2.4 A COMPARISON OF BIOMETEOROLOGICAL COMFORT INDICES AND HUMAN MORTALITY DURING HEAT WAVES IN THE UNITED STATES

Robert E. Davis*, David Knight, David Hondula
University of Virginia, Charlottesville, Virginia

Paul C. Knappenberger
New Hope Environmental Services, Inc., Charlottesville, Virginia

1. INTRODUCTION

Numerous biometeorological indices have been developed to account for the thermal stresses encountered by the human body in varying ambient environments. The models from which these indices have been devised are often based on fundamental physics and incorporate factors such as the fluxes of heat and moisture from bare skin, the influence of clothing on these fluxes, and the impacts of radiant absorption and metabolism. The resulting indices provide some measure of the relative level of comfort experienced by a modeled human body in a given environment. Examples include Physiological Equivalent Temperature, Perceived Temperature, and Predicted Mean Vote.

Other measures, not designed specifically for the purposes of examining human biometeorology, have also been widely applied in human comfort studies. One current popular example is the Spatial Synoptic Classification (SSC) (Sheridan 2002), an air-mass-like categorical classification that serves as the basis of heat and human health watch/warning systems in implementation world-wide (Kalkstein et al. 1997).

To date, no study has compared the relative efficacy of these measures in identifying heat wave deaths. In this paper, we examine seven different well-known indices that are commonly-used in human comfort studies with respect to summer heat-wave mortality in four major U.S. cities.

2. DATA

Daily total deaths (from all causes) are calculated from a database acquired from the National Center for Health Statistics for April through September, 1964–1998 (excluding the years 1967–1972 because of data availability issues) (NCHS, 1978). Data are gathered by county, grouped into Metropolitan Statistical Areas (MSAs), and then age-standardized (using ten standard age classes) to the average U.S. population using common epidemiological techniques (Anderson and Rosenberg, 1998). Four MSAs are included in this analysis: Boston, Massachusetts (BOS); Baltimore, Maryland (BAL); Philadelphia, Pennsylvania (PHI); and Minneapolis, Minnesota (MIN). Hourly weather data were acquired from a major first-order weather station within each MSA. All of the comfort index calculations except for the SSC air mass classification are derived using 1600 LST observations. The Spatial Synoptic Classification requires diurnal information and utilizes four observations per day.

3. INDICES

Apparent Temperature (AT)

Apparent Temperature (AT) (Steadman 1979) quantifies the physiological effects of high heat and high humidity. While AT can easily be calculated as a function of the ambient temperature and moisture, the index includes environmental and physiological variables important in determining human response to environmental stresses. These variables include heat generation and loss, fabric resistance, vapor pressure, wind speed, solar radiation, terrestrial radiation, proportion of body clothed, and other factors (Steadman 1984). When constants are input for these parameters, the index combines temperature and humidity into a single variable. AT is related to the commonly-used heat index in the United States.

Corresponding author address:
Robert E. Davis, University of Virginia,
Department of Environmental Sciences, P.O.
Box 400123, Charlottesville, VA 22904-4123;
e-mail: red3u@virginia.edu.

Perceived Temperature (PT)

The perceived temperature, with units of degree Celsius, describes a reference environment with fixed parameters in which physiological perception is identical to the experienced environment (Staiger et al.1997).

The meteorological variable inputs to PT are air temperature, dew point temperature, wind velocity, total cloud cover, and cloud cover of low, medium and high-level clouds. Because of our lack of specific cloud height information, we developed a parameterization scheme that related surface dew point depression and total cloud amount to the cloud amount at different levels. Further, we performed a sensitivity analysis that demonstrated that the value of PT was much more sensitive to total cloud amount (an observed value that we had available) than to how the clouds were divided between the various levels (values that we derived). Therefore, our determination of PT should be mostly unaffected by the limitations of the data that we had available.

Physiological Equivalent Temperature (PET)

Physiological Equivalent Temperature equates the heat balance of the body in the actual environment to that which is experienced indoors under light activity. PET is the temperature value in degrees Celsius of the indoor environment when the heat balances are identical (Höppe and Mayer 1987; Höppe 1999). PET is calculated from the mean radiative temperature, air temperature, air velocity, and water vapor pressure. Höppe **(**1999) suggests that a benefit of the PET is that it enables a layperson to make judgments about climate based on personal experience because it is reported in degrees Celsius, whereas other indices do not report a temperature-based value.

Predicted Mean Vote (PMV)

Predicted Mean Vote quantifies discomfort based on human-assessed response to physiological stresses. The index was developed by Fanger (1970) and is considered to be one of the most widely used comfort indices today (McGregor et al., 2002). In its simplest form, it may be considered a product of a function of partial vapor pressure, dry bulb air temperature, mean radiant temperature, and clothing surface temperature times the exponential of metabolic rate (McGregor et al., 2002).

Relative Strain (RS)

Lee (1979) presents the Belding and Hatch Relative Heat Strain (referred to here as Relative Strain) as an improvement of the former Heat Strain Index. The HSI, in its time, marked a considerable achievement in quantifying the human body's reaction to heat in that it included several important environmental variables as well as the metabolic rate, was based on the physics of heat exchange, and took a relatively simple computational form (Lee 1979). However, the HSI lacked consideration for the resistance of clothing to the loss of both sensible heat and evaporated water vapor, which were included in the refined RS model as calculated by Burton (Lee 1979). RS is “relative” as it is based on standard values for a person performing an established amount of work, with a given rate of ambient air movement. The revision allows the calculation of an RS value for any combination of air temperature, humidity, air movement, activity, radiation load, and clothing insulation (Lee, 1979).

Spatial Synoptic Classification (SSC)

The Spatial Synoptic Classification is a site-specific daily discretization of multivariate weather input variables observed diurnally. The resulting classification, developed using discriminant analysis and modified by user knowledge, identifies six primary synoptic weather types described primarily by temperature and moisture, and an additional transition type. In some cases, such as Moist Tropical (warm, moist), additional subcategories are determined to identify extreme days.

Prior research (Kalkstein and Greene 1997) has identified the weather situations primarily linked to high mortality in our four study cities. The so-called offensive weather types are always associated with high temperatures; Moist Tropical plus (very humid and warm air) and Dry Tropical (warm, dry air). Both synoptic types are deemed offensive in all three cities except for Baltimore, where Dry Tropical is the only one linked to high mortality.

Standard Effective Temperature (SET)

The principles behind the calculation of SET are somewhat similar to those of the PMV. Principally used by the American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc. (ASHRAE), the SET Index compares individual physiological comfort to a reference environment. The reference environment has a temperature equal to the mean radiant temperature in the ambient environment, wind velocity of zero, and is located at sea level (Ye et al., 2003).

In addition to these comfort variables, air temperature (T) and dew point temperature (DT) are included to provide low complexity measures for comparison to the more complex indices. Our list of indices is not meant to be exhaustive but was chosen to reflect a portion of the diversity of comfort measures available. AT was calculated according to procedures outlined by Davis et al. (2003). SSC was acquired from a web site monitored by Scott Sheridan at Kent State University (Sheridan, 2002). PMV, PET, and SET were calculated using Matzarazkis's (2000) RayMan model, version 1.2. Perceived temperature was derived using FORTRAN code for the MEMI model (Jendritzky 1990) kindly provided by Gerd Jendritzky. When needed, human input parameters chosen were a 1.75 m, 75 kg, 35-year-old male wearing clothing rated at 0.9 clo who is engaged in activity generating 80 W.

4. HEAT WAVE DETERMINATION

We begin by defining the sample of "heat events" for each city. "Heat events" are defined as all periods of two or more consecutive days, ending between April 1 and September 30, in which the daily 0700 LST temperature departure exceeds the centered 30-day running mean. In Philadelphia this procedure identified 555 heat events in the 29-year record. Next, we characterize anomalous mortality conditions. To do this, we define "extreme mortality events" when the total mortality for any given heat event exceeds the mean (calculated over all heat events) by at least one standard deviation. Because of declining mortality over time and a related trend in declining variance, we calculate this standard deviation threshold within a moving, five-year window. Similarly, for each index except SSC, we calculate a "heat wave" using 1600 LST values of the comfort index under consideration. If the maximum value of the index on any day within the heat event exceeds the one standard deviation threshold over a moving 5-year window, then that event is characterized as a "heat wave." In other words, a "heat wave" is a two or more day period of above normal temperatures that contains at least one unusually high value of the comfort index under consideration.

Heat waves are considered in two different contexts in this study. An *absolute heat wave* is defined as a heat wave that is determined using the actual values of each comfort index (as described above), while a *relative heat wave* is defined as a heat wave that is determined using anomalies of each index calculated as departures from a centered, 30-day moving average. For example, a series of days with above average temperatures in April might not qualify as an absolute heat wave but could be considered a relative heat wave if it is significantly warmer than the days preceding or following it. A combination of these two categories is also examined by taking the union of the absolute and relative heat waves.

For example, Figure 1 is a time series of AT in 1977 at Philadelphia. Using an absolute threshold, a July heat wave is evident (Figure 1a), but use of a relative threshold also identifies several additional heat waves in April (Figure 1b) that did not also exceed the absolute threshold criterion. In this example, the July heat wave qualifies as both relative and absolute.

The use of two thresholds, one for total mortality and the other for the comfort index, effectively divides all of the heat events into four categories—a high comfort index and high mortality, or a "killer heat wave"; a high comfort index and normal (or low) mortality, or a "non-killer heat wave"; a normal (or low) comfort index and high mortality; or a normal (or low) comfort index and normal (or low) mortality.

Figure 2a shows an example of this division as determined using absolute AT in Philadelphia over our 29-year period of record. Note the moving AT threshold (dashed-red line in Figure 2a). Alternatively, we can plot AT vs. total mortality for the same data (Figure 2b). The four categories can also be entered into a contingency table (Figure 2c) which assumes threshold exceedence can be used as a "forecast" of high mortality. In the contingency tables, Box a, high AT and high mortality, our "killer heat waves" represents a "hit" (n=21); Box b, high AT and below normal mortality, our "non-killer heat waves" is a "miss" (n=57); Box c, low AT and high mortality, is also a "miss" (n=46); and Box d, low AT and low mortality, is another type of "hit" (n=431).

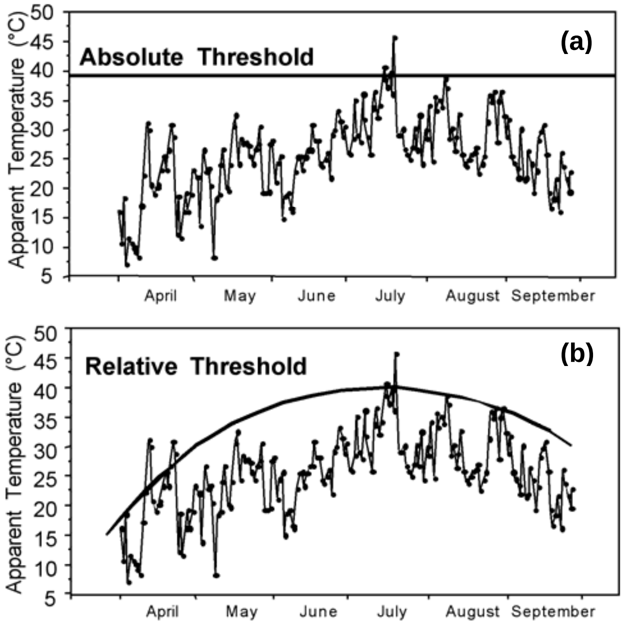
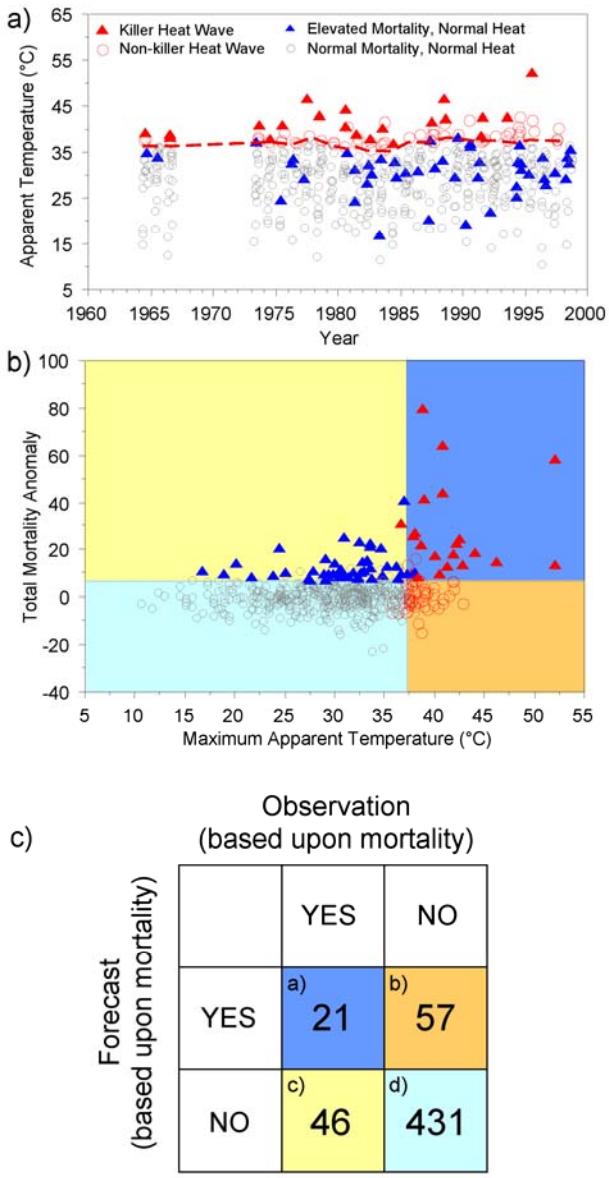


Figure 1 (a): Example of heat waves based on apparent temperature (AT) in Philadelphia, 1977, determined by applying an absolute threshold AT and (b) a relative threshold AT.



	YES	NO
YES	a) 21	b) 57
NO	c) 46	d) 431

Figure 2 (a) Categories of heat events by year for Philadelphia using apparent temperature as the comfort index; (b) categories of heat events vs. total mortality – note that because the thresholds vary over time, they are not accurately represented by a fixed value; (c) contingency tables built from same data.

To compare indices, we calculate a variety of accuracy and skill measures from the contingency tables, including the Hit Rate (H=((a+d)/N, where N=sample size), Probability of Detection (POD=a/(a+c)), False Alarm Rate (FAR=b/(a+b)), Bias (B=(a+b)/(a+c)) and Yule's Skill Score (Q= (a*d-b*c)/(a*d+b*c)) (e.g., Wilks, 1995). For the SSC, the weather types previously linked to high mortality for each city were used to determine hits and misses. In all MSAs examined here, the critical synoptic types (indicating a hit) were hot and very moist (Maritime Tropical Plus) and/or hot and dry (Dry Tropical). A SSC "heat wave" required that at least one day of the critical synoptic type occurred during a "heat event."

A chi-square test for goodness-of-fit is used to compare the number of killer heat waves that were deemed hits by each index to a random forecast of hits. The following calculation is made to determine the random number of hits that would be expected: ((a+b)/n)*((a+c)/n)*n (please refer to figure 2c). Then, for each city and heat wave type, we use a chi-square test to compare the number of hits determined by each index to this random number of hits at $\alpha \leq 0.05$.

A second chi-square test, also using $\alpha \leq 0.05$, is then carried out to determine if the number of hits identified by one index is statistically significantly different from the number of hits determined by the other indices. For example, for Philadelphia relative heat waves, the number of hits predicted by AT is compared to the number of hits predicted by each of the other indices. If the number of hits determined by AT is significantly greater than the number of hits determined by DT, for example, then AT receives a "win" and DT receives a "loss." If the number of hits determined by AT is not statistically significantly different than the number of hits determined by DT, then both AT and DT receive a "tie." These results can be tabulated in a standings format with a calculated winning percentage (a tie equals half a win) so the indices can be directly compared within or between cities for relative, absolute, and combined heat waves.

5. RESULTS

Statistical Skill Measures

In general, Table 1 shows no clear pattern among the comfort indices with respect to our accuracy and skill measures, including the simple measure of T. Dew point temperature (DT) did not perform as well overall in these statistical tests: this result is expected because a strong correlation between moisture and mortality has not been established. As we assume DT will perform poorly, this occurrence is not noted in our subsequent discussion.

Hit Rate

The hit rate (H) ranges from zero to one with a maximum hit rate of one. Hit rates vary from roughly 0.72 to 0.84 but are driven by the high number of weak heat events with below normal mortality and a low comfort index value (cell d in the contingency table), which is, by design, the most likely result. The indices generally vary together by city and classification method with no discernable pattern.

Probability of Detection

The probability of detection (POD) varies from zero to one, with a POD=1 indicating that all high mortality events are also associated with a high index value. There is an observable difference in the POD between cities; for example, POD values in Boston are generally higher than those for the other three cities, whereas Baltimore has the lowest POD. Dew point temperature has a consistently low POD, but the other indices tend to be comparable. POD is always higher for combined heat waves as compared to absolute and relative heat waves.

False Alarm Rate

False alarm rate (FAR) has a low value when most of the high mortality forecasts are also high mortality days. To simplify reading the table, we show the values for 1–FAR, so higher values indicate a good result, as with the previous accuracy measures. The relatively low values of 1-FAR indicate that many hot, humid days occur that are not associated with high mortality. However, 1-FAR is a useful indicator of the relative ability of each comfort index to identify killer heat waves. In general, killer heat waves are easiest to detect in Boston and most difficult to identify in Baltimore. Relative heat waves are harder to identify than absolute heat waves, and it follows that combined heat waves typically fall between these two values. With respect to index differences, SSC has the highest value of 1-FAR in all three cities except for Philadelphia, where RS and SET perform better.

Bias

The bias (B) = 1 for unbiased forecasts, is less than one for underforecasts, and exceeds one for overforecasts. Bias is generally positive for most cities, heat-wave types, and indices, indicating that mortality is often not high despite uncomfortable conditions. This is particularly true for the combined heat wave category, where in some cases the bias exceeds two.

Forecast Skill (Yule's Skill Score)

The forecast ability of a predictor increases as the Yules's Skill Score (Q) increases from zero to one. Skill varies from city to city, with highest skill values in Boston and lowest skill in Baltimore. Relative heat waves are more difficult to forecast, in almost all cases, than absolute or combined heat waves. SSC has the highest skill in Baltimore and Minneapolis, but other indices, PMV, RS, and SET, are preferable in Boston and Philadelphia.

Chi Square Analysis

Comparisons with Random Predictor

The number of positive hits (high mortality/high comfort index) identified using dew point temperature was not statistically significantly greater than the hits predicted randomly in Baltimore and Minneapolis for absolute heat waves, all four cities for relative heat waves, and Baltimore and Minneapolis for the combined category (Table 2). Six additional comfort indices showed no statistical improvement in predicting the number of hits for relative heat waves in Baltimore, and SET also showed no improvement over a random predictor for relative heat waves in Minneapolis. This suite of comfort indices is inefficient at predicting mortality during relative heat waves in Baltimore. As previously explained, there is no established relationship between moisture alone and mortality, and dew point temperature alone is not distinguishable from a random predictor.

Index Comparisons

Table 3 shows the “standings” for the three types of heat waves compiled, by comfort index, using a chi-square test to determine if the number of hits for one index is statistically significantly different from that of the other indices. The winning percentage is the average across the four cities for each comfort index.

SSC is ranked highest for absolute heat waves followed by PET (Table 3a). SSC is also ranked highest for relative heat waves followed by RS, as might be expected because these two indices explicitly consider relative meteorological conditions (Table 3b). In the combined category, PET and PMV tie for the highest ranking (Table 3c). The failure of SSC to perform well for combined heat waves is discussed later.

The overall winning percentage (the mean of the three heat wave categories) as well as the combined won/loss record for each index is shown in Table 4. PET performed the best overall when compared to the other indices and DT ranked last. It is of interest that T performed better than PT and SET, an observation that will be elaborated upon in the discussion.

6. DISCUSSION

Overall, the differences between comfort indices (other than dew point temperatures) are relatively small. There are clear inter-city differences, but in general a city with poor forecast skill, for example, will express this difficulty across all indices. It is noteworthy that the simple use of air temperature alone provides an estimate of killer heat waves that is comparable, and in some cases superior to, several complex biometeorological indices. This finding alone should encourage the development of simple, temperature-based models as a reasonable first estimate of the impacts of heat on mortality. However, there also appears to be some marginal utility gained from the addition of other parameters to a temperature model, especially some form of humidity measure.

Most measures are not skillful at identifying all three types of heat waves. Several of the relative measures, like relative strain and SSC, as expected, perform better in relative heat waves than do indices based on absolute conditions. The SSC is essentially a relative index, as it is possible to identify dry, polar air in Philadelphia in January and July even though it will have markedly different properties in different seasons. The SSC approach does have the advantage of identifying two air masses, dry tropical and moist tropical plus, that always occur at the high end of the temperature or temperature/humidity spectrum for both relative and absolute conditions. Thus, the SSC approach achieves better forecast skill than most other indices which, unlike the SSC, are explicitly designed to model human comfort.

Table 1		**Baltimore**			**Boston**			**Minneapolis**			**Philadelphia**		
Stat	CI	Abs.	Rel.	Comb	Abs.	Rel.	Comb	Abs.	Rel.	Comb	Abs.	Rel.	Comb
H	T	0.817	0.783	0.728	0.844	0.824	0.792	0.820	0.797	0.772	0.845	0.802	0.766
	DT	**0.835**	0.757	0.718	0.806	0.758	0.698	0.779	0.765	0.706	0.838	0.775	0.735
	AT	0.801	0.771	0.726	0.829	0.822	0.783	0.813	0.784	0.761	0.814	0.813	0.755
	PT	0.806	0.792	0.723	0.826	0.813	0.785	0.795	0.777	0.743	0.820	0.791	0.746
	RS	0.794	0.778	0.737	0.804	0.82	0.779	0.783	0.790	0.765	0.834	0.841	**0.823**
	PMV	0.812	0.789	0.726	**0.856**	0.822	0.806	**0.827**	0.793	0.763	**0.850**	0.798	0.769
	PET	0.813	0.79	0.737	0.84	0.842	0.804	0.820	0.788	0.770	0.840	**0.820**	0.780
	SET	0.821	0.783	0.726	0.849	0.811	0.779	0.818	0.775	0.742	**0.850**	0.800	0.766
	SSC	0.815	**0.815**	**0.815**	0.853	**0.853**	**0.853**	0.825	**0.825**	**0.825**	0.805	0.805	0.805
POD	T	0.254	0.149	0.373	0.465	0.352	0.62	0.304	0.266	0.494	0.299	0.269	0.507
	DT	0.09	0.149	0.239	0.211	0.169	0.324	0.152	0.190	0.329	0.209	0.164	0.358
	AT	0.254	0.209	0.388	0.451	0.423	0.62	**0.354**	0.329	0.468	0.313	0.343	0.478
	PT	0.209	0.209	0.373	0.437	0.338	0.606	0.316	0.266	0.443	0.299	0.239	0.493
	RS	0.269	0.254	0.373	0.394	0.479	0.577	0.291	0.304	0.392	0.313	0.343	0.448
	PMV	0.224	0.209	0.388	0.493	0.338	**0.662**	0.329	0.291	**0.532**	0.358	0.269	**0.552**
	PET	0.299	0.194	**0.418**	**0.535**	**0.507**	**0.662**	0.38	0.266	0.519	0.388	0.328	**0.552**
	SET	0.269	0.194	**0.418**	0.465	0.310	0.634	0.304	0.228	0.481	0.373	0.224	0.522
	SSC	**0.358**	**0.358**	0.358	0.465	0.465	0.465	0.354	**0.354**	0.354	**0.403**	**0.403**	0.403
1 - FAR	T	0.243	0.133	0.184	0.402	0.325	0.331	0.343	0.273	0.307	0.339	0.228	0.26
	DT	0.158	0.111	0.129	0.224	0.136	0.161	0.174	0.181	0.188	0.275	0.137	0.187
	AT	0.215	0.156	0.187	0.364	0.341	0.319	0.341	0.277	0.287	0.269	0.277	0.241
	PT	0.200	0.179	0.180	0.352	0.296	0.319	0.291	0.239	0.259	0.274	0.198	0.236
	RS	0.212	0.185	0.191	0.298	0.351	0.306	0.258	0.276	0.270	0.313	**0.343**	0.33
	PMV	0.217	0.175	0.187	**0.443**	0.316	0.359	0.371	0.277	0.304	0.375	0.222	0.274
	PET	0.256	0.169	0.204	0.404	0.404	0.356	0.366	0.256	0.311	0.351	0.286	**0.287**
	SET	0.257	0.16	0.196	0.418	0.282	0.317	0.338	0.217	0.268	**0.379**	0.203	0.263
	SSC	**0.282**	**0.282**	**0.282**	0.429	**0.429**	**0.429**	**0.373**	**0.373**	**0.373**	0.284	0.284	0.284
B	T	1.045	**1.119**	2.03	1.155	1.085	1.873	0.886	0.975	1.608	0.881	1.179	1.955
	DT	0.567	1.343	1.851	0.944	1.239	2.014	0.873	**1.051**	1.747	0.761	1.194	1.91
	AT	1.179	1.343	2.075	1.239	1.239	1.944	**1.038**	1.190	1.633	1.164	1.239	1.985
	PT	1.045	1.164	2.075	1.239	1.141	1.901	1.089	1.114	1.709	1.09	1.209	2.09
	RS	1.269	1.373	1.955	1.324	1.366	1.887	1.127	1.101	**1.456**	**1.000**	**1.000**	**1.358**
	PMV	**1.030**	1.194	2.075	1.113	**1.070**	1.845	0.886	**1.051**	1.747	0.955	1.209	2.015
	PET	1.164	1.149	2.045	1.324	1.254	1.859	**1.038**	1.038	1.671	1.104	1.149	1.925
	SET	1.045	1.209	2.134	1.113	1.099	2.000	0.899	1.051	1.797	0.985	1.104	1.985
	SSC	1.269	1.269	**1.269**	**1.085**	1.085	**1.085**	0.949	**0.949**	0.949	1.418	1.418	1.418
Q	T	0.479	0.075	0.347	0.771	0.639	0.758	0.611	0.467	0.627	0.661	0.440	0.612
	DT	0.176	-.046	0.06	0.382	0.046	0.187	0.144	0.176	0.237	0.526	0.088	0.347
	AT	0.408	0.187	0.365	0.726	0.687	0.743	0.626	0.498	0.578	0.551	0.577	0.556
	PT	0.35	0.281	0.332	0.707	0.587	0.736	0.523	0.383	0.505	0.555	0.342	0.551
	RS	0.403	0.312	0.373	0.611	0.721	0.705	0.443	0.487	0.507	0.629	**0.681**	**0.700**
	PMV	0.405	0.265	0.365	**0.814**	0.62	**0.807**	0.660	0.486	0.641	0.724	0.425	0.661
	PET	0.525	0.238	0.436	0.797	0.787	0.805	0.670	0.428	0.645	0.706	0.588	0.683
	SET	0.517	0.205	0.408	0.785	0.55	0.748	0.603	0.309	0.542	**0.733**	0.354	0.626
	SSC	**0.598**	**0.598**	**0.598**	0.794	**0.794**	0.794	**0.671**	**0.671**	**0.671**	0.613	0.613	0.613

Table 1 (previous page). Values of the hit rate (H), probability of detection (POD), one minus the false alarm rate (1— FAR), bias (B), and Yule's Skill Score (Q) for absolute (Abs.), relative (Rel.), and combined (Comb) heat waves for each city. The best comfort index for each city/heat wave type is indicated in bold for each variable.

Table 2 (below). Chi-square values for test of the observed number of killer heat wave "hits" against random expectation for each city, comfort index, and heat wave type. Significant differences are indicated in bold.

		BAL			BOS			MIN			PHI	
	Absolute	Relative	Comb.	Absolute	Relative	Comb.	Absolute	Relative	Comb.	Absolute	Relative	Comb.
T	**11.5**	0.13	**6.65**	**61.57**	**26.19**	**56.38**	**22.44**	**10.56**	**31.73**	**26.96**	**7.52**	**26.6**
DT	0.22	0.11	0.09	**4.58**	0.11	1.86	0.46	0.88	3.4	**11.72**	0.12	**6.96**
AT	**8.21**	0.98	**6.38**	**47.44**	**38.83**	**50.31**	**25.15**	**15.56**	**25.97**	**18.48**	**19.86**	**21.02**
PT	**5.11**	3.21	**5.04**	**43.03**	**22.81**	**52.28**	**16.61**	**7.96**	**17.74**	**15.53**	**4.23**	**20.18**
RS	**7.52**	**3.92**	**6.65**	**25.67**	**48.54**	**44.55**	**9.21**	**14.15**	**17.63**	**23.99**	**31.94**	**39.26**
PMV	**6.96**	1.88	**6.38**	**72.75**	**22.81**	**69.61**	**29.31**	**11.89**	**36.66**	**36.34**	**7.52**	**36.21**
PET	**15.53**	2.05	**11.82**	**67.79**	**67.23**	**69.61**	**31.84**	**7.96**	**33.54**	**37.09**	**21.69**	**36.21**
SET	**14.19**	1.06	**9.54**	**61.57**	**16.76**	**54.25**	**22.44**	3.54	**21.69**	**41.02**	**4.62**	**29.64**
SSC	**23.04**	**23.04**	**23.04**	**61.57**	**61.57**	**61.57**	**30.52**	**30.52**	**30.52**	**27.84**	**27.84**	**27.84**

Table 3. Average winning percentage (across cities) for each comfort index and heat have type (absolute, relative, and combination).

Table 3a

Absolute	
Index	**WP**
AM	0.626
PET	0.579
SET	0.563
AT	0.547
RS	0.547
PMV	0.547
T	0.532
PT	0.532
DT	0.031

Table 3b

Relative	
Index	**WP**
AM	0.782
RS	0.641
PET	0.594
AT	0.594
PMV	0.469
T	0.453
PT	0.407
SET	0.344
DT	0.188

Table 3c

Combined	
Index	**WP**
PET	0.625
PMV	0.625
SET	0.610
AT	0.594
T	0.594
RS	0.500
PT	0.500
AM	0.313
DT	0.063

Table 4. Overall won/loss/tie records and winning percentage for each comfort index tallied across all cities and heat wave types.

Index	**W**	**L**	**T**	**WP**
PET	20	1	75	0.599
AT	17	2	77	0.578
AM	28	14	54	0.573
RS	16	4	76	0.563
PMV	15	6	75	0.547
T	11	6	79	0.526
PT	9	8	79	0.505
SET	12	11	73	0.505
DT	0	78	18	0.094

It is useful to remember that high skill is not anticipated in forecasting killer heat waves. Many people die for reasons completely unrelated to the weather. Assuming these high death days are randomly distributed both above and below the comfort index threshold, we would expect a low POD in general. Similarly, deaths do not typically rise linearly with increasing index value. Often, a heat wave requires some period of extreme conditions before a mortality signal becomes evident, and then it often occurs as a single or multiple-day spike followed by several days of below-normal mortality ("mortality displacement"). Therefore, the FAR is also expected to be fairly high (or 1-FAR is low).

This preliminary research tends to identify differences in the mortality response to discomfort in different cities. Although this four-city study is not extensive, prior research indicates that all four of these locations have "heat-sensitive" populations (e.g., Kalkstein and Davis, 1989; Davis et al., 2002, 2003). Even though we allow the threshold for each comfort index to vary between cities, there nevertheless exist what appear to be substantially different responses. This speaks to the need for location-specific models and strategies to identify how heat and humidity impact a given populace.

Our experimental design creates a problem in the comparison of the SSC to the comfort indices. Because the SSC identifies uncomfortable days based on one or two categorical variables, we cannot differentiate between absolute and relative heat waves. This leads to difficulties in comparisons under the combined classification, in which each of the other indices effectively has two chances at getting an elevated mortality "hit" where the SSC only has one chance. The result is that while SSC generally performs better than the other indices in predicting both absolute and relative killer heat waves when considered separately, it falls behind when these two heat wave types are combined into a single pool. To date, we have not devised a more appropriate comparison.

We made a number of subjective decisions in developing this analysis—other researchers may arrive at different conclusions using essentially the same data. For example, the variable we chose to quantify the magnitude of a heat wave, and thus determine its forecast of mortality, was the maximum daily index value observed during the heat wave. This determination could create some inconsistency in the classification. For example, using the PMV index, a five-day heat wave may have only one day with the PMV above the threshold value with the remainder of the days below the threshold. In our method, this series of days is considered a heat wave and is comparable to heat wave during which the PMV was above the threshold every day. An alternative approach would be to use the average daily index value as the representative value for the heat wave. Even more complex methods might be able to take into account the pattern of the index values. These differences in approaches are emblematic of the uncertainty of specifically how heat and humidity influence human mortality and the importance of prolonged vs. short but extreme exposure.

This research supports the hypothesis that all comfort indices have some ability to predict heat-induced human mortality. While statistical skill measures do not demonstrate distinct differences between the indices, preliminary research suggests that some indices may exhibit marginally better skill than others. Our study is the first comparative analysis of the human thermal comfort indices for U.S. heat wave and this is not comprehensive. But these preliminary results that show large differences in skill between cities suggests that the development of a single index of comfort may need to be adapted locally because of spatially varying relationships between human comfort and mortality.

7. REFERENCES

Anderson, R.N., Rosenberg, H.M., 1998: Age standardization of death rates: Implementation of the year 2000 standard. *National Vital Statistics Reports*. **47**, Hyattsville, MD, National Center for Health Statistics.

Davis, R.E., Knappenberger, P.C., Michaels, P.J., Novicoff, W. M., 2002: Decadal changes in summer mortality in U.S. cities. *Int. J. Biometeorol.*, **47**, 166-175.

Davis, R.E., Knappenberger, P.C., Michaels, P.J., Novicoff, W.M., 2003: Changing Heat-Related Mortality in the United States. *Environ. Health Perspect.*, **111**, 1712-1718.

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

Höppe, P., 1999: The physiological equivalent temperature universal index for the biometeorological assessment of the thermal environment. *Int. J. Biometeorol.*, **43**, 71-75.

Höppe P., Mayer, H., 1987: Thermal comfort of man in different urban environments. Theoretical and Applied Climatology., **38**, 43-49.

Jendritzky, G., Menz, G., Schmidt-Kessen, W., Schirmer, H., 1990: *Methodik zur raumlichen Bewertung der thermischen Milieus in der Stadt- und Landschaftsblanung (Klima-Michel-Modell). Beitrage der Akademie fur Raumforschung und Landesplanung*. Hermann Schroedel, Hannover.

Kalkstein, L. S., Davis, R. E., 1989: Weather and human mortality: an evaluation of demographic and interregional responses in the United States. *Ann. Assoc. Am. Geogr.*, **79**, 44-64.

Kalkstein, L.S., Greene, J. S., 1997: An evaluation of climate/mortality relationships in large U.S. cities and the possible impacts of a climate change. *Environ. Health Perspect.*,**105**, 84-93.

Kalkstein, L.S., Jamason, P. F., Green, J. S., Libby, J., Robinson, L., 1997: The Philadelphia Hot Weather-Health Watch/Warning System: Development and Application, Summer 1995. *Bull. Am. Meteorol. Soc.*, **77**,1519-1528.

Lee, D.H., 1979: Seventy-five Years of Searching for a Heat Index. *Environ. Res.*, **22**, 331-356.

Matzarakis, A, 2000: *RayMan*. Meteorological Institute of the University of Freisburg, Germany.

McGregor, G.R., Markou, M. T., Bartzokas, A., Katsoulis, B.D., 2002: An evaluation of the nature and timing of summer human thermal discomfort in Athens, Greece. *Clim. Res.,* **20**, 83-94.

NCHS, 1978: *Standardized micro-data tape transcripts*. Department of Health, Education, and Welfare, Washington, DC, 34 pp.

Sheridan, S.C., 2002: The redevelopment of a weather-type classification scheme for North America. *Int. J. Clim.*, **22**, 51-68.

Staiger, H., Bucher, K., Jendritzky, G., 1997: Gefühlte Temperatur. Die physiologisch gerechte Bewertung von Wärmebelastung und Kältestress beim Aufenthalt im Freien in der Maßzahl Grad Celsius. *Annalen der Meteorologie*, **33**, 100-107.

Steadman, R.G., 1979: The assessment of sultriness. Part I: A temperature-humidity index based on human physiology and clothing science. Journal of Applied Meteorology, **18**, 861-873.

Steadman, R.G., 1984: A Universal Scale of Apparent Temperature. Journal of Climate and Applied Meteorology, **23**, 1674-1687.

Ye, G., Yang, C., Chen, Y., Li,Y., 2003: A new approach for measuring predicted mean vote and standard effective temperature. *Building and Environment*, **38**, 33-44.

Wilks, D.S., 1995: Statistical Methods in Atmospheric Sciences. Academic Press, **Ch. 7**, 233-281.