Estimating Thermal Sensation Probability Distribution Using Three Models

Dayi Lai¹, Chuanming Chen², Ruoyu You³, Chun Chen⁴,*
¹Shanghai Jiao Tong University, Shanghai, China
Tianjin University, Tianjin, China
³The Hong Kong Polytechnic University, Hong Kong SAR, China
⁴The Chinese University of Hong Kong, Hong Kong SAR, China

chunchen@mae.cuhk.edu.hk

Abstract
This study compares the linear regression model, ordered probability model, and multinomial logit model for prediction of the individual thermal sensation votes (TSVs) and TSV distributions under given conditions. A thermal comfort dataset from “Pakistan project” was used to develop and evaluate the models. The predictive capability of the three models were systematically evaluated and compared. The results showed that the ordered probability model and the multinomial logit model correctly predicted around 50% of the individual TSVs, whereas the accuracy of the linear regression model was only around 40%. In addition, the chi-square statistics show that the ordered probability model and the multinomial logit model better predicted the TSV distributions than the linear regression model.

Introduction
Indoor thermal comfort is strongly associated with the productivity and well-being (Lan et al., 2011) of the occupants. To maintain a thermally comfortable environment, a large amount of energy is needed for heating and air-conditioning in both residential and commercial buildings (Yang et al., 2014). Therefore, numerous studies have been conducted to explain indoor thermal comfort (Takasu et al., 2017; Martinez-Molina et al., 2017). Appropriate and effective thermal comfort models are in great need for the design of thermally comfortable spaces (Holmes and Hacker, 2007).

The most widely used indicator of occupants’ thermal comfort is thermal sensation, which is represented by several ordered categorical responses (ASHRAE, 2009). Many thermal comfort models were developed by correlating thermal sensation with some variables. For instance, the predicted mean vote (PMV) model (Fanger, 1972), which has been extensively used for indoor environments, correlates the human body’s thermal load with the mean thermal sensation of a group of occupants. However, large discrepancies were found between PMV and actual thermal sensation in previous field studies in different regions (van Hoof, 2008; de Dear et al., 1998b). Thus, to address local variations, linear regression models (Han et al., 2007; Givoni et al., 2003; Nikolopoulou, 2004) were developed by fitting thermal sensation with environmental and individual parameters. Except for linear regression models, researchers have also used other advanced statistical models, such as decision tree (Choi and Yeom, 2017) and support vector machine (SVM) (Dai et al., 2017), to fit the thermal sensation data.

Due to individual differences, occupants can experience different thermal sensations, even in the same thermal environment (Wang et al., 2018). Such variation indicates that the thermal comfort for a group of people is actually a distribution instead of a single value (Coley et al., 2017). Therefore, it is worthwhile to develop advanced models that can predict the probability distribution of thermal comfort. Fanger (1972) proposed the predicted percentage of dissatisfied (PPD) model to consider differences in potential individual thermal comfort. However, the PPD model can only determine the percentage of people who feel that they are outside their comfortable range (i.e., slightly cool, neutral, and slightly warm). Although linear regression models are able to predict the distribution of thermal sensation, to our best knowledge, its accuracy has not been well examined. Also, the linear regression models treated categorical thermal sensations as continuous data, which may result in a large error. Thus, more advanced models are needed to comprehensively describe the distribution of thermal sensations.

This study presents two advanced models, the ordered probability model and multinomial logit model, to predict the distribution of thermal comfort. Both models have been widely used in economic studies when the distribution information is to be modeled. It is worthwhile to evaluate their performances in prediction the distribution of thermal sensation. Traditional linear regression models were also developed to serve as the comparison baseline. A thermal comfort dataset taken from an indoor thermal comfort survey conducted in Pakistan was used to develop and evaluate the models. The predictive capability of these three models was systematically evaluated and compared to identify the best one for predicting the distribution of thermal comfort.

Methods
This section first presents the linear regression model, the ordered probability model, and the multinomial logit model. The thermal comfort databases used to develop the models are then briefly introduced.

Linear regression model
The linear regression model has been widely used in thermal comfort studies (de Dear et al., 1998b; Han et al., 2007; Givoni et al., 2003; Nikolopoulou, 2004). Therefore,
this study used it as the baseline model, which can be described as:

$$\mu = \alpha + \beta x$$  \hspace{1cm} (1)

where $\mu$ is the mean of thermal sensation vote (TSV). In this study, the American Society of Heating, Refrigerating, and Air-conditioning Engineers (ASHRAE) seven-point thermal sensations ($-3$ = cold, $-2$ = cool, $-1$ = slightly cool, $0$ = neutral, $1$ = slightly warm, $2$ = warm, and $3$ = hot) were adopted. $x$ is the explanatory variable, $\beta$ is the coefficient corresponds to $x$, and $\alpha$ is the intercept. Through least-squares regression, the coefficients $\beta$ and $\alpha$ can be estimated. Apart from the mean value $\mu$, the standard deviation $\sigma$ was calculated as:

$$\sigma = \sqrt{\frac{\sum (Y_i - \bar{Y})^2}{n-2}}$$  \hspace{1cm} (2)

where $Y_i$ and $\bar{Y}$, are the observed and predicted TSV for observation i, respectively, and $n$ is the total number of observations.

With the mean $\mu$, standard deviation $\sigma$, and the normal distribution assumption for the predicted outcomes, the probabilities for different categories of TSV can be obtained via:

$$P(Y = j) = \frac{\Phi(\mu - 2.5/\sigma)}{1 - \Phi(\mu - 2.5/\sigma)} \quad \text{if } j = -3$$

$$P(Y = j) = \frac{\Phi(\mu - j - 0.5/\sigma) - \Phi(\mu - j + 0.5/\sigma)}{1 - \Phi(\mu - 2.5/\sigma)} \quad \text{if } j = -2 \text{ to } 2$$

$$P(Y = j) = \frac{\Phi(\mu - 2.5/\sigma)}{1 - \Phi(\mu - 2.5/\sigma)} \quad \text{if } j = 3$$  \hspace{1cm} (3)

where $j$ represents each specific TSV ($j$ from $-3$ to $3$), and $\Phi$ is the cumulative distribution function of the standard normal distribution. The probabilities in Eq. (3) imply an assumption that the predicted TSV follows the rule defined in Eq. (4):

$$Y = \begin{cases} 
-3 & \text{if } \mu < -2.5 \\
 j & \text{if } -0.5 \leq \mu < j + 0.5, \text{ for } j = -2 \text{ to } 2 \\
 3 & \text{if } \mu \geq 2.5
\end{cases}$$  \hspace{1cm} (4)

Ordered probability model

The ordered probability model predicts the probability distribution of TSVs (Agresti, 2007; Washington et al., 2010; McKelvey and Zavoina, 1975; Lai et al., 2018; Lim et al., 2018). According to Agresti (2007) (page 180, Eq. 6.4), the logit transformation of the probability that $Y$ at or falls below a particular category $j$ can be calculated by the following expression:

$$\text{logit} [P(Y \leq j)] = \alpha_j + \beta_j x, \quad \text{for } j = -3 \text{ to } 2$$  \hspace{1cm} (5)

For various category $j$, the constant terms $\alpha_j$ are different, but the coefficient $\beta$ is the same. The $\alpha$ and $\beta$ are estimated via the maximum likelihood estimation (MLE) method. Please note Eq. (5) did not specify $P(Y \leq 3)$ because it simply equals 1. The probability for each level of thermal sensation can be calculated by Eq. (5).

The ordered probability model has several key features. Firstly, the outcome $Y$ is treated as an ordinal variable. Secondly, for each value of $Y$, the linear predictor has a specific intercept term, which represents the different thresholds of thermal sensation for going from one TSV category to the next. Thirdly, the regression coefficient, $\beta$, is the same for each category $Y$. Therefore, with one explanatory variable, the ordered probability model fits $6 + 1 = 7$ parameters compared with the two parameters for the linear regression model.

Multinomial logit model

The multinomial logit model can also be used to predict the probability distribution of TSVs (Agresti, 2007; Washington et al., 2010; Hausman and McFadden, 1984). According to Agresti (2007) (page 180, Eq. 6.3), the probability for thermal sensation category $j$ in the multinomial logit model is calculated as:

$$P(Y = j) = \frac{\exp(\alpha_j + \beta_j x)}{\sum_{j=-3}^{3} \exp(\alpha_j + \beta_j x)} \quad j = -3,...,3$$  \hspace{1cm} (6)

Note that in the multinomial logit model, the constant terms and coefficient terms are different for categories $-3$ to $3$. Similar to the ordered probability model, these constants and coefficient terms were also estimated by MLE method.

Compared with the ordered probability model, in the multinomial logit model, the linear predictor not only has a separate intercept but also has a separate slope for each outcome $Y$ category. Therefore, with one explanatory variable, there are $6 + 6 = 12$ parameters fitted for the multinomial logit model. Note that there is no implied ordering of the categories.

To make the model descriptions simpler and clearer, Eqs. (1), (4), and (6) used only one explanatory variable, but the models can be easily extended to include multiple predictor variables. All three models were developed using the commercial statistical software NLOGIT 6 (Greene, 2002).

Case descriptions

This study used the database from the “Pakistan project” (Nicol and Roaf, 1996) to develop the three models for prediction of indoor thermal comfort. The data were collected under a wide range of environmental conditions indoors. The full range of thermal sensation responses make the dataset suitable for model performance evaluation. The field survey was conducted in July 1993 and January 1994 in five cities in Pakistan, namely Karachi, Peshawar, Quetta, Saidu, and Multan. In each city, five local people were recruited. Each subject carried a portable data logger that recorded the air temperature, globe temperature, relative humidity, and air velocity every 5 minutes over a period of 1 week. During the test period, each subject took the questionnaire survey every hour when he or she was awake. The questions included the seven-point thermal sensations, clothing, and activity.

All of the measurements and the survey were conducted in naturally ventilated buildings. More details regarding the field survey can be found in Nicol and Roaf (1996). The data were downloaded from the ASHRAE RP-884 adaptive model project database (de Dear, 1998a). Because the observations in Multan were in summer only, this study excluded the Multan data to avoid potential skewed results. Since the surveys were conducted
repeatedly on a small number of subjects, to reduce the time series effects, this study selected only two observations per day to develop and validate the proposed models. The observations during 9:00 to 10:00 and 15:00 to 16:00 were used as the training dataset, and those from 12:00 to 13:00 and 18:00 to 19:00 were used to form the validation dataset. The training and validation datasets included 456 and 462 observations, respectively. The training dataset was used to develop the three models. The developed models were then used to predict the TSVs using the predictor variables of the observations in the validation dataset. The predicted TSVs were then compared with the observed TSVs in the validation dataset to validate the models.

This study selected four predictor variables, including air temperature ($T_a$), water vapor pressure ($P_v$), metabolic rate of the activity ($M_{act}$), and the clothing insulation ($I_c$), for the model development. These four variables were included as they are important factors of thermal sensation and can be measured or estimated easily in real applications. Although in naturally ventilated buildings, radiation and wind speed are also important factors influencing thermal sensation, these two parameters are relatively hard to obtain compared with $T_a$, $P_v$, $I_c$, and $M_{act}$. Thus, this study did not include radiation and wind speed in the model estimation. Table 1 descriptively summarizes the variables in the training and validation datasets for the indoor thermal comfort case. The datasets included a wide range of air temperature from 7.7 to 39.4 °C. Water vapor pressure also had a wide range (mean, 1.4 kPa; standard deviation (SD), 0.8 kPa). These wide ranges of the environmental parameters led to a wide range of clothing insulation values, from 0.3 to 2.2 clo. Metabolic heat generation ranged from 0.6 to 1.9 met, which corresponded to the metabolic heat generation of reclining and walking according to the ASHRAE handbook (ASHRAE, 2009). It can be seen from Table 1 that the distributions of $T_a$, $P_v$, $M_{act}$, and $I_c$, in the training and validation dataset were similar.

### Table 1: Summary of the predictor variables in the training and validation datasets.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Training dataset</th>
<th>Validation dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (std.)</td>
<td>Min.</td>
</tr>
<tr>
<td>$T_a$ (°C)</td>
<td>24.0 (7.2)</td>
<td>7.7/39.4</td>
</tr>
<tr>
<td>$P_v$ (kPa)</td>
<td>1.4 (0.8)</td>
<td>0.2/4.5</td>
</tr>
<tr>
<td>$M_{act}$ (met)</td>
<td>1.2 (0.4)</td>
<td>0.6/1.9</td>
</tr>
<tr>
<td>$I_c$ (clo)</td>
<td>1.1 (0.4)</td>
<td>0.4/2.2</td>
</tr>
<tr>
<td>Number of obs.</td>
<td>456</td>
<td>462</td>
</tr>
</tbody>
</table>

#### Evaluation criteria

The three models will first be evaluated in terms of predicting individual TSVs (point predictions) for the validation datasets. For the linear regression model, Eq. (4) was used to calculate the individual TSVs. For the ordered probability and multinomial logit models, this study considered the predicted individual TSV as the category that corresponded to the highest probability. This study used the percentage of correct predictions to assess the accuracy of the models in predicting the individual TSV for each observation. The percentage of correct predictions was simply calculated by the number of correct predictions over the number of total observations.

This study will then evaluate the three models in terms of predicting TSV distributions under given conditions. After the three models were developed, several subsets of data with different temperature levels were selected from the validation database. For example, three subsets with the temperature levels of 12.1±0.6, 19.4±0.8, and 30.1±0.5 °C (mean ± standard deviation), respectively, were selected. There were 46, 50, and 73 observations, respectively, in the three subsets. The probabilities for each of the seven TSVs were then calculated from the models using the mean value of the observed air temperatures in the subset of data. The calculated TSV distributions under the given temperature levels were compared with the actual vote distributions (transformed to percentages) among the observations in the corresponding subsets of data. For the evaluation of accuracy in the distribution predictions, this study used the chi-square test from Press et al. (2007) (page 622, Eq. 14.3.2) to examine the level of closeness between the actual distribution and the predicted distribution. The chi-square statistic, $\chi^2$, was calculated by:

$$\chi^2 = \sum_{j=-3}^{3} \frac{(R_j - S_j)^2}{R_j + S_j}$$

where $R_j$ and $S_j$ are the predicted and the actual numbers of votes for thermal sensation $j$ from −3 to 3, respectively. A smaller chi-square value indicates a better match between the predicted and actual TSV distribution.

#### Results

This section first presents the development of the linear regression model, ordered probability model, and multinomial logit model using the training dataset. The developed models were then used to predict the TSVs in the validation dataset. The predicted results were compared with the observed TSVs to examine how well the models can predict new cases.

In the first stage, we used only the air temperature as the predictor variable to develop the models. That was to make the description of the models, the presentation of the results, and the comparisons between models simpler. In the second stage, four variables were included in the models to examine the improvement in model predictive capability.

#### Model development

For the indoor case, this study first developed a linear regression model as the baseline by using only $T_a$ as the single predictor variable. Table 2 summarizes the estimated coefficients $\alpha$ and $\beta$, their standard errors, and p-values. The model shows that the increase in $T_a$ led to an increase in TSV. Despite of only using $T_a$ as the
predictor variable, the developed linear regression model had a moderate $R^2$ value of 0.548.

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Estimated value</th>
<th>Standard error</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>15.130</td>
<td>3.289</td>
<td>0.000</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.549</td>
<td>0.058</td>
<td>0.000</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>5.872</td>
<td>3.168</td>
<td>0.064</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.833</td>
<td>0.087</td>
<td>0.000</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.990</td>
<td>0.114</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**Comparison of the model predictive capability**

In this subsection, the three developed models were used to predict the TSVs for the validation dataset. The predicted TSVs were then compared with the observed TSVs to examine the model performances of predicting new cases.

The predicted TSVs were first compared with the observed TSVs to examine how well the models predict new cases. The confusion matrices shown in Tables 5 to 7 were generated to demonstrate the accuracy of the models in predicting individual TSVs in the validation dataset. The most frequently occurred TSV in the survey was neutral ($n = 192$). The ordered probability model and multinomial model were able to correctly predict 153 and 155 of the neutral votes, while the linear regression model only had 81 correct predictions. It can also be seen from the confusion matrices that the ordered probability and multinomial models were better than the linear regression model in predicting the extreme TSV such as “cold” and “hot”. Table 5 shows that the linear regression model never predicted a TSV of $-3$ or $3$ even though there were 40 observed TSV of $-3$ and 27 of observed TSV of $3$. Interestingly, Tables 6 and 7 shows that both the ordered probability and multinomial models never predicted a TSV of $-1$, while there were 43 observed TSV of $-1$.

**Table 5: Confusion matrix of the predicted TSVs from the developed linear regression model v.s. the observed TSVs in the validation dataset.**

<table>
<thead>
<tr>
<th>TSV</th>
<th>Survey</th>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>-2</td>
<td>17</td>
<td>28</td>
<td>12</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>62</td>
</tr>
<tr>
<td>1</td>
<td>17</td>
<td>7</td>
<td>9</td>
<td>50</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>86</td>
</tr>
<tr>
<td>-1</td>
<td>5</td>
<td>1</td>
<td>21</td>
<td>81</td>
<td>19</td>
<td>10</td>
<td>1</td>
<td>0</td>
<td>138</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>55</td>
<td>49</td>
<td>25</td>
<td>19</td>
<td>1</td>
<td>0</td>
<td>151</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>10</td>
<td>7</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>40</td>
<td>37</td>
<td>43</td>
<td>192</td>
<td>78</td>
<td>45</td>
<td>27</td>
<td>0</td>
<td>462</td>
</tr>
<tr>
<td>Total</td>
<td>40</td>
<td>37</td>
<td>43</td>
<td>192</td>
<td>78</td>
<td>45</td>
<td>27</td>
<td>0</td>
<td>462</td>
</tr>
</tbody>
</table>
Table 6: Confusion matrix of the predicted TSVs from the developed ordered probability model v.s. the observed TSVs in the validation dataset.

<table>
<thead>
<tr>
<th>TSV</th>
<th>-3</th>
<th>-2</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>17</td>
<td>27</td>
<td>11</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>57</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>11</td>
<td>33</td>
<td>44</td>
<td>14</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>40</td>
<td>37</td>
<td>43</td>
<td>192</td>
<td>78</td>
<td>45</td>
<td>27</td>
</tr>
</tbody>
</table>

Table 7: Confusion matrix of the predicted TSVs from the developed multinomial logit model v.s. the observed TSVs in the validation dataset.

<table>
<thead>
<tr>
<th>TSV</th>
<th>-3</th>
<th>-2</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>16</td>
<td>26</td>
<td>11</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>57</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>22</td>
<td>30</td>
<td>30</td>
<td>55</td>
<td>45</td>
<td>14</td>
<td>36</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>28</td>
<td>25</td>
<td>14</td>
<td>9</td>
<td>15</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>8</td>
<td>16</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>40</td>
<td>37</td>
<td>43</td>
<td>192</td>
<td>78</td>
<td>45</td>
<td>27</td>
</tr>
</tbody>
</table>

By summarizing the information from the confusion matrices, the absolute error distributions for the three models were calculated and presented in Figure 1. The percentages of correct predictions of ordered probability and multinomial logit models were 45% and 47%, respectively, while that of the linear regression model was 39%. Again, the ordered probability and multinomial logit models were more accurate than the linear regression model in predicting individual TSVs for new cases.

Figure 1: Comparisons of the distribution of the absolute error between the surveyed and predicted individual TSVs by the developed linear regression (Linear), ordered probability (OP), and multinomial logit (ML) models for the validation dataset.

In addition to individual TSVs, this study also assessed the accuracy of the three models in predicting TSV distribution under given conditions. Three scenarios with the air temperature levels at 12.1±0.6, 19.4±0.8, and 30.1±0.5 °C (mean ± standard deviation), respectively, were selected from the validation dataset. The three scenarios had 46, 50, and 73 observations, respectively. The probabilities for each of the seven TSVs were calculated from the models using the mean value of the observed air temperatures for each scenario and compared with the actual vote percentage distributions. For the first scenario presented in Figure 2(a), the actual TSV distribution was concentrated on the “cold”, “cool”, and “slightly cool” categories. However, the predicted distributions by the three models had larger variances than the actual situation. For the second and third scenarios in Figure 2(b) and Figure 2(c), respectively, the ordered probability and multinomial logit model clearly provided better predictions of the TSV distribution than the linear regression model. This was also quantitatively reflected in the chi-square values summarized in Table 8. The chi-square statistics for the linear regression model in the second and third scenario were 25.9 and 10.9, respectively, while the other two models had chi-square values of around 5.

By summarizing the information from the confusion matrices, the absolute error distributions for the three models were calculated and presented in Figure 1. The percentages of correct predictions of ordered probability and multinomial logit models were 45% and 47%, respectively, while that of the linear regression model was 39%. Again, the ordered probability and multinomial logit models were more accurate than the linear regression model in predicting individual TSVs for new cases.

Figure 2: Comparisons between the surveyed and predicted TSV distribution by the developed linear regression (Linear), ordered probability (OP), and multinomial logit (ML) models in the three scenarios for indoor validation dataset: (a) T<sub>a</sub> = 12.1 °C; (b) T<sub>a</sub> = 19.4 °C; and (c) T<sub>a</sub> = 30.1 °C.
To examine the effect of more predictor variables on the model fitting and predictive capability, this study developed models based on the training dataset by using air temperature, water vapor pressure, metabolic rate, and clothing insulation as predictor variables. Due to the limited space, the four variable models were not presented in this paper. The variance inflation factor (VIF) was used to examine the level of multicollinearity in the four-variable models. The VIFs were all less than 1.6, indicating an acceptable level of multicollinearity. The model performances of the four-variable models were compared with those of one-variable models.

This study examined how well the developed models predict the TSVs for new cases in the validation dataset. Figure 3 shows the percentages of correct predictions from the three four-variable models in predicting individual TSVs. Table 9 lists the chi-square statistics of the four-variable models for predicting the TSV distributions in indoor validation dataset. Both comparisons show no significant improvements in the model accuracy in the linear regression and ordered probability models. For the multinomial logit model, the percentage of correct prediction was improved by 8%.

**Four variable models**

The results in this study show that the ordered probability and multinomial logit models were 10% more accurate than the linear regression model in predicting the individual TSVs. The results also show that the ordered probability and multinomial models had much lower chi-square values than that of the linear regression model. Thus, the two models gave more precise TSV distribution predictions than the linear regression model. Interestingly, the multinomial logit model was slightly more accurate than the ordered probability model in predicting individual TSVs. However, the ordered probability model was slightly more accurate than the multinomial logit model in predicting TSV distributions under given conditions.

Thermal sensation has orders, that is, from cold (−3) to hot (3). The ordered probability model treats the TSV as an ordered categorical variable. The probabilities for each possible TSV value are specified by areas under a logistic curve. Furthermore, the thresholds for each categorical TSV defining those areas are estimated from the data along with the regression coefficients. However, for the linear regression model, the thresholds for each categorical TSV are fixed and equally spaced. Thus, the probabilities are only adjustable by varying the regression coefficients and the error standard deviation. That was one of the reasons why the ordered probability model fitted the data better than the linear regression model.

Different from the ordered probability model, the multinomial logit model treats all thermal sensations equally without considering their ordered properties. This model is effectively a set of separate binary regression models. Each separate regression model is fitted to each categorical TSV value regarded as a binary variable. Therefore, the equation system in the multinomial logit model is more complex than that of the ordered probability model. Each thermal sensation has its own function with different combinations of variables and coefficients in the multinomial logit model. This may be the reason why the multinomial logit model performed better than the ordered probability model in predicting individual TSVs. However, the extra order information in the ordered probability model may be the reason why the ordered probability model performed better than the multinomial logit model in terms of TSV distribution prediction.

It is worthy of note that the developed ordered probability model and multinomial logit model cannot be applied.
outside the scope in which the data were collected. In contrast, physiological models, such as the predicted mean vote (PMV) (Fanger, 1972), are usually considered as universally applicable. However, the development of these physiological models still uses linear regressions in some steps because thermal sensations cannot be described using a purely physical model. If these linear regressions can be replaced by ordered probability or multinomial logit regressions, then the new physiological models may be more realistic and informative because they take advantage of both the physical derivation approach and the advanced probability-based statistical approach. Furthermore, this study randomly split the data into training and validation sets. However, for the Pakistan data, different times of days were used for the training and validation datasets. As the subjects’ thermal sensations might change systematically with time, the random splitting method may lead to certain levels of errors. This issue deserves further investigation in the future. In addition, it is also worthwhile to include interaction terms, such as a combination of clothing and activity, in the models to improve the model fitting and predictive capability in future studies. As a normal practice in literature (Washington et al., 2010), the ordered probability model and multinomial model regard the category with the highest probability as the predicted thermal sensation in point estimation. However, it is hard to ascertain such assumption, especially if the largest probability is not very large. Finally, in addition to the methods introduced in this paper, another way of predicting the thermal sensation distribution is through the Monte Carlo method, if the distribution of the predictor variables is available.

This study extended the model from one predictor to four predictors. Although the accuracy improvement was limited, the four variable model may provide additional benefits. For example, with more predictors, it may be possible to create a model with good prediction accuracy from a smaller dataset.

Conclusions

This study systematically compared the linear regression model, the ordered probability model, and the multinomial logit model for predicting the distribution of thermal comfort. The data from an indoor thermal comfort survey in Pakistan was used for the model development and evaluation. Within the scope of this study, the following conclusions can be drawn:

- The ordered probability and multinomial logit models were 10% more accurate than the linear regression model in predicting individual TSVs.
- Based on the quantitative chi-square comparison, the ordered probability and multinomial logit models were more accurate than the linear regression model in predicting TSV distributions under given conditions.
- The multinomial logit model was slightly more accurate than the ordered probability model in predicting individual TSVs.
- The ordered probability model was slightly more accurate than the multinomial logit model in predicting TSV distributions under given conditions.

Acknowledgement

This work was partially supported by the project #RNE-p1-18 of the Shun Hing Institute of Advanced Engineering, The Chinese University of Hong Kong.

References


