Using Robust Locally Weighted Regression with Adaptive Bandwidth to Predict Occupant Comfort

Carlo Manna[1] **Nic Wilson**[2] **Kenneth N. Brown**[3]

Abstract. One of the main consumers of energy in buildings are the HVAC systems intended to maintain the internal environment for the comfort and safety of the occupants. Occupant satisfaction, is influenced by many different factors, including air temperature, radiant temperature, humidity, the outdoor environment, activity levels and clothing. Occupant thermal comfort is traditionally measured by the Predicted Mean Vote (PMV) metric, which estimates the expected response of the occupants on a seven point scale. PMV is a statistical measure, which holds for large populations. For small groups, however, the actual thermal comfort could be significantly different, and so energy may be wasted trying to achieve unwanted conditions. In this paper, we apply Locally Weighted Regression with Adaptive Bandwidth (LRAB) to learn individual occupant preferences based on historical reports. As an initial investigation, we attempt to do this based on just one input parameter, the internal air temperature. Using publicly available datasets, we demonstrate that this technique can be significantly more accurate in predicting individual comfort than PMV, relies on easily obtainable input data, and is much faster to compute. It is therefore a promising technique to be used as input to adpative HVAC control systems.

1 INTRODUCTION

One of the primary purposes of heating, ventilating and air conditioning (HVAC) systems is to maintain an internal environment which is comfortable for the occupants. Accurately predicting comfort levels-for the occupants should avoid unnecessary heating or cooling, and thus improve the energy efficiency of the HVAC systems. A number of thermal comfort indices (indicators of human comfort) have been studied for the design of HVAC systems [1,2]. However, the most widely used thermal comfort index is the predicted mean vote (PMV) index [1]. This conventional PMV model predicts the mean thermal sensation vote on a standard scale for a large group of persons in a given indoor climate. It is a function of two human vari-ables and four environmental variables, i.e. clothing insulation worn by the occupants, human activity, air temperature, air relative humidity, air velocity and mean radiant temperature, respectively. The values of the PMV index have a range from -3 to +3, which corresponds to the occupants thermal sensation from cold to hot, with the zero value of PMV meaning neutral. The conventional PMV model has been an international standard since the 1980s [3,4]. It has been validated by many studies, both in climate chambers and in buildings [5,6]. The standard approach to comfort-based control involves regulating the internal environment variables to ensure a PMV value of 0 [7,8]. Though the conventional PMV model predicts thermal sensations well, it is a nonlinear relation, and it requires iteratively computing the root of a nonlinear equation, which may take a long computation time. Therefore, Fanger [1] and ISO [4] suggest using tables to determine the PMV values of various combinations between the six thermal variables. Secondly, PMV is a statistical measure, based on field studies over large populations. For small groups of people, within a single room or zone of a building, PMV may not be an accurate measure. Moreover, computing PMV requires knowledge of a number of variables that may be hard to obtain (i.e., activity level and clothing). Therefore, in this study we consider an alternative practical approach to predicting thermal comfort through the automatic learning of the comfort model of each user based on his historical records.

[1] Cork Constraint Computation Centre, Dept. of Computer Science, University College Cork, e-mail: c.manna@4c.ucc.ie
[2] Cork Constraint Computation Centre, Dept. of Computer Science, University College Cork, e-mail: n.wilson@4c.ucc.ie
[3] Cork Constraint Computation Centre, Dept. of Computer Science, University College Cork, e-mail: k.brown@cs.ucc.ie

2 THE METHOD

Locally weighted regression is a technique with many strengths, some discussed in detail [9]. Let (x_i, y_i) denote a response, y_i, to a recorded value x_i, for $i = 1, \dots, n$. The aim is to assess the response y for a new value x. The approach is concerned to estimate a local mean, fitting the recorded data by means of a local linear regression. This involves solving the least squares problem, where α and β are the values that minimize (1):

$$min \sum_{k=1}^{n} \big(y_i - \alpha - \beta(x_i - x)^2\big)\omega(x_i - x; h) \tag{1}$$

Then α is the response y for the new entry point x. The kernel function $\omega(x_i - x; h)$, is generally a smooth positive function which peaks at 0 and decreases monotonically as increases in size. The smoothing parameter h controls the width of the kernel function and hence the degree of smoothing applied to the data.

There are many criteria to choose the kernel function based on the theoretical model of the function that has to be fitted. We follow the method in [10]. The kernel function for our problem is:

$$\omega = \left(1 - \left(\frac{|x_i - x|}{h}\right)^3\right)^2 \tag{2}$$

for $|x_i - x| \le h$; otherwise $\omega = 0$.

2.1 Adaptive bandwidth

Finally, we need to choose the bandwidth h. The choice here needs to take into account the fact that the density of the recorded data may be variable. In particular, there may be areas in which the data are clustered closely together (which suggests a narrow bandwidth), while, on other hand, other areas may be characterised by sparse data (in which a choice of a large bandwidth is more appropriate). In view of this, it would be appropriate to have a large smoothing parameter where the data are sparse, and a smaller smoothing parameter where the data are denser (Figure 1). In this situation an adaptive parameter has been introduced. Let the ratio k/n describe the proportion of the sample which contributes positive weight to each local regression (for example if the ratio is 0.7, it means that 70% of the recorded data contributes to the regression). Once we have chosen k/n (that means we have chosen k, as n is fixed), we select the k nearest neighbour from the new entry point x. Then, the smoothing parameter h is denoted by the distance of the more distant neighbour among the k neighbour selected. It should be noted that the entire procedure requires the choice of a single parameter setting.

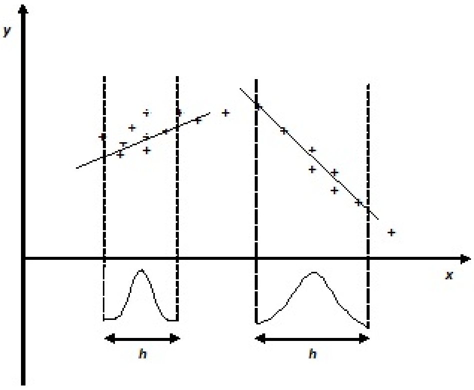
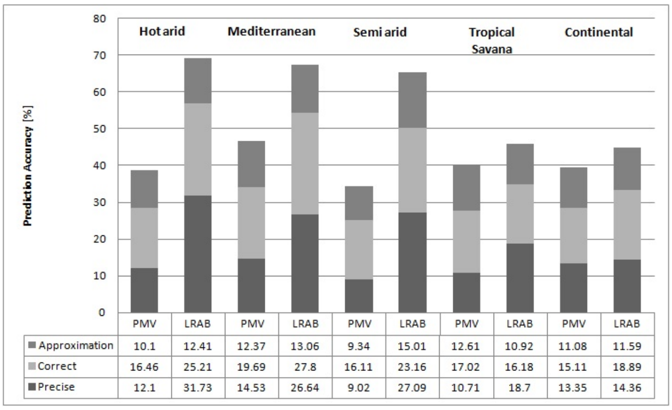


Figure 1. In locally weighted regression, points are weighted by proximity to the current x in question using a kernel function. A linear regression is then computed using the weighted points. Here, an adaptive bandwidth h based on the density of the recorded data is proposed.

3 EXPERIMENTS

The proposed LRAB has been compared with the PMV index on real data from ASHRAE RP-884 database [11]. This collection contains 52 studies with more than 20,000 user comfort votes from different climate zones. However, some of these field studies contain only few votes for each user. Thus they are not well suited for testing the proposed algorithm. This is because our approach seeks to learn the user preferences based on their votes, and it requires sufficiently many data records. For this reason, only the users with more than 10 votes have been used to compute the proposed LRAB. After removing the studies and records as described above we were left with 5 climate zones, 223 users and 7552 records. As a starting point, only one environmental variable (i.e. inside temperature) has been taken into account in order to evaluate the proposed LRAB. The proposed LRAB has been implemented in Matlab$^{\text{TM}}$, using the *trust-region method* to minimize the problem in (1), with a termination tolerance of 10^{-6}. As with the field study [11, 12], the algorithms are evaluated considering the difference ΔV between the computed votes by both LRAB (evaluated) and PMV (reported in the database) and the actual vote (reported in the database) on a three-level accuracy scale [11, 12] as reported below:

- *Precise*: $\Delta V < 0.2$
- *Correct*: $0.2 \leq \Delta V < 0.5$
- *Approximation*: $0.5 \leq \Delta V < 0.7$



	PMV	LRAB	PMV	LRAB	PMV	LRAB	PMV	LRAB	PMV	LRAB
Approximation	10.1	12.41	12.37	13.06	9.34	15.01	12.61	10.92	11.08	11.59
Correct	16.46	25.21	19.69	27.8	16.11	23.16	17.02	16.18	15.11	18.89
Precise	12.1	31.73	14.53	26.64	9.02	27.09	10.71	18.7	13.35	14.36

Figure 2. Average accuracy of predicting user comfort in 5 different climate zones.

Figure 2 illustrates how accurately the LRAB predicts the actual comfort vote of each user compared with PMV. In all 5 climate zones, LRAB is able to predict the actual vote better than PMV especially in the accuracy level < 0.2 and has up to 200% of the number of occupants for whom a correct value is predicted. Since this value can be computed quickly, and requires only a single setting parameter that is easily obtained, this method is feasible for use as a comfort measure in real time control. The next step will be the extension of the method to accept multiple environment variables (for example humidity etc.) in order to improve the above results.

REFERENCES

[1] Fanger, P. O. Thermal comfort: analysis and applications in environmental engineering. New York: McGraw-Hill; 1972.

[2] Gagge AP, Fobelets AP, Berglund LG. A standard predictive index of human response to the thermal environment. ASHRAE Trans 1986;92(2B):70931.

[3] ASHRAE-55. Thermal environmental conditions for human occupancy. American Society of Heating, Refrigerating and Air-conditioning Engineers Inc.; 1992.

[4] International Standard ISO 7730, Moderate thermal environments-estimation of the PMV and PPD indices and specification of the conditions for thermal comfort, Geneva; 1987.

[5] Busch JF. A tale of two populations: Thermal comfort in air conditioned and naturally ventilated offices in Thailand. Energy Buildings 1992;18:23549.

[6] de Dear Rj, Leow KG, Ameen A. Thermal comfort in the humid tropics. Part I. Climate chamber experiments on temperature preferences in Singapore. ASHRAE Trans 1991;97(1):8749.

[7] Funakoshi S, Matsuo D. PMV-based train air-conditioning control system. ASHRAE Trans 1995;101(1):42330.

[8] Yang KH, Su CH. An approach to building energy savings using the PMV index. Building Environment 1997;32(1):2530.

[9] H Astie , T. and Loader, C., Local regression: automatic kernel carpentry, Statist. Science 8, 120-143, 1993.

[10] Cleveland, W.S., Robust Locally Regression and Smoothing Scatterplots, J. of the American Statistical Association,Vol.74, No. 368, 829-836, 1979.

[11] de Dear, R.J., A global database of thermal comfort experiments, ASHRAE Trans., 104(1):1141-1152, 1998.

[12] A. Schumann, N. Wilson, M. Burillo, Learning User Preferences to Maximise Occupant Comfort in Office Buildings, IEA/AIE (1) 2010: 681:690.