Supplementary material

Effects of microclimatic and human parameters on outdoor thermal sensation in the high-density tropical context of Dhaka

Tania Sharmin[a,1] and Koen Steemers[1]

Study area

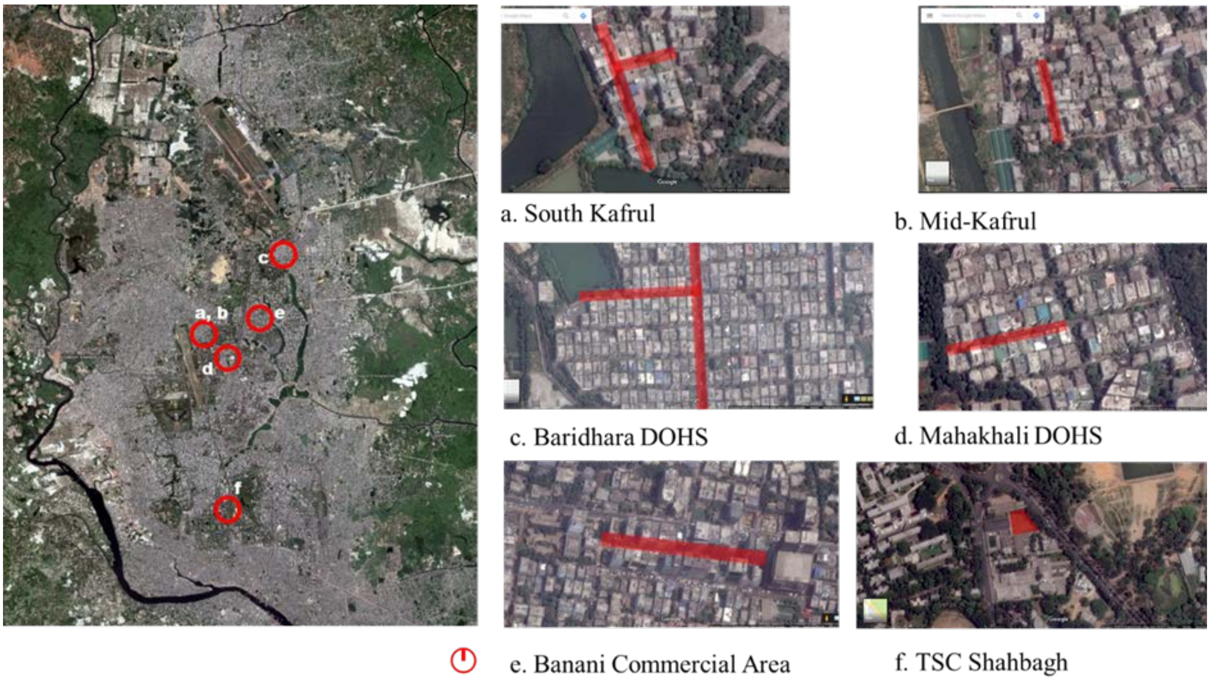

Figure 1. Overview of the case study area (a) South Kafrul (b) Mid Kafrul (c) Baridhara DOHS (d) Mahakhali DOHS, (e) Banani Commercial Area and (f) TSC Shahbagh

The case study areas were selected on the basis of providing distinct “local climate zones” (Stewart and Oke, 2012) as explained in (Sharmin, Steemers and Matzarakis, 2015). The residential areas *South Kafrul* and *Mid-Kafrul* are traditional residential areas of “spontaneous urban development”, whereas, *Mahakhali DOHS* and *Baridhara DOHS* are “formally planned residential areas”. *Banani Commercial Area* is a “high-rise development” for commercial activity and *TSC Shahbagh* is a comparatively “low-density area” with educational establishments. The areas correspond to the climate zones described by Stewart and Oke (2012) as: ‘Compact mid-rise-traditional’(sub-class), ‘Compact mid-rise-formal’(sub-class) ‘Compact high-rise’ and ‘Open mid-rise’ respectively.

[1] The Martin Centre for Architectural and Urban Studies, Department of Architecture, University of Cambridge, 1-5 Scroope Terrace, CB2 1PX, Cambridge, United Kingdom.

[a]Corresponding author: ts531@cam.ac.uk

Table 1. Description of the sample

Gender		Age			Body Type			Activity		
Female	311 (24%)	30 - 50	487	38%	Normal	971	76%	Light walking	479	37%
Male	975 (76%)	Over 50	65	5%	Obese	69	5%	Moderate walking	20	2%
		Below 16	47	4%	Skinny	229	18%	Sitting	131	10%
		16 - 30	678	53%	NA's	17	1%	Standing, light work	636	49%
		NA's	9	1%				NA's	20	2%

Body exposure to the sun		Clothing (clo)		Time living in Dhaka		Profession		Profession type	
No	92%	Min.	0.10	Above 5 years	76%	Student	31%	Indoor type	73%
Yes	5%	Median	0.50	Below 1 year	5%	Office job	37%	Outdoor type	26%
NA's	2%	Mean	0.50	Below 5 years	19%	Business	7%	NA's	1%
		Max.	1.20	NA's	0%	Housewife	6%		100%
		NA's	11.00			(Other)	18%		
						NA's	1%		

Exposure to air-conditioned space in the last 30min		Travelling in last 30min		Reason for being on the site	
No	71%	No	79%	Close to home, office, school or station	80%
Yes	29%	Yes	21%	To take rest and enjoy environment (sunshine, breeze, nice view)	6%
				For No particular reason	4%
				For prayer	2%
				(Other)	9%
				NA's	0%

Ordinal Logistic Regression (OLR)

The ordinal logistic regression is a form of logistic regression which is used to model the relationship of an ordinal dependent variable and a set of independent variables that are either categorical or continuous. In an ordinal logistic regression model, the outcome variable is ordered and has more than two levels. The distance between the levels is generally unknown (Christensen, 2011). In this study, the ordinal outcome variable is the thermal sensation vote (TSV), which is coded on a 7-point scale as: Cold = -3; Cool = -2; Slightly cool = -1; Neutral = 0; Slightly warm = 1; Warm = 2; Hot = 3.

In ordinal regression, instead of modelling the probability of an individual event, as done in logistic regression, the probability of that event and all others below it in the ordinal ranking are considered. That means it deals with cumulative probabilities rather than probabilities for discrete categories. This process in logistic regression to deal with the ordinal response is known as the proportional odds model (Fitrianto and Maziatul, 2014). The odds are modelled as:

$$\theta_j = \frac{P(Y \leq j)}{1 - P(Y \leq j)} \qquad \textbf{(1)}$$

The ordinal regression model can be expressed in the logit form as the following:

$$logit\ [\theta_j] = ln\left(\frac{P(Y \leq j)}{1 - P(Y \leq j)}\right) = \alpha_j + (-\beta_1 X_1 - \beta_2 X_2 - \cdots - \beta_i X_i) \qquad \textbf{(2)}$$

Here, P is the probability, Y is the ordinal variable, $Y \leq j \mid (X_1, X_2, \ldots, X_i)$ is the probability of being at or below category j, given a set of predictors $(X_1, X_2, \ldots, X_i)$, j is the number of ranking of the dependent variable = $j = 1, 2, \ldots, J-1$, αj are the cut-points and $\beta 1$, $\beta 2, \ldots, \beta i$ are logit coefficients.

By exponentiation of Equation (8) the following equation can be found:

$$\left(\frac{P(Y \leq j)}{1 - P(Y \leq j)}\right) = e^{\alpha_j + (-\beta_1 X_1 - \beta_2 X_2 - \cdots - \beta_i X_i)} = e^{\alpha_j - \sum \beta_i X_i}$$

$$=> P(Y \leq j) = \frac{e^{\alpha_j - \sum \beta_i X_i}}{1 + e^{\alpha_j - \sum \beta_i X_i}} = \frac{1}{1 + e^{-(\alpha_j - \sum \beta_i X_i)}} \qquad \textbf{(3)}$$

In the model, each logit has its own α_j which is called the threshold value and its values do not depend on the values of the independent variable for a particular case. The model has the same effects $\beta_i X_i$ for each logit. The equal logit slope or the proportional odds assumption (the assumption of the same coefficients across all cut-points/ thresholds) must be tested before applying an ordinal regression. In this study, the proportional odds assumption was tested by the method suggested in Agresti (2010) and is determined by chi-square test statistics before conducting each ordinal regression analysis. The probability of each category was determined: $P(Y \leq -1)$, $P(Y \leq 0)$, $P(Y \leq 1)$, $P(Y \leq 2)$ and $P(Y \leq 3)$, with $P(Y \leq 3) = 1$, as per the condition of cumulative probability. While determining the category of each response, the category with the highest probability was chosen. For example, if the probability of a respondent feeling 'Slightly cool', 'Neutral', 'Slightly warm', 'Warm' or 'Hot' are 0.012 (1%), 0.141 (14%), 0.407 (41%), 0.289 (29%) and 0.151(15%) respectively, then the respondent is considered to belong in the 'Slightly warm' category with the highest probability of 41%. The probability of each category is predicted by the difference of the cumulative probability of each consecutive category. For example, *P (Y≤ 3)* is the cumulative probability of the earlier four categories plus its own (-1, 0, 1, 2 and 3). Therefore, *P (3) =P (≤3)-P (≤2).*

To understand the process, OLR was carried out with a single predictor variable air temperature first. The coefficients and standard errors can be found in Table 2.

Table 2. TSV_predicted ordinal single meteorological model

Parameters	**Coefficients**	**Standard error**		**Further model parameters**	
Slightly cool=−1	10.799	0.927		Pseudo R2	0.229
Neutral=0	13.495	0.912		gamma	0.524
Slightly warm = 1	15.448	0.933		std. error	0.031
Warm=2	16.936	0.952		CI	0.463 - 0.584
Hot=3	0.000			cor.test	0.380
Air temperature	0.478	0.029			
CI (2.5%, 97.5%)	0.422 - 0.535				
Exponentiated Coefficient/ odds ratio	1.612				

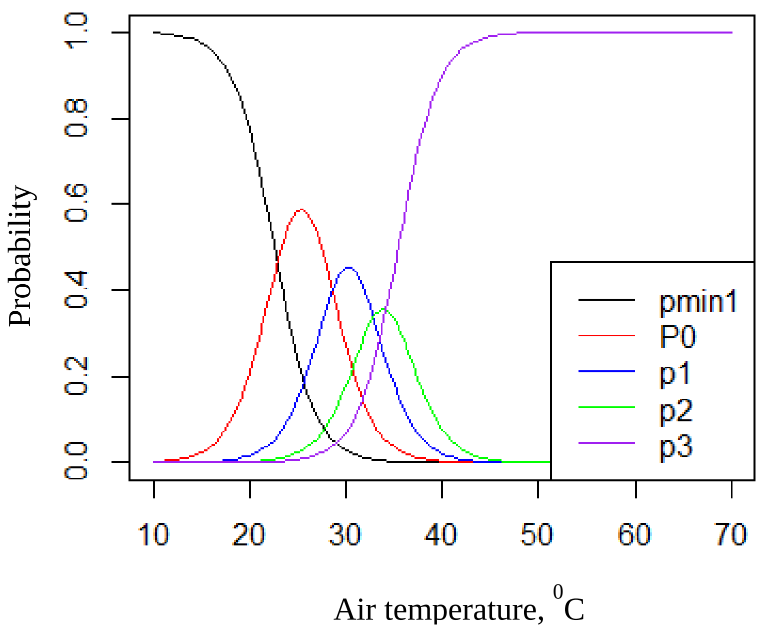

Figure 2. Probability diagram of air temperature against TSV[2]

The probability diagram for the ordinal single meteorological model can be seen in Figure 2. As the temperature increases, the probability of feeling 'Slightly cool' decreases and

[2] Since there was no response in 'Cool' (-2) and 'Cold' (-3) categories, these are not presented in the figure.

reaches P (probability) = 0 at 35.0^{0}C. This means the probability of feeling 'Slightly cool' becomes zero at 35.5^{0}C. Likewise, the probability of feeling 'Hot' increases when air temperature exceeds 30.0^{0}C and reaches P (probability) = 1 at 45.0^{0}C. The probability of feeling 'Neutral', 'Slightly warm' and 'Warm' reaches highest levels at 26.0^{0}C, 31.0^{0}C and 35.0^{0}C respectively. The odds ratio, that is simply the inverse log (i.e. the exponential) of the estimated coefficient, can be read from Table 2. The interpretation of the odds ratio is that, groups greater than *j* versus those who are in groups less than or equal to *j* are compared, where *j* is the level of the response variable. This means that for a one-unit change in the predictor variable, the odds for cases in a group that is greater than *j* versus less than or equal to *j* are the proportional odds times larger. For example, when air temperature moves 1 unit, the odds of TSV being in the 'Hot' category are 1.612 times greater than TSV being in 'Warm' and lower category.

Bibliography

Agresti, A. (2010) 'Modeling Ordinal Categorical Data', pp. 1–75. Available at: http://www.stat.ufl.edu/~aa/ordinal/agresti_ordinal_tutorial.pdf.

Christensen, R. (2011) 'Analysis of ordinal data with cumulative link models—estimation with the ordinal package', *R-package version*, pp. 1–31. Available at: http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Analysis+of+ordinal+data+with+cumulative+link+models+?+estimation+with+the+R+-package+ordinal#0.

Fitrianto, A. and Maziatul, M. I. (2014) 'Modeling employments rate data using ordinal logistic regression', *Middle - East Journal of Scientific Research*, 20(12), pp. 1715–1720. doi: 10.5829/idosi.mejsr.2014.20.12.21085.

Sharmin, T., Steemers, K. and Matzarakis, A. (2015) 'Analysis of microclimatic diversity and outdoor thermal comfort perceptions in the tropical megacity Dhaka , Bangladesh', *Building and Environment*. Elsevier Ltd, 94(November), pp. 734–750. doi: 10.1016/j.buildenv.2015.10.007.

Stewart, I. D. and Oke, T. R. (2012) 'Local climate zones for urban temperature studies', *Bulletin of the American Meteorological Society*, 93(12), pp. 1879–1900. doi: 10.1175/BAMS-D-11-00019.1.