Predicting Older People’s Thermal Sensation by a New Integrated Physiological-based and Data-driven Model

Lili Ji, Chang Shu, Abdelaziz Laouadi

Liangzhu (Leon) Wang, Michael Lacasse

ABSTRACT

Due to age-related physiological changes, older people are more vulnerable than young people to heat or cold conditions. Predicting older people’s thermal sensations is essential for controlling the built environment and avoiding extreme heat/cold injuries. Previous studies mainly focused on predicting the thermal sensation of young people, and the data-driven methods are often not constrained by physiological responses. This study proposes a new integrated model to combine the two-node physiological model and the data-driven method random forest classifier. The surveyed data of older people come from ASHRAE Global Thermal Comfort Database II. The dataset has collected the environmental conditions, subjects’ factors, and survey results of thermal sensation vote (TSV). In this study, with the environmental conditions (air temperature, mean radiant temperature, relative humidity, and airspeed) and subject factors (clothing insulation, height, and weight) as inputs, core and skin temperatures, water loss, and standard effective temperature (SET) can be calculated by the two-node model of older people. The above physiological parameters and building operation mode (naturally-ventilated/air-conditioned - NV/AC), older people’s gender, surveyed seasons, and climate zones are used to train the data-driven model. The results show that the overall accuracy classification score of the integrated model is 90%, which is more accurate than the PMV model and the majority of other data-driven studies. The integrated model can also reach above 80% accuracy classification score under different building operation modes (NV/AC), older people’s gender, surveyed seasons, and climate zones. The correlation between SET and TSV is better than the traditional linear regression. It is found that there is a possibility that older people’s core temperature increase to a dangerous level (>38°C) even when they just feel slightly warm.

Lili Ji and Chang Shu are PhD students in the Department of Building, Civil and Environmental Engineering, Concordia University, Montreal, Canada. Liangzhu (Leon) Wang is a professor in the Department of Building, Civil and Environmental Engineering, Concordia University, Montreal, Canada. Abdelaziz Laouadi and Michael Lacasse are senior research officers at the Construction Research Centre of the, National Research Council Canada, Ottawa, Canada.
INTRODUCTION

The world population aging process has been escalating; the number of people aged 60 years and over was tripled in the last 50 years and expected to reach over 2.1 billion in the next thirty years (Issahaku & Neysmith, 2013; Mba, 2010). Meanwhile, the projected climate change and global warming in the twenty-first century would intensify the exposure to deadly ambient conditions (Mora et al., 2017). Older people are more vulnerable than young people to heat or cold conditions due to age-related physiological changes (Rida et al., 2014). Older people's thermal sensation is also different from young people, especially under hot/cold exposures (Schellen et al., 2010; Soebarto et al., 2019). It is essential to predict older people’s thermal sensations under different situations and further link it to their physiological responses to investigating the relationship between physical and subjective factors.

Thermal sensations affected by physical and environmental factors have been explored for the past decades. The predicted mean vote (PMV) model built by (Fanger, 1970) is based on human body heat balance and chamber experiments, which has been widely used but its accuracy remains a contested topic (Zhou et al., 2020). In recent years, many studies have tried to apply data-driven methods to predict occupants' thermal sensation vote (TSV) and showed good prediction accuracy between 80% to 90% (Zhou et al., 2020). Different data-driven methods, including Decision Tree (DT), Support Vector Machine (SVM), Artificial Neural Network (ANN), Random Forest (RF), Adaboost (Ab), and Gradient Boosting Machine (GBM) have been investigated for thermal sensation predictions (Chai et al., 2020; Katić et al., 2020; Luo et al., 2020). However, these data-driven methods lack the constraints of physiological parameters, raising concerns over applying those models to broader situations.

Therefore, this study aims to propose a method to integrate the physiological model with the data-driven method to predict older people's thermal sensation under various situations and further investigate the relationship between older people's physiological responses and their thermal sensation to evaluate their thermal risk under cold/hot exposures.

METHODOLOGY

Figure 1 shows the framework of the proposed method. With the inputs of air temperature ($T_a$), relative humidity ($RH$), air velocity ($V_a$), metabolic rate ($MET$), and clothing insulation ($CLO$), the physiological model can calculate people's physiological responses, including skin temperature ($T_{sk}$), core temperature ($T_{cr}$) and body water loss. After that, the data-driven model is used to establish their relationship with thermal sensation. In this step, additional parameters, including climate zones, seasons, gender, outdoor temperature, and building operations strategies, are also inputs of the data-driven model. In the integrated model, the data-driven model is constrained by physiological parameters.

![Figure 1 Framework of the proposed method.](image-url)
Data source and pre-processing

The thermal sensation data of older people (age>60) are from ASHRAE Global Thermal Comfort Database II (Földváry Ličina et al., 2018), which including 107,584 samples of thermal comfort data, among them there are 1566 for older people aged above 60 years old. The seven-point scale $TSV$ (3: hot; 2: warm; 1: slightly warm; 0: neutral; -1: slightly cool; -2: cool; -3: cold) (ASHRAE-55, 2017) is chosen as the targeted comfort index. Eleven variables, including the $T_a$, RH, $V_a$, $MET$, $CLO$, Age, Gender, Climate zones (Koppen climate classification (Kottek et al., 2006)), Seasons, Outdoor monthly temperature, Building operation strategy, are input features. In total, 1413 samples with all the above variables were qualified. This sample size is comparable with previous researches on the data-driven thermal comfort model. Wang et al. (2019) adopted 1040 samples of older people and split them as 80% training data and 20% testing data. Megri and El Naqa (2016) used 793 samples for training and 18 samples for testing. As shown in Table 1, the unit and data distribution vary with input features. There are numeric numbers including $T_a$, RH, $V_a$, $MET$, $CLO$, Age, and categorical variables including Climate zones, Seasons, Gender, and Building operation strategies.

It is essential to do the pre-coding to transform the formats of categorical data into numerical data before analyzing their impact on $TSV$ prediction. Label encoding can convert categorical data into ordinal data, which does not add any new columns to the data but implies an order to the variable (Potdar et al., 2017). However, for our categorical variables, there is no quantitative relationship between the individual values. For example, there is no order between different climate zones. In this situation, using label encoding can potentially create a fictional ordinal relationship in the data. One-hot encoding is another way to deal with categorical data (Rodríguez et al., 2018). A one-hot is a group of bits among which the legal combinations of values are only those with a single high (1) bit and all the others low (0). For example, if one sample has a value of 'Autumn', one-hot encoding will set its 'Season' column as '1' and other columns as '0'. They will be fixed in a matrix vector and transformed into an array. Therefore, we apply one-hot encoding to deal with our nominal categorical variables, to improve the algorithm performance.

The data are split into training data and test data. The 80% training proportion rule is favored by many researchers (Luo et al., 2020). The training data (80%) are used to train our data-driven model, and the test data (20%) are used to validate the model.

<table>
<thead>
<tr>
<th>Category</th>
<th>Variables</th>
<th>Unit</th>
<th>Range</th>
<th>Number of samples</th>
<th>Number of missing samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target thermal comfort index</td>
<td>Thermal sensation</td>
<td>7-point scale units</td>
<td>1527</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>Indoor environment</td>
<td>$T_a$</td>
<td>°C</td>
<td>13-41</td>
<td>1474</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>RH</td>
<td>%</td>
<td>10.4-86.3</td>
<td>1169</td>
<td>397</td>
</tr>
<tr>
<td></td>
<td>$V_a$</td>
<td>m/s</td>
<td>0-2.1</td>
<td>1073</td>
<td>493</td>
</tr>
<tr>
<td></td>
<td>$MET$</td>
<td>met</td>
<td>0.7-5</td>
<td>1143</td>
<td>423</td>
</tr>
<tr>
<td>Occupant factors</td>
<td>$CLO$</td>
<td>clo</td>
<td>0-2.87</td>
<td>1532</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>Female, Male</td>
<td>1557</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>60-99</td>
<td>1566</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Outdoor environment</td>
<td>Monthly Tout</td>
<td>°C</td>
<td>(-15)-39.9</td>
<td>1342</td>
<td>224</td>
</tr>
<tr>
<td></td>
<td>Climate zones</td>
<td>Koppen climate classification</td>
<td>1566</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Seasons</td>
<td>Spring, Summer, Autumn, Winter</td>
<td>1504</td>
<td>62</td>
<td></td>
</tr>
<tr>
<td>Building information</td>
<td>Operation strategy</td>
<td>AC, NV, mix mode</td>
<td>1531</td>
<td>35</td>
<td></td>
</tr>
</tbody>
</table>

A physiological model of older people

The physiological model of older people is developed based on the two-node model (Ji et al., 2021). The two-node model treats the human body as two concentric cylinders for the core and skin layers. The core and skin layers are represented by one node each. A uniform layer of clothing covers the skin layer throughout the body. Metabolic heat is generated at the core layer. A small portion of that heat is dissipated through respiration by convection and evaporation, and the remainder is transported by conduction and skin blood flow to the skin surface. The heat exchange between the skin surface and the environment is
divided into two parts: 1) the sensible heat by conduction, radiation, and convection from the skin surface to the clothing layer and then to the environment; 2) the insensible heat by the evaporation of sweat and moisture diffusion from the skin surface. The heat balance equations for the core and skin layers are given below.

\[ m_{cr}c_{cr} \frac{dT_{cr}}{dt} / A_b = M_{cr} + SHIV - Work - Q_{res} - h_{sk}(T_{cr} - T_{sk}) \quad (1) \]

\[ m_{sk}c_{sk} \frac{dT_{sk}}{dt} / A_b = M_{sk} + h_{sk}(T_{cr} - T_{sk}) - Dry - Evap \quad (2) \]

Where \( m_{cr} \) and \( c_{sk} \) are the mass (kg) of the core and skin nodes, respectively; \( c_{cr} \) and \( c_{sk} \) are the specific thermal capacity (W/kg°C) of the core and skin nodes, respectively; \( T_{cr} \) and \( T_{sk} \) are the core and skin node temperatures (°C), respectively; \( dt \) is the time step (1 min); \( A_b \) is the Dubois body surface area (m²); \( M_{cr} \) is the metabolic rate of the core node (W/m²); \( M_{sk} \) is the metabolic rate of the skin node (W/m²); \( SHIV \) is the shivering metabolic rate (W/m²); Work is the mechanical work done by the body (W/m²); \( Q_{res} \) is the heat loss through respiration (W/m²); \( h_{sk} \) is the skin thermal conductance that accounts for the blood flow (W/m²°C); \( Dry \) and \( Evap \) are the sensible and evaporative heat exchanges of the skin node (W/m²), respectively. The steady-state of heat balance can be achieved after 1 h exposure.

Under heat or cold conditions, the deviation of the \( T_{cr} \), \( T_{sk} \) or \( T_b \) (body temperature) from their threshold values (\( T_{cr0}, T_{sk0} \) or \( T_{b0} \)) would be \( SET \) as the thermoregulatory control signals. The warm signal is given by \( \Delta T_w = (T - T_b) + \) while the cold signal is given by \( \Delta T_c = (T_b - T) + \), where (+) means the only positive value will be taken. Those signals would trigger the regulatory sweating, vasodilation, and vasoconstriction, and shivering. As the ambient conditions become hotter and/or more humid with an increased activity level, the human body tends to depend on sweat evaporation from the skin surface to cool and maintain its core temperature. The sweating rate may be expressed as a function of the body and skin temperature control signals, as shown in Equation 3.

\[ SWR = CSWE \times CSW \times (\Delta T_{b,sw} + Acof \cdot \Delta T_{sk,sw}) \times \exp \left( \frac{\Delta T_{sk,sw}}{10.7} \right) \quad (3) \]

where \( CSWE \) is the sweating attenuation coefficient of older people, \( Acof \) and \( CSW \) are model constants, \( SWR \) is the sweating rate (g/m²h), \( \Delta T_{b,sw} \) is the body temperature control signal for sweating (°C), \( \Delta T_{sk,sw} \) is the skin temperature control signal for sweating (°C).

Skin blood flow rate (\( SBF \)) depends on the body’s thermal state and varies between the minimum and maximum values. Under heat-stressful conditions, skin blood flow is increased by vasodilation. Under cold conditions, skin blood flow is, however, controlled by vasoconstriction. The skin blood flow rate is expressed by the following formulation:

\[ SBF = [SBF_{basal} + CDE \times CDIL \times \Delta T_{cr,ctl}] / [1 + CSTR \times CSTR \times \Delta T_{sk,cons}] \quad (4) \]

where \( SBF \) is the skin blood flow rate (L/m²/hr), \( SBF_{basal} \) is the basal (neutral) skin blood flow rate (L/m²/hr), \( CDIL \) and \( CSTR \) are model constants, \( CDE \) and \( CCE \) are the vasodilation attenuation coefficient and Vasoconstriction attenuation coefficient of older people, respectively, \( \Delta T_{cr,ctl} \) is the core temperature control signal for vasodilation (°C), \( \Delta T_{sk,cons} \) is the skin temperature control signal for vasoconstriction (°C).

As body cooling progresses, when the non-shivering thermogenesis during vasoconstriction is not enough to maintain the body heat balance, the second line of defense is shivering. Shivering is the random involuntary contraction of superficial muscle fibers, which increases heat production. The original two-node model adopted the shivering rate model developed by Stolwijk et al. (1971). Shivering is triggered by a multiplicative error signal, which means shivering starts until both skin and core vasoconstriction thresholds are exceeded. However, shivering might be triggered by additive control signals, which means shivering starts until the skin or core vasoconstriction threshold is exceeded. With this assumption, the shivering metabolic rate is expressed by the following Equation:
\[ M_{shiv} = CSHE \times 19.4 \times \Delta T_{sk,cons} + Cof_{sc} \times \Delta T_{cr,shiv} + Cof_{ss} \times \Delta T_{sk,cons} \] (5)

Where \( CSHE \) is the shivering attenuation coefficient of older people, \( \Delta T_{cr} \) is the core temperature control signal for shivering (°C), \( Cof_{sc} \) and \( Cof_{ss} \) are model constants.

Table 2 lists the threshold values of thermoregulatory functions and the attenuation coefficients in the older people's model. The maximum sweating rate changes, minimum and maximum skin blood flow rates are also listed.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Older</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vasodilation threshold ( T_{cr0,di} ) (°C)</td>
<td>37.29</td>
</tr>
<tr>
<td>Vasoconstriction threshold ( T_{sk0,cons} ) (°C)</td>
<td>33.20</td>
</tr>
<tr>
<td>Sweating threshold ( T_{cr0,sw} ) (°C)</td>
<td>36.90</td>
</tr>
<tr>
<td>Sweating threshold ( T_{sk0,sw} ) (°C)</td>
<td>33.90</td>
</tr>
<tr>
<td>Shivering threshold ( T_{cr0,sh} ) (°C)</td>
<td>36.75</td>
</tr>
<tr>
<td>Vasodilation attenuation coefficient CDE</td>
<td>0.75</td>
</tr>
<tr>
<td>Vasoconstriction attenuation coefficient CCE</td>
<td>0.50</td>
</tr>
<tr>
<td>Sweating attenuation coefficient CSWE</td>
<td>0.75</td>
</tr>
<tr>
<td>Shivering attenuation coefficient CSHE</td>
<td>1</td>
</tr>
<tr>
<td>Min SBF rate (L/h/m²)</td>
<td>0.75</td>
</tr>
<tr>
<td>Max SBF rate (L/h/m²)</td>
<td>63</td>
</tr>
<tr>
<td>Max sweating rate factor</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Data-driven model and model evaluation

As the target output is the seven-point scale \( TSV \), which can be regarded as a classification problem. So we selected the Random forest classifier as the data-driven method in this study. A random forest is a meta estimator that fits a number of decision tree classifiers on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting. The complexity and size of the trees are controlled by tree numbers and tree depth. We adopt 'Scikit-learn Ensemble' package in python. To optimize the model performance, the number of trees and the depth of trees were tuned by grid-search in the range. The model can achieve maximum global accuracy with 180 trees and 18 tree depths.

For model evaluation, we utilized two indices, the Accuracy Classification Score (\( ACS \)) and Mean Absolute Error (\( MAE \)). \( ACS \) is the ratio of the number of correct predictions to the total number of input samples, as shown in Equation 1. \( MAE \) is the average of the difference between the original values and the predicted values. It gives us the measure of how far the predictions were from the actual output. Generally, the higher the \( ACS \) and the lower \( MAE \), the better the predicted \( TSV \)s are. They are calculated as follows:

\[
ACS = \frac{N}{N} \quad \text{(6)}
\]

\[
MAE = \frac{1}{N} \sum_{j=1}^{N} |y_j - y'_j| \quad \text{(7)}
\]

Where \( N \) is the number of correct prediction (predicted \( TSV = \) actual \( TSV \)), \( N \) is the total number of predictions made, \( y_j \) and \( y'_j \) are the actual \( TSV \) and predicted \( TSV \) values, respectively.

RESULTS

Overall prediction results

Figure 2a shows how \( TSV \) and its predicted value vary with \( T_{aw} \). The predicted \( TSV \) is also compared with the \( PMV \) to
evaluate its performance compared to the PMV model. The shades are the range of TSV and predicted TSV at each T_{air}. The bars at the bottom stand for sample sizes at every 1 °C temperature bin. The ACS is 90.2% for predicted TSV, higher than the 34.7% for PMV, which is increased by 55.4%. Similarly, MAE is 0.08 for predicted TSV, much lower than 0.85 for PMV. The comparison of these two indices indicates that the predicted TSV from the proposed integrated model has a smaller error and fits much better with actual TSV than the PMV model. Besides, under hot and cold exposures, the deviations from actual TSV and PMV values increase. When the temperature is higher than 30 °C, PMV begins to deviate from actual TSV. The maximum deviation increased to 3.63 at 37 °C, while the maximum deviation between predicted and actual TSV values is 0.67 at 38°C. Therefore, compared with the PMV model that matches with actual TSV only in a narrow neutral temperature range, the predicted TSV fit well with actual TSV, which is because while PMV model is based on steady-state human heat balance, the two-node model within the proposed model includes the thermoregulatory actions under hot/cold exposure. The data-driven model considers more factors beyond the physiological responses.

Figure 2 TSV, predicted TSV, and PMV varied with (a) Indoor air temperature and (b) older people’s SET. The shaded areas are the ranges of TSV and predicted TSV at each T_{air}.

Figure 2b shows how TSV and predicted TSV vary with the Standard Effective Temperature (SET). SET is the thermal comfort index and a heat/cold stress screening tool. The traditional widely used correlation between SET and TSV is the linear regression from McIntyre (1980), as shown in Equation 8. According to the data-driven regression and linear regression, the SET at neutral condition (TSV=0) is about 24°C. However, with the variation of SET, the predicted TSV by the linear regression deviates from the actual TSV rapidly, while the predicted TSV with the proposed integrated model matches well with the actual TSV.

\[ TSV = 0.25SET - 6.03, R^2 = 0.998 \]  

Model evaluation under various situations

Many studies have reported the significant differences between thermal comfort in naturally ventilated (NV) and air-conditioned (AC) buildings. Figure 3a shows the results of the NV situation. In the NV situation, the ACS of the integrated model is 85.9%, which is much higher than the PMV model of 33.1%. The MAE of the integrated model is 0.13, while the MAE of the PMV model is 0.94. Figure 3b shows the AC situation. The ACS of the integrated model is 82.2%, which is much higher than the PMV model of 34.9%. The MAE of our integrated model is 0.13, while the MAE of the PMV model is 0.72. Based on the above two, the integrated model can achieve better performance than the PMV model under both AC and NV situations.
Figure 3 shows the comparison of TSV prediction results for males and females. For males, the prediction ACS is 92.8%, and MAE is 0.06. For females, the prediction ACS is 88.3%, and MAE is 0.11. When the indoor air temperature is between 20°C and 30°C, the TSV of males and females are comparable. Under the hot exposure when the indoor temperature is above 30°C, males tend to feel hotter than females; under the cold exposure, when the indoor air temperature is below 17°C, females tend to feel cooler than males.

Figure 4 shows the comparison of TSV prediction for male and female. Besides building operation mode and gender, the model is also evaluated under different climate zones and seasons, as shown in Figure 5. According to the surveyed cities provided by Database II, ASHRAE climate zones from extremely hot to very cold are classified. Figure 5 shows that the ACS of the proposed integrated model is above 80% in various situations while the ACS of the PMV model is lower than 50%. The MAE of the proposed model is also much less than the MAE of the PMV model under various situations. This shows that the integrated model could be applied to various cases and broader situations.
Figure 5 Comparison of the (a) ACS and (b) MAE of the integrated model and PMV model under different situation

**Relationship between physiological parameters and TSV**

To further evaluate older people's thermal risk under cold/hot exposures, we investigated the relationship between older people's physiological responses and their thermal sensation. Figure 6 shows the core temperature \( T_c \) and skin temperature \( T_s \) vary with TSV. Within the TSV range -3 to 3, the core temperature of older people vary from 36.5°C to 38.1°C. According to ISO 7933, a core temperature exceeding 38.0°C is defined as dangerous. It should be noted that the maximum core temperature (38.1°C) happened when TSV is 1 (slightly warm), which indicates that when older people do not feel hot, their core temperature may be already increased to a dangerous level. This result is consistent with the chamber experiment of Tsuzuki and Iwata (2002); older subjects felt slightly warm or warm when conducting light exercise under the environment of \( T_{air} = 27°C \), while their core temperature varied from around 37.5°C to 38.0°C. Within the TSV range of -3 to 3, the skin temperature of older people varies from 30.2°C to 37.9°C. Even under warm to hot exposure (TSV >0), older people tend to have skin temperature under the normal skin temperature of 33.5°C.

CONCLUSION

This study proposed a method to integrate the two-node physiological model and the data-driven method, random forest classifier, to predict the TSV of older people under various situations. The results show that the overall accuracy of the integrated model is 90.1%, which is more accurate than the PMV model and the majority of other data-driven studies. The integrated model can also reach above 80% accuracy under different building operation modes (NV/AC), older people's gender, surveyed
seasons, and climate zones. Another contribution of the integrated model is its linking of the thermal index \( SET \) and physiological parameters to \( TSV \). The correlation between \( SET \) and \( TSV \) is better than the traditional linear regression. It is also found that older people's core temperature may increases to a dangerous level (>38°C) when they feel slightly warm, which is consistent with the previous chamber experimental study.

ACKNOWLEDGMENTS

The research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) through the Discovery Grants Program [#RGPIN-2018-06734] and the "Advancing Climate Change Science in Canada" program the Advancing Climate Change Science in Canada Program [#ACCP] 535986-18, and the Construction Research Centre of the National Research Council of Canada through funding from Infrastructure Canada in support of the Pan Canadian Framework on Clean Growth and Climate Change. The authors were very thankful for their supports.
REFERENCES


https://doi.org/10.1016/j.enbuild.2020.109795