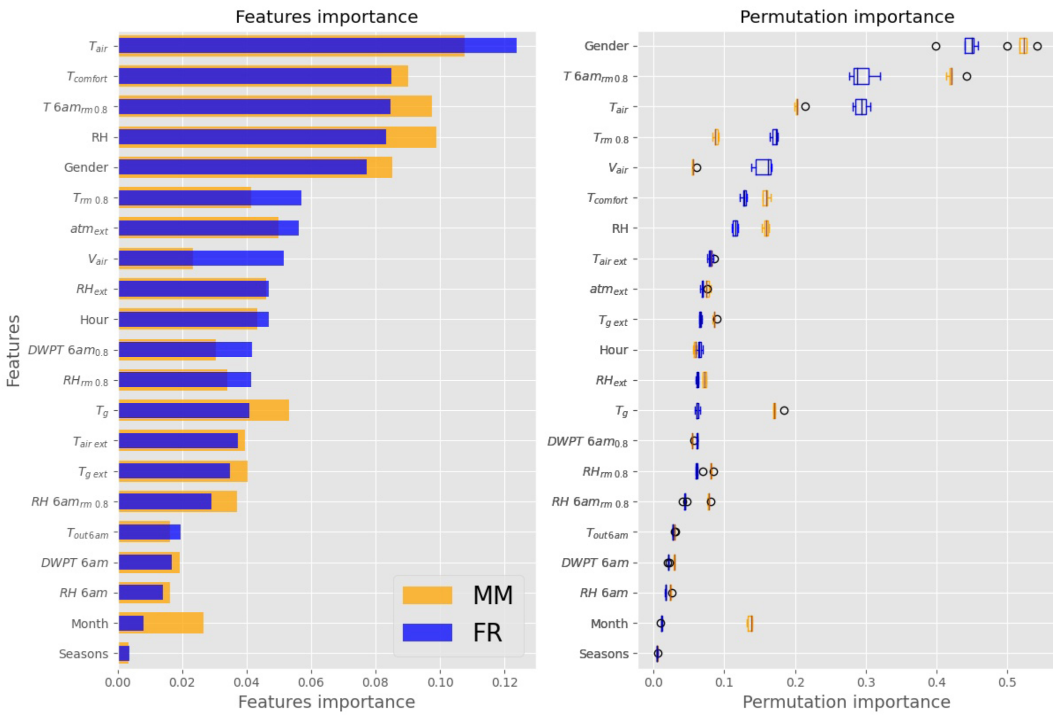


Fig. 4 Feature importance and permutation importance of input variables

as "Important Feature", which shows a relevance rank of the features and "Permutation Importance" (Table 3) and they are used to calculate our input characteristics to the model.

Based on the previous study, which results are shown in Table 3 and Figure 4, it is noted that, in the present work, the most relevant features and with the highest relevance are the running mean temperature at 6 o'clock in the morning, the indoor air temperature, the indoor relative humidity and the comfort temperature, as well as the gender and the barometric pressure. There are no special differences between the results obtained with free-running and mixed-mode data, although it can be noted that the variable "month" is more important and significant in the second one.

It is observed that gender has a significant impact on predictions and the usefulness and impact of the temperature at 6 a.m., which agrees with the work of De Carli et al. (2007). Moreover, it could be highlighted that the indoor air temperature is an outstanding variable for the model.

4.1.1 Principal component analysis (PCA) transformation

In order to further analyse the relevance of input variables and eliminate redundant information, we also considered the PCA transformation. Based on such technique, we selected 9 components from the original 21 variables-dataset, which would represent the 88% of the variability of the problem (Figure 5). It can be seen that, in the first component, all the items have mixed signs involving 21 characteristics, which means that there is a general correlation between the characteristics.

PCA helps to select the components that provide the best value, although it assumes that the data are linearly dependent, and this is not always the case.

In the following sections, it will be analysed which models improve their performance by the application of PCA transformation.

4.2 Comparison of machine learning algorithms for predicting clothing insulation level

A comparison based on the performance of each algorithm was carried out. The results are presented independently, for regression and classification algorithms, considering the whole 21 input variables and applying the PCA transformation (9 input components).

4.2.1 Regression algorithms

Figure 6 depicts the accuracy of the regression models analysed, based on the R2 value and considering the whole 21 input variables. The boxplot shows the mean score of each model, including the minimum score, the first quartile, the median, the third quartile as well as the maximum score.

The ordinary least squares logistic regression (LR)

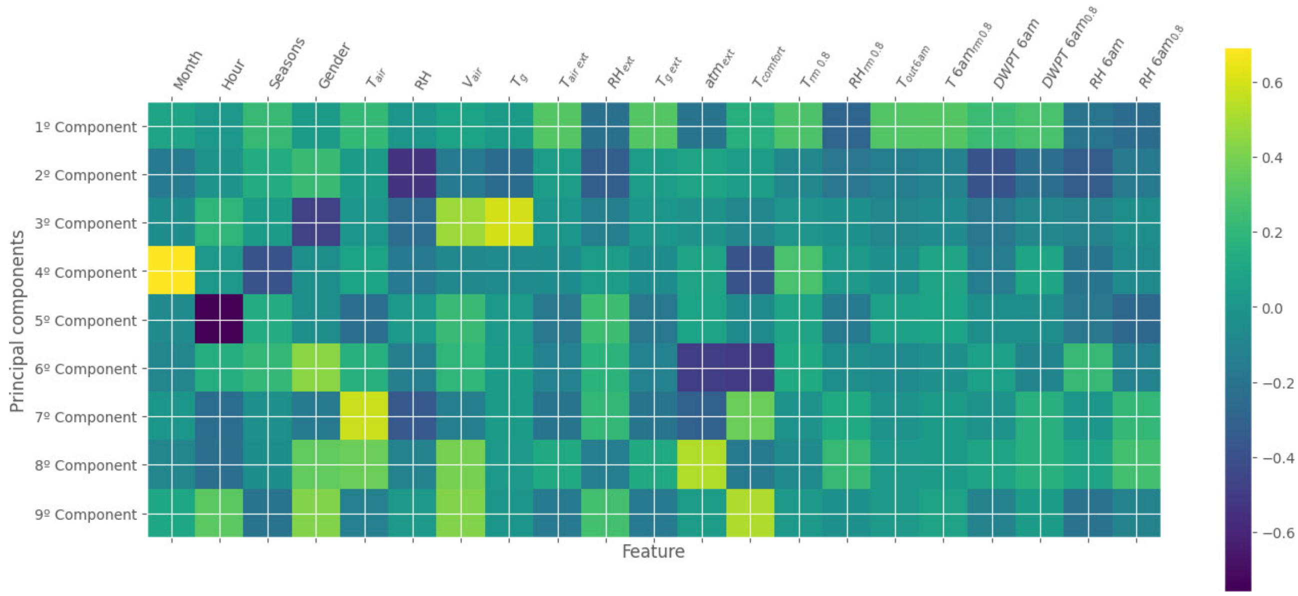


Fig. 5 Feature of each variable in each principal component

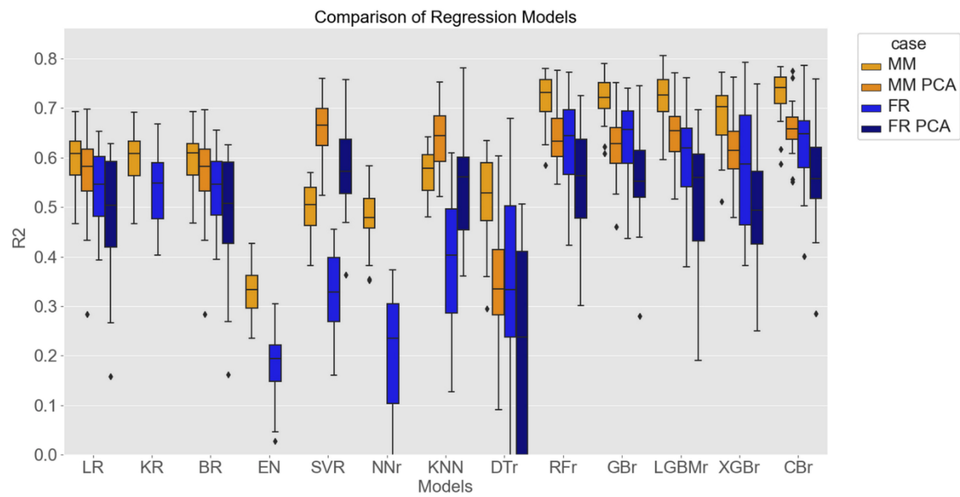


Fig. 6 Comparison of regression models based on the R2 value

presents a good solution, with R2 0.59 MM and R2 0.54 FR, versus other models in the family of linear regression models, as Elastic-Net (EN) (Murphy 2012).

Regarding KR and BR, the weights in the models not only have a predisposition to have the smaller absolute values, but they have a tendency to penalize the extremes of the weights. Both models have similar results to LR, R2 of 0.6 MM and 0.54 FR. So, Ridge is used to analyse any data that suffers from multicollinearity and could be considered as a good regularized of LR method. In the case of EN, its R2 was 0.33 MM and 0.18 FR, lower than LR.

K-nearest neighbour regressor (KNN) model obtained similar results to LR, being the accuracy of this model equal to 0.57 MM and 0.40 FR.

Two models obtained worse solutions than LR and KNN: decision tree (DT) and support vector machine (SVM). The simple decision tree, when there are large number of features with less datasets, may perform worse than linear regression. In general cases, decision trees would achieve better average accuracy, but in the present study, its R2 was 0.51 MM and 0.36 FR. In the case of support vector machine, R2 of 0.50 MM and 0.33 FR, a similar resolution was obtained. The SVM model generally supports, both linear and non-linear solutions, and handles outliers better than LR, but in this case, the model is not better than LR.

The performance of random forest (RFr) (Cutler et al. 2012) is more accurate than decision tree, since a mean of R2 of 0.72 MM and 0.63 FR is observed. Similar solutions are obtained when the results are compared with Gradient-boosted regression (GBr, R2: 0.72 MM, 0,64 FR), Light-gradient machine (LGBMr, R2: 0.72 MM, 0.61 FR), CatBoost (CBr, R2: 0.73 MM, 0.63 FR) or XGBoost (XGBr, R2: 0.68 MM, 0.57 FR). Those models have merged as the most optimized boosting techniques for gradient-boosted tree algorithms (Hancock and Khoshgoftaar 2020).

The last regression model analysed is neural networks (NNr). If this model is compared to LR, LR is focused on linear dependencies, since NN deals with non-linearities. In general, when the model has some nonlinear dependencies, neural networks should perform better than LR, but in this case, the dataset has enough linear relationships and the LR, the RF, as well as the models belonging to the same family, perform all of them better than NN (R2: 0.48 MM, 0.23 FR).

The result of applying Principal Component Analysis is also shown in Figure 6.

After performing the PCA transformation as pre-processing method, it could be noted that some models enhance their performance. This is the case of K-nearest neighbour and support vector machine, but the rest of the regression algorithms did not improve the results.

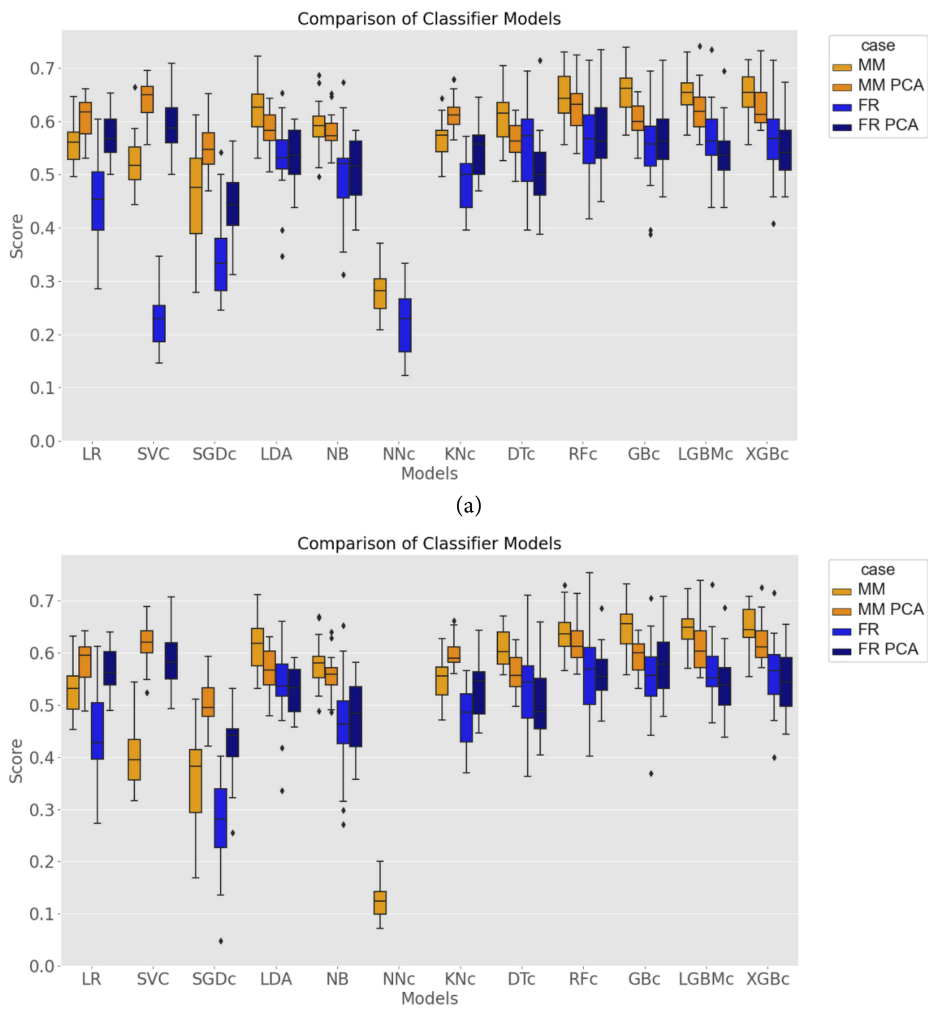
KR and EN were eliminated for the comparison analysis because they are used to determine multicollinearity and the method PCA penalized the resolution.

4.2.2 *Classification algorithms*

Figure 7 illustrates the performance of the classification algorithms for the whole 21 input variables dataset.

The logistic regression (LR) is quite similar to linear regression, basically it is a linear regression with an activation function which would lead to predict one class or another. LR presents a good solution, with accuracy 0.56 MM, 0.46 FR and f1-score of 0.53 MM, 0.45 FR versus other models like support vector machine classifier (accuracy: 0.52 MM, 0.23 FR; f1-score: 0.4 MM, 0.09 FR) and SGDc (accuracy: 0.46 MM, 0.34 FR; f1-score 0.36 MM, 0.28 FR).

The linear discriminant analysis (LDA) model (accuracy: 0.63 MM, 0.53 FR; f1 score: 0.62 MM, 0.53 FR) improved the LR model results. As for the K-nearest neighbour classifier (KNc) algorithm, it is based on classifying a certain



(a)

(b)

Fig. 7 Comparison of classification models based on accuracy (a) and f1-score (b) point looking for its nearest neighbours and it usually achieves good results when enough information is available. Such model obtained an accuracy 0.57 MM and 0.49 FR and f1-score 0.55 MM and 0.48 FR.

Concerning the decision tree classifier (DTc) algorithm, it is feasible to distribute the observations based on their attributes and thus, to predict the outdoor variable with accuracy 0.60 MM, 0.55 FR value and f1-score 0.61 MM, 0.53 FR value. However, it performed worse than random forest classifier (accuracy: 0.65 MM and 0.56 FR and f1-score: 0.64 MM and 0.56 FR), which is an ensemble of decision trees combined with bagging. Random forest classifier usually achieves better results, since each tree is trained with different data and when combining their results, some errors are compensated.

The gradient boosting classifier (GBc) model involves different decision trees, which try to enhance the errors of the previous ones. For this reason, this algorithm leads to an improvement in the solution compared to the previous ones (accuracy: 0.66 MM, 0.55 FR and f1-score: 0.65 MM, 0.55 FR). Likewise, the extreme gradient boosting classifier (XGBc) (accuracy 0.65 MM and 0.56 FR and f1-score 0.65 MM and 0.56 FR) achieved similar results as well as the LGBM classifier (LGBMc) or light GBM classifier with accuracy 0.65 MM, 0.57 FR and f1-score 0.65 MM, 0.57 FR, probably due to its higher training speed and efficiency, and therefore precision (Ke et al. 2017).

It is also possible to perform classification approaches with naive bayes algorithms, which unlike the previous ones, it is based on historical probability to obtain a future probability. Accuracy 0.59 MM and 0.50 FR and f1-score 0.58 MM and 0.46 FR were achieved, considering a normal distribution or a GaussianNB method.

The last algorithm exposed is neural networks classifier (NNc), which accomplished worse results for classification (accuracy: 0.28 MM, 0.22 FR and f1-score: 0.12 MM, 0.09 FR) than for regression.

After performing the PCA technique as pre-processing method, Figure 7 also shows the results of the classification algorithms considered.

It can be noted that some models improved their performance with the PCA transformation, as for the linear regression and support vector machine. SGDc also enhanced its accuracy, as well as K-nearest neighbours (KNc). The rest of the classification algorithms did not improve their predictions.

Based on previous results in the literature, it could be firstly concluded that random forest (RFr and RFc) and other models of the same family, as gradient-boosted (GBr and GBc), light-gradient machine (LGBMr and LBMc), CatBoost (CBr and CBc) and XGBoost (XGBr and XGBc) achieve the best results, for both, regression and classification techniques. Moreover, PCA transformation enhanced the performance of certain algorithms, without detriment of their accuracy.

Comparing the results obtained with previous work, random forest (RFr and RFc) and gradient-boosted seem to perform better that statistical models in order to predict individual comfort clothing level of the occupants. In such kind of comparison, it is noted that different references were obtained from different experiments, dataset and data processing strategies.

Based on the R2 value, linear regression-based algorithms could reach an accuracy in the same order of magnitude as other machine learning algorithms. In Schiavon and Lee (2013), the multivariable linear mixed models predicted 19%–28% of the total variance based on ASHRAE Database. In Su et al. (2022), the multivariable linear mixed model could explained 57.5% of the total variance considering the Chinese Thermal Comfort Database.

For naturally ventilated buildings, in Wang et al. (2019a) the best performance was obtained with power statistical models, achieving a R2 of 0.557. In Nakagawa and Nakaya (2021), it raised up to 0.69 considering weighted regression analysis with binned data, being 0.14 for the raw dataset.

Regarding papers which contemplate mixed-mode buildings, the linear regression model for women and men developed in De Carli et al. (2007) had a R2 equal to 0.25 and 0.39 respectively. In the work carried out by Rupp et al. (2021b), the type of building ventilation was considered for predicting the ensemble insulation based on multiple linear regression analysis with an R2 of 0.22. Taking only into account the data from naturally ventilated situations, it improved up to 0.31.

In Ngarambe et al. (2019), the prediction of indoor clothing insulation levels was made through a deep learning approach using a deep neural network. The model obtained an accuracy of 0.90 considering mean observations, decreasing up to 0.33 for individual values.

5 Conclusions, limitations and future trends

5.1 Conclusions

This paper aims to delve into the study between clothing insulation levels and the variables which it is related to and, also, into the prediction of the individual clothing insulation level in which users would feel thermal comfort with. For that, it centres on an analysis of the relevance of the input for such prediction and on comparing 13 regression machine learning algorithms and 12 classification machine learning algorithms.

The main contributions of this study can be summarized as follows:

- Outdoor temperature at 6 a.m., indoor air temperature, indoor relative humidity, comfort temperature, as well as gender are the most relevant features for predicting the clothing insulation level.
- In particular, the outdoor running mean temperature at 6 a.m. and the running mean dew point variables seem to be better predictors for the clothing ensemble than just outdoor temperature at 6 a.m. and dew point temperature at 6 a.m. Similar to the outdoor running mean temperature variable, the outdoor temperature at 6 a.m. is an exponentially weighted running mean of the daily mean air temperature at 6 a.m., which takes into account the historical values of outdoor temperatures at 6 a.m. Likewise, an improvement of the running mean dew point is based on an exponentially weighted running mean of the daily dew point. Moreover, running mean values have been previously proved to have relationship with clothing changing (Humphreys 1979).
- Principal component analysis (PCA) transformation does not generally enhance the performance of the regression machine learning, nor the classification regression machine learning algorithms for the dataset. Only KNNr and SVMr (regression algorithms) and KNNc, SVMc, LRc, and SGDc (classification algorithms) slightly improved the results.
- Different machine learning models for classification and regression have been compared, since it is often unclear which one would perform better. It has been observed that for regression models, RFr, GBr, LGBMr, CBr and XGBr achieved the best results with a mean of R2 of 0.72 and all of them for MM mode. As for the classification models, RFc, GBc, XGBc, LGBMc performed better than the other algorithms, with a mean accuracy and f1-score of 0.65 (for MM mode).
- The category which comprises decision tree-based algorithms with boosting techniques, or models that improve such techniques, achieve the highest accuracy when predicting the clothing insulation. The feature importance analysis provides insights into the key factors influencing clothing insulation (Figure 4), where decision three based on boosting techniques, might bring about performance improvements and by combining multiple weak models, enhance the overall predictive power. This ensemble approach might be particularly effective for capturing subtle nuances in thermal insulation prediction, and on the other hand, decision trees can naturally handle non-linear relationships in the data.

The study reinforces the use of machine learning algorithms in order to improve the indoor comfort through passive adaptive actions regarding clothing insulation, expanding the actual knowledge based on indoor level clothing insulation considering a new combination of type of buildings and kind of building conditioning. Moreover, since systems are currently being developed that analyse users' clothing and their clothing insulation, e.g., using cameras, machine learning approach would be useful in order to predict an optimal ensemble and to assist the occupants of a building in terms of recommendations regarding their ensemble for the indoor environment, avoiding the overuse of heating, ventilating and air conditioning systems. Recommendations regarding clothing insulation level would be useful, for example, in terms of dress codes for achieving thermal comfort and energy efficiency.

5.2 Limitations and future trends

Analysis of clothing insulation using a thermal manikin is one of the most accurate methods to obtain it, but it is challenging and not usually included in an adaptive thermal comfort field. Some researchers have obtained thermal insulation data for a large number of garments and ensembles by experimenting with thermal manikins and extensive clothing insulation databases have aided in the development of standard insulation values for some typical garments and ensembles in regulations, such as ISO 7730 and ASHRAE 55, which are widely used in comfort studies and field experiments.

Due to the above, and although traditional models obtain good results in a concise way, analysis techniques are currently being developed to obtain clo values in a rapid and easier way, but that are capable of estimating the reality of field studies in buildings, being machine learning algorithm among them.

However, since machine learning is a recent area of study, there are limitations to its real application, which should be studied in the future. Likewise, gaps in research in this area are suggested below, as well as the main recommendations for future studies:

- It would be advisable to carry out studies to search for solutions that simplify or facilitate data collection and the development of ML-based models.
- The training data of this research are based on a field study carried out during a year, which includes winter and summer periods in a Mediterranean climate and it is based on a particular climatology and type of building. Although clo was analysed with various ML algorithms, further studies would be helpful in order to consider different climatic conditions, buildings, as well as conditioning systems and age ranges (elderly and children), which should be further investigated.
- Future studies are recommended focused on the impact of a smaller number of input variables to further investigate variable reduction.

- Moreover, given the rapid and dynamic evolution of techniques in machine learning, there is a need to expand these analyses to encompass additional methods and further explore alternative algorithms in future studies, since there are few works that present analysis of the models based on field studies.
- It would be also interesting to analyse the possibilities of predicting the thermal resistance distinguishing between the upper and lower part of the body, or other segments, for inclusion in new thermal comfort models.

Acknowledgements

The authors wish to acknowledge the financial support of the SICODE project (Ref. US-1380581) funded by the I+D+i FEDER project in Andalusia 2014-2020 and the CONFORES project (Ref. TED2021-130659B-I00) funded by Proyectos de Transición Ecológica y Transición Digital.

Funding note: Funding for open access publishing: Universidad de Sevilla/CBUA.

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

Author contribution statement

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Pablo Aparicio-Ruiz and Elena Barbadilla-Martín. The first draft of the manuscript was written by all authors and all authors commented on previous versions of the manuscript and approved the final manuscript.

References

Aït-Sahalia Y, Xiu D (2019). Principal component analysis of high-frequency data. *Journal of the American Statistical Association*, 114: 287–303.

ANSI/ASHRAE (2020). ANSI/ASHRAE Standard 55-2020: Thermal Environmental Conditions for Human Occupancy. Atlanta, GA, USA: American Society of Heating, Refrigerating and Air-Conditioning Engineers.

Antweiler D, Harmening M, Marheineke N, et al. (2022). Machine learning framework to predict nonwoven material properties from fiber graph representations. *Software Impacts*, 14: 100423.

Arens E, Humphreys MA, de Dear R, et al. (2010). Are 'class A' temperature requirements realistic or desirable? *Building and Environment*, 45: 4–10.

Barbadilla-Martín E, Salmerón Lissén JM, Guadix Martín J, et al. (2017). Field study on adaptive thermal comfort in mixed mode office buildings in southwestern area of Spain. *Building and Environment*, 123: 163–175.

Bishop CM (2006). Pattern Recognition and Machine Learning. New York: Springer.

Breiman L (2001). Random forests. *Machine Learning*, 45: 5–32.

Chai Q, Wang H, Zhai Y, et al. (2020). Using machine learning algorithms to predict occupants' thermal comfort in naturally ventilated residential buildings. *Energy and Buildings*, 217: 109937.

Chen T, Guestrin C (2016). XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA.

Choi H, Na H, Kim T, et al. (2021). Vision-based estimation of clothing insulation for building control: A case study of residential buildings. *Building and Environment*, 202: 108036.

Cui J, Watanabe T (2014). A study on air-conditioning load calculation method considering thermal sensation of residents—Quantitative evaluation on the load reduction effect of WARM BIZ and COOL BIZ. *Journal of Environmental Engineering (Transactions of AIJ)*, 9(695): 73–81.

Cutler A, Cutler DR, Stevens JR (2012). Random forests. In: Zhang C, Ma Y (eds), Ensemble Machine Learning. New York: Springer. pp. 157–175.

De Carli M, Olesen BW, Zarrella A, et al. (2007). People's clothing behaviour according to external weather and indoor environment. *Building and Environment*, 42: 3965–3973.

de Carvalho PM, da Silva MG, Ramos JE (2013). Influence of weather and indoor climate on clothing of occupants in naturally ventilated school buildings. *Building and Environment*, 59: 38–46.

de Dear RJ, Brager GS (1998). Developing an adaptive model of thermal comfort and preference. *ASHRAE Transactions*, 104(1): 145–167.

Deng Z, Chen Q (2018). Artificial neural network models using thermal sensations and occupants' behavior for predicting thermal comfort. *Energy and Buildings*, 174: 587–602.

Duhirwe PN, Ngarambe J, Yun GY (2022). Geographically extended occupant clothing behavior model using convolutional neural networks with hyperband technique. *Journal of Building Engineering*, 49: 104023.

EN (2020). EN 16798-1: 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. European Committee for Standardization (CEN).

Fan C, Chen M, Wang X, et al. (2021). A review on data preprocessing techniques toward efficient and reliable knowledge discovery from building operational data. *Frontiers in Energy Research*, 9: 652801.

Friedman JH (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29: 1189–1232.

Gagge AP, Burton AC, Bazett HC (1941). A practical system of units for the description of the heat exchange of man with his environment. *Science*, 94: 428–430.

Gholamy A, Kreinovich V, Kosheleva O (2018). Why 70/30 or 80/20 relation between training and testing sets: A pedagogical explanation. Departmental Technical Reports (CS), The University of Texas At El Paso.

Hancock JT, Khoshgoftaar TM (2020). CatBoost for big data: an interdisciplinary review. *Journal of Big Data*, 7: 94.

Hastie T, Tibshirani R, Friedman JH (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. New York: Springer.

Humphreys MA (1970). A simple theoretical derivation of thermal comfort conditions. *JIHVE*, 38: 95–98.

Humphreys MA (1979). The influence of season and ambient temperature on human clothing behaviour. In: Fanger PO, Valbjørn O (eds), Indoor Climate. pp 699–713.

Humphreys MA, Rijal HB, Nicol JF (2013). Updating the adaptive relation between climate and comfort indoors: New insights and an extended database. *Building and Environment*, 63: 40–55.

ISO (2002). ISO 7726: Ergonomics of the thermal environment—Instruments for measuring physical quantities.

ISO (2006). 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.

ISO (2009a). ISO 9920: Ergonomics of the thermal environment—Estimation of thermal insulation and water vapour resistance of a clothing ensemble.

ISO (2009b). ISO 11079: Ergonomics of the thermal environment—Determination and interpretation of cold stress when using required clothing insulation (IREQ) and local cooling effects.

James G, Witten D, Hastie T, et al. (2023). An Introduction to Statistical Learning with Applications in Python. Cham: Springer.

Jiao Y, Yu H, Wang T, et al. (2017). The relationship between thermal environments and clothing insulation for elderly individuals in Shanghai, China. *Journal of Thermal Biology*, 70: 28–36.

Joshi A, Psikuta A, Bueno MA, et al. (2019). Analytical clothing model for sensible heat transfer considering spatial heterogeneity. *International Journal of Thermal Sciences*, 145: 105949.

Joshi A, Wang F, Kang Z, et al. (2022). A three-dimensional thermoregulatory model for predicting human thermophysiological responses in various thermal environments. *Building and Environment*, 207: 108506.

Ke G, Meng Q, Finley T, et al. (2017). LightGBM: A highly efficient gradient boosting decision tree. In: Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA.

Kottek M, Grieser J, Beck C, et al. (2006). World map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, 15: 259–263.

Kuhn M, Johnson K (2019). Feature Engineering and Selection: A Practical Approach for Predictive Models. Boca Raton, FL, USA: Chapman and Hall/CRC.

Li X, Chen Q (2021). Development of a novel method to detect clothing level and facial skin temperature for controlling HVAC systems. *Energy and Buildings*, 239: 110859.

Liu W, Yang D, Shen X, et al. (2018). Indoor clothing insulation and thermal history: A clothing model based on logistic function and running mean outdoor temperature. *Building and Environment*, 135: 142–152.

Liu J, Foged IW, Moeslund TB (2022a). Automatic estimation of clothing insulation rate and metabolic rate for dynamic thermal comfort assessment. *Pattern Analysis and Applications*, 25: 619–634.

Liu T, Jin L, Zhong C, et al. (2022b). Study of thermal sensation prediction model based on support vector classification (SVC) algorithm with data preprocessing. *Journal of Building Engineering*, 48: 103919.

Luo M, Xie J, Yan Y, et al. (2020). Comparing machine learning algorithms in predicting thermal sensation using ASHRAE Comfort Database II. *Energy and Buildings*, 210: 109776.

Medina A, Méndez JI, Ponce P, et al. (2022). Embedded real-time clothing classifier using one-stage methods for saving energy in thermostats. *Energies*, 15: 6117.

Ming R, Yu W, Zhao X, et al. (2020). Assessing energy saving potentials of office buildings based on adaptive thermal comfort using a tracking-based method. *Energy and Buildings*, 208: 109611.

Morgan C, de Dear R (2003). Weather, clothing and thermal adaptation to indoor climate. *Climate Research*, 24: 267–284.

Murphy KP (2012). Machine Learning: A Probabilistic Perspective. Cambridge, MA, USA: MIT Press.

Mutlu G, Acı Çİ (2022). SVM-SMO-SGD: A hybrid-parallel support vector machine algorithm using sequential minimal optimization with stochastic gradient descent. *Parallel Computing*, 113: 102955.

Nakagawa A, Nakaya T (2021). A survey of clothing insulation for university students in Japan: Effect of clothing insulation distribution between the upper and lower body in the winter. *Journal of Building Engineering*, 44: 103287.

Ngarambe J, Yun G, Kim G (2019). Prediction of indoor clothing insulation levels: A deep learning approach. *Energy and Buildings*, 202: 109402.

Nicol F, Humphreys M (2010). Derivation of the adaptive equations for thermal comfort in free-running buildings in European standard EN15251. *Building and Environment*, 45: 11–17.

Nicol F, Roaf S (1996). Pioneering new indoor temperature standards: The Pakistan Project. *Energy and Buildings*, 23: 169–174.

Prokhorenkova L, Gusev G, Vorobev A, et al. (2018). CatBoost: Unbiased boosting with categorical features. In: Proceedings of the 32nd International Conference on Neural Information Processing Systems, Montréal, Canada.

Rocha M, Cortez P, Neves J (2007). Evolution of neural networks for classification and regression. *Neurocomputing*, 70: 2809–2816.

Rupp RF, Andersen RK, Toftum J, et al. (2021a). Occupant behaviour in mixed-mode office buildings in a subtropical climate: Beyond typical models of adaptive actions. *Building and Environment*, 190: 107541.

Rupp RF, Kazanci OB, Toftum J (2021b). Investigating current trends in clothing insulation using a global thermal comfort database. *Energy and Buildings*, 252: 111431.

Schiavon S, Lee KH (2013). Dynamic predictive clothing insulation models based on outdoor air and indoor operative temperatures. *Building and Environment*, 59: 250–260.

Sokolova M, Japkowicz N, Szpakowicz S (2006). Beyond accuracy, F-score and ROC: A family of discriminant measures for performance evaluation. In: Sattar A, Kang Bh (eds), AI 2006: Advances in Artificial Intelligence. Heidelberg: Springer Berlin Heidelberg. pp. 1015–1021.

Song Y, Liang J, Lu J, et al. (2017). An efficient instance selection algorithm for k nearest neighbor regression. *Neurocomputing*, 251: 26–34.

Su X, Wang Z, Zhou F, et al. (2022). Comfortable clothing model of occupants and thermal adaption to cold climates in China. *Building and Environment*, 207: 108499.

Tang R, Fan C, Zeng F, et al. (2022). Data-driven model predictive control for power demand management and fast demand response of commercial buildings using support vector regression. *Building Simulation*, 15: 317–331.

Thai HT (2022). Machine learning for structural engineering: A state-of-the-art review. *Structures*, 38: 448–491.

Thapa S (2022). Investigation of thermal comfort and adaptation among the residents of cold climate in the lower Himalayan region of eastern India. *Indoor and Built Environment*, 31: 1613–1635.

Vapnik VN (1999). The Nature of Statistical Learning Theory. New York: Springer.

Wang L, Kim J, Xiong J, et al. (2019a). Optimal clothing insulation in naturally ventilated buildings. *Building and Environment*, 154: 200–210.

Wang Z, Yu H, Luo M, et al. (2019b). Predicting older people's thermal sensation in building environment through a machine learning approach: Modelling, interpretation, and application. *Building and Environment*, 161: 106231.

Wang Z, Wang J, He Y, et al. (2020). Dimension analysis of subjective thermal comfort metrics based on ASHRAE Global Thermal Comfort Database using machine learning. *Journal of Building Engineering*, 29: 101120.

Yu C, Li B, Wu Y, et al. (2022). Performances of machine learning algorithms for individual thermal comfort prediction based on data from professional and practical settings. *Journal of Building Engineering*, 61: 105278.

Zhang S, Li X, Zong M, et al. (2017). Learning *k* for kNN classification. *ACM Transactions on Intelligent Systems and Technology*, 8: 43.

Zang M, Xing Z, Tan Y (2019). IoT-based personal thermal comfort control for livable environment. *International Journal of Distributed Sensor Networks*, 15(7): 155014771986550.

Zhao W, Chow D, Yan H, et al. (2022). Influential factors and predictive models of indoor clothing insulation of rural residents: A case study in China's cold climate zone. *Building and Environment*, 216: 109014.

Zhuang L, Huang J, Li F, et al. (2022). Psychological adaptation to thermal environments and its effects on thermal sensation. *Physiology & Behavior*, 247: 113724.