Nat Hazards (2014) 74:771–788
DOI 10.1007/s11069-014-1213-5

ORIGINAL PAPER

The predictability of high-risk zones for heat-related mortality in seven US cities

David M. Hondula · Robert E. Davis

Received: 14 November 2013 / Accepted: 27 April 2014 / Published online: 24 May 2014

Abstract Heat-related mortality remains a public health challenge in the United States. The objective of this study was to determine the temporal consistency of high-risk zones for heat-related mortality using historical georeferenced mortality data from seven US cities. A generalized additive model was used to identify city-specific threshold temperatures associated with increased mortality, and then the mortality rate on threshold-exceeding days was calculated for each postal code comprising each study city. This process was iterated by withholding subsets of data from the model and assessing predictability via cross-validation. In all cities, the average mortality rate in postal codes targeted for intervention by the statistical model was higher than that in non-targeted areas. Targeted areas for interventions in the study data accounted for 50 % of excess heat-related deaths despite only accounting for 25 % of total mortality. Focusing intervention measures at certain geographical zones within urban areas could be an effective means of combating heat-related mortality because there is temporal consistency in places where the death rate is most sensitive to heat.

Keywords Heat · Mortality · Spatial · Prediction · Urban

D. M. Hondula (✉)
Center for Policy Informatics, Arizona State University, 411 N. Central Avenue, Suite 400, Phoenix, AZ 85004, USA
e-mail: david.hondula@asu.edu

D. M. Hondula
School of Geographical Sciences and Urban Planning, Arizona State University, P.O. Box 875302, Tempe, AZ 85287, USA

R. E. Davis
Department of Environmental Sciences, University of Virginia, 291 McCormick Road, Charlottesville, VA 22904, USA

1 Introduction

The human response to high heat and humidity varies from individual to individual, but over large populations there is clear evidence that extreme summertime conditions lead to elevated mortality and morbidity rates (Davis et al. 2003; Sheridan et al. 2009; Guo et al. 2012). This information has contributed to the adoption of a number of measures aimed at protecting public health when dangerous conditions are forecast, including heat-health watch-warning systems (Hondula et al. 2013b). There is now a growing capacity to improve specificity in assessments of heat-related risk across both time and space. Recent research has contributed to increased awareness of geographic variability in the response to heat within large metropolitan areas (e.g., Vaneckova et al. 2010; Hondula et al. 2012) and of intra-annual variability in the metropolitan aggregate-scale response to heat (e.g., Guo et al. 2012). But little effort has been directed at the intersection of these sources of variability—exploring how the intra-city human response to heat, in particular morbidity and mortality, changes from one year to another.

Within metropolitan areas, the heat-health risk is expected to vary spatially as the urban heat island and related microclimates create different levels of exposure for residents of different communities. Similarly, the underlying risk of the population varies as the demographic profile of city residents is spatially heterogeneous. Heat vulnerability indices have been proposed and mapped at fine spatial resolutions to highlight geographic variability in expected risk arising from factors that are believed to impact heat sensitivity (e.g., Reid et al. 2009). Other research has mapped health outcomes to identify regions within cities that have historically been associated with higher rates of heat-related mortality and then deduced the association with certain variability factors from the observations (Smargiassi et al. 2009; Hondula et al. 2012; Laaidi et al. 2012; Reid et al. 2012).

In addition to exploring spatial variability in intra-city risk, it is also important to consider the temporal dynamics in the response. It is known that the heat-mortality response varies across time (e.g., Davis et al. 2003; Rocklöv et al. 2009; Guo et al. 2012). Although considerable debate continues regarding future climate change and population adaptation, the increase in temperatures in recent years combined with urban heat island impacts has resulted in individuals being more frequently exposed to thermally stressful conditions in some locations (Arnfield 2003; Wilby 2003). At the same time, infrastructure improvements including the use of air conditioning and building design might be decreasing the risk for some (Davis et al. 2003; Sheridan et al. 2009). Aside from these long-term trends, there is evidence that warm-season mortality might be influenced by antecedent wintertime conditions, and further evidence suggests that the characteristics of individual heat events including their timing, duration, and intensity, impact the mortality rate (Rocklöv et al. 2009; Anderson and Bell 2011). In the light of these sources of variability at different time scales, it is possible that specific populations within cities most affected by heat might vary based on heat event characteristics, time in season, or longer-term trends. From a geographical perspective, then, the *location* of high-risk zones (e.g., Reid et al. 2009; Hondula et al. 2012) may not be consistent over time.

An examination of temporal variability in the intra-city response to heat is warranted because of the insights such a study could provide regarding the effectiveness of various intervention strategies, particularly those where resources are directed toward particular locations. Here, we investigate the predictability of high-risk zones for heat-related mortality using multi-decadal geo-coded mortality records from seven climatically diverse cities in the United States. The goal of the study is to determine whether spatial patterns in heat-related mortality are non-random temporally.

2 Data

We obtained mortality records from seven urban areas across the United States: Atlanta, Georgia (period of record 1994–2007, abbreviation ATL), Boston, Massachusetts (1987–2007, BOS), Minneapolis, Minnesota (1992–2008, MSP), Philadelphia, Pennsylvania (1983–2008, PHL), Phoenix, Arizona (1989–2007, PHX), Seattle, Washington (1988–2008, SEA), and St. Louis, Missouri (1980–2008, STL). The mortality records contained the postal code of residence of each decedent, enabling us to determine the daily number of deaths within each postal code over the entire period of record. As data were sought from individual states, complete records for entire metropolitan statistical areas (MSAs) were not available. Instead, contiguous postal codes were selected such that each study city would have relatively comparable data sets. The average number of daily deaths within each city ranged from 30 to 53, and each city contained between 48 and 101 postal codes. Postal codes were selected as the spatial unit for this study because they were the smallest subdivision available across all cities. Descriptive statistics of the mortality data sets used are available in Table 1. These data were obtained from the respective state departments of health. Because the data were de-identified, no IRB approval was necessary for this research under Title 45 Part 46 exemption category 4.

Hourly meteorological data were obtained from the first-order weather stations closest to the study cities through the U.S. National Climatic Data Center. These data are publicly available online (www.ncdc.noaa.gov). We calculated the daily maximum afternoon temperature, defined as the maximum temperature observed in the 5-h window centered at the hour of the average daily maximum temperature. There were 43 days (<0.1 % of available data) where no data were available from this time window; these days were given a missing value for temperature.

Geographic boundary files for the study cities were sourced from the United States Census Bureau accessed through the National Historical Geographical Information System portal online (www.nhgis.org) (Minnesota Population Center 2011). We downloaded year-2000 Zip Code Tabulation Areas (ZCTAs, henceforth, "postal codes") for the states comprising the study cities and extracted the boundaries corresponding to the postal codes available from the mortality records.

3 Methods

3.1 Model training and threshold temperatures

We adopted a cross-validation framework to examine the predictability of mortality within postal codes in each of the study cities. The general procedure was to identify target zones within each city using all years of data except a single withheld year (y_w), and then evaluate how mortality in the target zones compared to that in the withheld year. This procedure was iterated for all years y in each city's period of record and across all postal codes.

In the first iteration, we calculated the relationship between temperature and mortality while accounting for seasonality and long-term time trends using a generalized additive model of the form:

$$\mathrm{Log}[E(M)] = \alpha + s(T_{PM}, df = 5) + s(Time, df = 7 * n(y)) \quad (1)$$

where $E(M)$ is the predicted daily mortality count, α is the model intercept, T_{PM} represents

Table 1 Descriptive statistics of the meteorological and mortality data sets used in the study for six different United States cities

City	Boston, MA	Minneapolis-St. Paul, MN	Philadelphia, PA	Phoenix, AZ	Saint Louis, MO	Seattle, WA
Period of record	1987–2007	1992–2008	1983–2008	1989–2007	1980–2008	1988–2008
Summer mean temperature in training periods (Min–Max)	25.5 (25.4–25.6)	26.4 (26.3–26.6)	29.1 (29.0–29.1)	39.9 (39.9–40.0)	30.3 (30.2–30.4)	22.7 (22.6–22.8)
Threshold temperature (Min–Max)	28.0 (27.8–28.2)	30.2 (30.0–30.6)	30.7 (30.6–30.8)	42.6 (42.4–43.0)	34.4 (33.8–35.8)	25.9 (25.7–26.1)
Model-predicted mortality increase at threshold (95 % CI)	1.018 (1.000–1.036)	1.023 (1.000–1.046)	1.015 (1.000–1.031)	1.019 (1.020–1.040)	1.016 (1.000–1.033)	1.021 (1.000–1.042)
Mean mortality on all above-threshold days relative to below-threshold days (Min–Max)	1.059 (1.056–1.063)	1.046 (1.041–1.050)	1.066 (1.060–1.070)	1.040 (1.035–1.055)	1.032 (1.026–1.046)	1.063 (1.059–1.067)
Number of postal codes	64	101	48	101	63	65
Mean number of target postal codes (Min–Max)	15.0 (13-16)	13.5 (10–15)	18.9 (16–21)	11.9 (10–15)	7.7 (4–11)	11.3 (10–14)

The table also includes information regarding the modeled temperature–mortality relationship for each city, the modeled threshold temperatures (°C), and the number of spatial units identified as targets during model training iterations. The minimum and maximum values shown represent the range across different cross-validation samples daily afternoon maximum temperature, *Time* is a variable to account for long-term time trends and seasonality, $n(y)$ is the number of years in the time series (excluding the withheld year y_w), and $s()$ represents thin plate smoothing splines for the temperature and time terms (Wood 2003, 2006). Five degrees of freedom was set for the temperature term and seven per year for the time term for all cities. We tested a number of other options for degrees of freedom for each variable made the decision to use five and seven, respectively, based on generalized cross-validation scores and visual examination of plots of model-predicted values for each term. All values for the withheld year were set to a missing value indicator. We then extracted model-predicted values and partial error for the temperature component of the model.

As we were interested in examining mortality when summertime weather deviates from normal conditions, the analysis was conducted relative to the mean summer (June–August inclusive) afternoon maximum temperature. We then examined the model-predicted values and standard errors and determined the first temperature higher than the summer mean at which the confidence interval of the relative risk (RR) of all-cause mortality between a given temperature T_x and the mean temperature T_m did not include one. The RR between a given temperature and the mean was calculated as:

$$\mathrm{RR}_{T_x} = e^{((\widehat{M}(T_x) - \widehat{M}(T_m)))} \tag{2}$$

where $\widehat{M}$ is the model-predicted mortality at a given temperature T_x or the summer mean temperature T_m. The 95 % credible interval of the RR was calculated as:

$$95\,\%\ \mathrm{CI}\ (\mathrm{RR}_{T_x}) = e^{((\widehat{M}(T_x) - \widehat{M}(T_m)) \pm 1.96 \times \sqrt{\varepsilon(T_x)^2 + \varepsilon(T_m)^2})} \tag{3}$$

where ε is the partial model error in the temperature term at a given temperature. Thus, the temperature threshold is defined here as the temperature at which model-predicted mortality significantly differs from that expected for mean summertime conditions. We then calculated the average mortality rate on all above-threshold days.

As smoothing spline estimates of seasonality derived directly from small samples of data can be unreliable, we examined the relationship between temperature and mortality within each postal code using a multi-stage model (Hondula et al. 2013a). In the first stage, seasonality was removed from the citywide data using a generalized additive model that includes terms for seasonality, long-term trends, and temperature. The seasonality and long-term time trends in the mortality data represent a mortality "baseline"—long-term variability that is (presumably) unrelated to short-term variability in temperature. This baseline curve was then adjusted to match the changing mortality rates within each postal code based on population growth and demographic changes. A daily mortality ratio (MR) was calculated for each postal code by calculating the ratio between each day's sum of deaths within the postal code and the expected number of deaths based on the mortality baseline. We then group the postal code-level ratios into different "bins" based on temperature thresholds (as defined above).

This method was adapted to the current investigation by omitting the withheld year. To ensure that intra-annual, rather than inter-annual differences in mortality rates were emphasized in the analysis, an additional standardization was made to set the mean summertime MR equal to exactly one for each year in the period of record. No adjustment was made to the variance. We obtained estimates of within-postal code mortality for two separate groups of days: those with temperatures above and below the citywide estimated threshold. We then used a randomization test (Sheridan and Dolney 2003; Hondula et al.

2013a) to identify postal codes where mortality was statistically significantly high on above-threshold days. These postal codes were labeled as "targets" for intervention activities.

3.2 Model testing

The target zones were determined using all years except the one withheld year, excluded from the training model to assess *predictability*. We quantified predictability by comparing the mortality in the target zones in the withheld year to that in the non-target zones. In general, if mortality was higher in the target zones than the non-targets in the withheld year, we concluded that there was predictability in the spatial response. As each postal code year is mean-standardized, there was no bias introduced from the methods that would cause a postal code identified as a target to be more likely to have high or low mortality in the testing year.

With only one year of withheld data and small sample sizes within each postal code, there were not enough data to generate a robust statistical model of within-postal code mortality for the withheld year analogously to what we have done for the rest of the time period. Instead, we calculated seasonality for the entire city in the withheld year using the model:

$$\mathrm{Log}[E(M)] = \alpha + s(Time, df = 7). \tag{4}$$

Temperature was not included in the model for the withheld year because effect estimates based on one year of data were found to be unstable (i.e., the magnitude, significance, and sign of the effect varied from one year to another). Although there is some concern that the time component of this model was incorrectly capturing temperature-related variability, in the cases we examined that there was no visual evidence of short-term deviations away from seasonality that might indicate over fitting.

Seasonality, as determined from Eq. 4, was removed from the postal code-level data by scaling the seasonal curve so that the mean mortality rate matched that within each postal code. We then calculated a mortality ratio MR for each day within each postal code for the withheld year y_{w} by dividing the observed mortality count by the scaled seasonality. As before, we then re-standardized, so the mean summertime *MR* in the withheld period is equal to one. A postal code was labeled as "high mortality" during the withheld year if the mean *MR* on above-threshold days exceeded the critical value for significance as determined from the randomization test described above.

3.3 Model evaluation

This entire procedure was repeated for each year y_j in the data set. Thus, for each year, we identified a set of target postal codes based on $(n(y) - 1)$ years of data and those postal codes that were associated with high mortality rates on above-threshold days in year j. We used 2×2 contingency tables to quantify the performance of the models in identifying postal codes with high mortality. For each testing iteration, we tabulated the number of postal codes with high and low mortality (based on the single withheld year) and the number of target and non-target postal codes (based on the training data set, excluding the single withheld year). Cross-tabulation of these quantities yields four distinct possibilities: (A) target zones with high mortality in the withheld year, (B) non-target zones with high mortality in the withheld year, (C) target zones without high mortality in the withheld year, and (D) non-target zones without high mortality in the withheld year (e.g., Table 2). Cells

Table 2 Contingency tables comparing the relative frequency at which postal codes were identified as targets and non-targets during model training and associated with high mortality during model testing

	BOS			MSP	
	Training sample			Training sample	
	Target	Non-targets		Target	Non-targets
Withheld year			Withheld year		
High	158	373	High	106	459
Not high	158	655	Not high	124	1,028
Relative risk	1.378 (1.202, 1.580)		Relative risk	1.493 (1.273, 1.751)	
	PHL			PHX	
	Training sample			Training sample	
	Target	Non-targets		Target	Non-targets
Withheld year			Withheld year		
High	228	288	High	82	565
Not high	264	468	Not high	145	1,127
Relative risk	1.216 (1.067, 1.387)		Relative risk	1.082 (0.899, 1.302)	
	STL			SEA	
	Training sample			Training sample	
	Target	Non-targets		Target	Non-targets
Withheld year			Withheld year		
High	79	571	High	108	434
Not high	140	1,037	Not high	129	694
Relative risk	1.016 (0.842, 1.226)		Relative risk	1.184 (1.012, 1.386)	

The data shown are summed across all years in the study period for each respective city. Relative risk values >1.0 indicate that postal codes identified as targets are more likely to be associated with high mortality in model testing compared to non-targets

"A" and "D" in the contingency table are associated with correct forecasts, "B" and "C" with missed forecasts. We then calculated the RR of a postal code being associated with high mortality in the withheld year if it was classified as a target zone compared to the risk of a postal code being associated with high mortality if it was classified as a non-target using the equation:

$$\mathrm{RR} = \frac{A/(A+C)}{B/(B+D)} \tag{5}$$

where the letters A–D refer to specific cells in the 2×2 contingency table (Table 2). A $\mathrm{RR} > 1.0$ indicated that target zones were more likely to be associated with high mortality in the withheld years. We determined that high mortality zones within each city were statistically predictable whether the lower bound of the confidence interval for RR (Gardner and Altman 1989) was >1.0.

Finally, we estimated the overall mortality burden attributable to heat for each city over the entire time period as well as for target postal codes exclusively during years they were identified as targets. "Excess mortality" was defined as the difference between the observed and expected number of deaths on a given day or set of days. We calculated the excess mortality ε for each postal code *i* within each year *j* on above-threshold days using the equation:

$$\varepsilon_{i,j} = M_{i,j} \times n(T_{PM} \geq T^{*})_{j} \times (\mu_{i,j} - 1) \tag{6}$$

where M is the mean daily summertime mortality specified for each postal code year, T^{*} is the city-specific threshold temperature for each respective year, n is a count of the number of days the threshold temperature is reached, and μ is the average mortality ratio MR for each postal code and year on above-threshold days. Because of the procedures discussed above to derive MR, μ represents average mortality ratios on above-threshold days standardized for seasonality and long-term time trends (see Eq. 4 and subsequent text). Summation of ε across all postal codes and years yielded the total excess mortality on above-threshold days for each city in aggregate.

We next sought to determine the portion of excess mortality that occurred in target zones. We repeated the calculation in Eq. 6 for observed mortality in target zones (ε^{t}) only using observations from each year when a postal code was identified as a target. The percentage of excess mortality occurring in the target zones is then determined by multiplying the ratio of ε^{t} (summed across all postal codes and years) to ε (summed across all postal codes and years) by 100. For comparative purposes, we also calculated total summertime mortality (all days regardless of temperature) for all postal codes and for target postal code years only using a variation of Eq. 6.

3.4 Example of training and testing methods

For illustrative purposes, we present the approach for Minneapolis, MN, for which data were available 1992 to 2008—thus 17 separate numerical "experiments" were conducted for that city. In the first iteration, the model was run using only data from 1993 to 2008, omitting 1992 (all values for 1992 were set to a missing value indicator). The mean summer temperature in the years included in the training period was 26.6 °C, and the threshold temperature was determined to be 30.2 °C. This threshold corresponded to a model-predicted increase in mortality of 2.3 % (Fig. 1). On the 296 days in the training period when the threshold temperature was exceeded, mortality increased 4.5 % relative to baseline.

Based on the 1993–2008 data, 15 postal codes in Minneapolis were identified as target zones (i.e., mortality in 15 postal codes was significantly greater than the baseline on above-threshold days in the training period). The mortality rate on above-threshold days in target zones averaged 23.4 % above the baseline, compared to 3.6 % below baseline in non-target zones.

In the withheld year, there were 5 days on which the temperature exceeded the threshold of 30.2 °C. On these days, after accounting for seasonality (Fig. 2), mortality was 15.6 % above baseline across the entire city. The average mortality rate across the 15 target zones on these 6 days was 23.2 % above baseline, whereas in the other postal codes mortality rates averaged 5.4 % above. Of the 15 target zones, nine were associated with high mortality in the withheld year (cell "A" in the contingency table) and six were not (cell "C"). Of the remaining 86 postal codes, 31 were associated with high mortality in the withheld year (cell "B") and 55 were not (cell "D"). The RR for this iteration was 1.66 (95 % confidence interval 1.01, 2.74), indicating that high mortality locations in 1992 could have been statistically reliably "predicted" using 1993–2008 data.

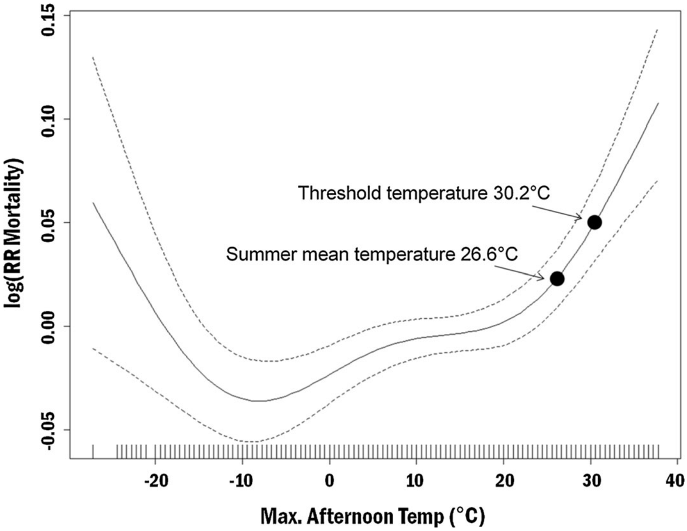

Fig. 1 The temperature–mortality relationship for Minneapolis, MN, based on a generalized additive model using data from 1993 to 2008. The threshold temperature is the lowest temperature at which model-predicted mortality significantly differs from that predicted at the summer mean temperature. The model-predicted value is the *solid center line,* and the confidence intervals are represented by the *dashed lines*. The vertical ticks on the horizontal axis indicate temperatures at which observations were available

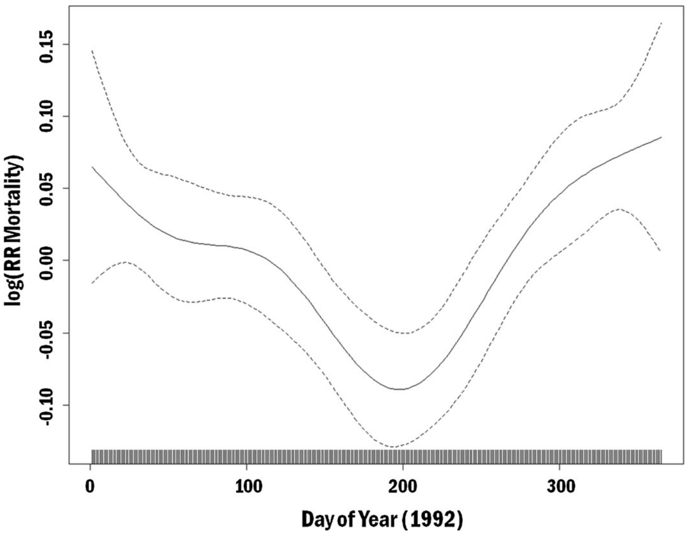

Fig. 2 The seasonality in mortality in Minneapolis, MN, during 1992. The pattern emerges from a generalized additive model using daily mortality data. The model-predicted value is the *solid center line,* and the confidence intervals are represented by the *dashed lines*

To demonstrate the calculation of excess mortality, we focus on postal code 55112, where the average daily mortality rate during the withheld year was 0.77 deaths per day. Mortality on the six above-threshold days was, in this postal code, 10.3 % above baseline, 0.85 deaths per day. The difference in these two daily mortality rates (0.08 deaths per day) is multiplied by the number of above-threshold days (five) for excess mortality in this postal code during the withheld year of 0.4 deaths. Across all target postal codes in 1992, there were 12.2 excess deaths.

4 Results

4.1 Model training

Statistically significant relationships between temperature and mortality were established for six of the seven study cities in all training iterations. The relationship between temperature and mortality was only significant in a portion of the training models for Atlanta; Atlanta was excluded from the remainder of the analysis. The threshold temperatures (Table 1) for the six other cities were between 1.6 °C (Philadelphia) and 4.1 °C (St. Louis) above the respective city mean summer afternoon maximum temperatures. The model-predicted increase in mortality at the threshold temperatures range from a 1.5 % increase (Philadelphia) to 2.3 % increase (Minneapolis-St. Paul). On all threshold-exceeding days mortality rates were 3.2 % (St. Louis) to 6.6 % (Philadelphia) above mortality rates on summer days with below-threshold temperatures. Depending on the year and city, between 4 and 31 postal codes were identified as targets (i.e., associated with high mortality during the training period). Across all cities, the average number of target postal codes ranged between 12 and 39 % of the total number of postal codes within the city.

For most cities, there was a clear separation between postal codes that tended to be identified as targets during training iterations and those that were not. In Philadelphia, for example, ten postal codes were identified as targets in every training iteration (i.e., they always had statistically high mortality rates on hot days), and another five were labeled as targets on over 90 % of the training runs. At the other end of the distribution, 26 postal codes were never associated with statistically significantly high mortality in any training model and one area was identified in only one iteration. A similar pattern was seen in the other cities. Mortality rates in the target zones on above-threshold days were markedly higher than that in non-threshold zones. On average mortality rates in the target zones were between 4.6 % (Philadelphia) and 20.5 % (Minneapolis) greater that in non-target zones during the training iterations (Fig. 3).

4.2 Model testing

After identifying target postal codes, we next evaluated how mortality in the target zones compared to that in the non-target zones during the single withheld year. As was the case during the training periods, mortality was higher in the target zones in the testing years relative to the non-target zones (Fig. 3). The ratio of target to non-target zone mortality was similar between the training and testing periods in Boston ($\sim$8 % higher), Minneapolis ($\sim$20 % higher), and Philadelphia ($\sim$4 % higher). In Phoenix, Seattle, and St. Louis there was a larger difference. In St. Louis, for example, mortality was 11.9 % higher in target zones using the training data but only 2.7 % higher than the non-target zones in the testing period. Of the six cities, Minneapolis showed the greatest difference in mortality between target and non-target zones in the withheld year and the difference was smallest in St. Louis.

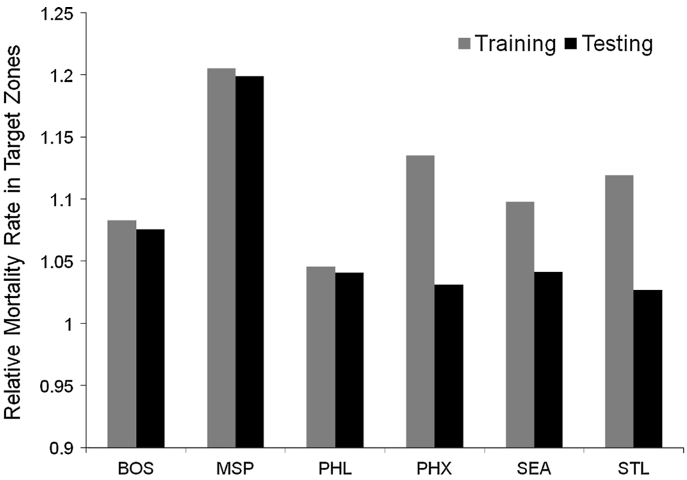

Fig. 3 The ratio of mortality rates on high-temperature days in target postal codes to non-target postal codes in each of the six cities examined. The ratios shown are the means across all training data sets (*gray bars*) and testing data sets (*black bars*). A ratio of 1.0 indicates that the mortality rate on hot days is equal between the two groups

Unlike the training periods, however, there was considerable variability evident in the difference between target and non-target zones in each sample of testing (Fig. 4). In total, we examined 130 city-years of withheld data; in 87 of these cases (67 %), mortality in the target zones was higher than that in the non-target zones. Mortality rates were higher in target zones in 16 of 17 (94 %) testing cases in Minneapolis, which was also the city associated with the greatest difference in mortality between target and non-target zones. St. Louis target zones were found to be associated with higher mortality in only 14 of 26 (54 %) testing cases, and this was the city associated with the smallest difference in mortality rates between targets and non-targets. Rates in the other four cities were between 58 and 76 %.

Contingency tables were employed for comparison of the likelihood of a postal code being associated with high mortality in the testing period if it was classified as a target relative to the likelihood of it being associated with high mortality if not classified as a target (Table 2). In all six cities we examined, the RR of high mortality in a target zone was greater than 1.0, but the RRs were statistically significant in only four (i.e., the confidence intervals did not include 1.0): Boston, Minneapolis, Philadelphia, and Seattle. Among these locations, the city-specific RRs varied from a low of 1.18 in Seattle to a high of 1.49 in Minneapolis; the overall average RR among these cities was 1.32. This means that, in these locations, the probability that a postal code will have high mortality in a given year is nearly 32 % higher for zones classified as targets compared to non-targets. The average for Phoenix and St. Louis was a 5 % increase in probability.

Finally, we compared two of the key metrics of this study: (1) the likelihood that each postal code would be labeled as a target, and (2) the likelihood that each postal code would be associated with high mortality in the withheld year. As evident from city-specific scatter-plots (Fig. 5) of these quantities compared to each other and the contingency tables, the overall tendency was for target zones to be associated with high mortality in the testing year. Correlations between the percentage of years each postal code was identified as a

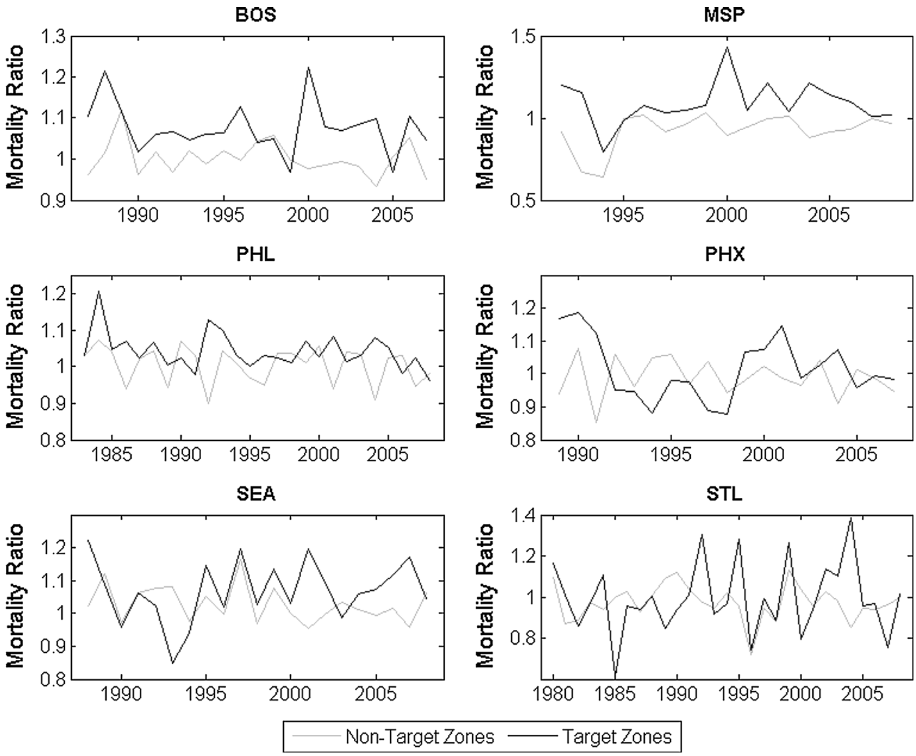

Fig. 4 A year-by-year comparison of mortality ratios on high-temperature threshold-exceeding days in target zones (*black lines*) and non-target zones (*gray lines*) in six study cities. A mortality ratio of 1.0 indicates that mortality on hot days was equal to the rate that would be expected for normal summer conditions

target and the percentage of withheld years that the postal code had high mortality were statistically significant for all cities based on the nonparametric Kendall tau test. Correlation coefficients ranged between 0.425 and 0.598.

4.3 Excess mortality

We calculated the number of excess deaths on above-threshold days to estimate the portion of the total heat-related mortality burden that could potentially be alleviated by targeting specific postal codes. Estimates of average annual excess mortality per city ranged from 7.9 for St. Louis to 46.1 in Philadelphia (Table 3). Over the study periods, we estimated 3,490 excess deaths on above-threshold days. Of these, nearly half (1,741: 49.9 %) occurred in postal codes labeled as targets. It should be reiterated that excess deaths were calculated using data from withheld years only; training data excluding the withheld year were used to identify target locations. The percent of excess deaths occurring in target postal codes ranged from a low of 11.2 % in Phoenix to a high of 103.4 % in Minneapolis. (Note: A percentage greater than 100 % indicates that mortality in non-target postal codes was below the summertime mean on above-threshold days).

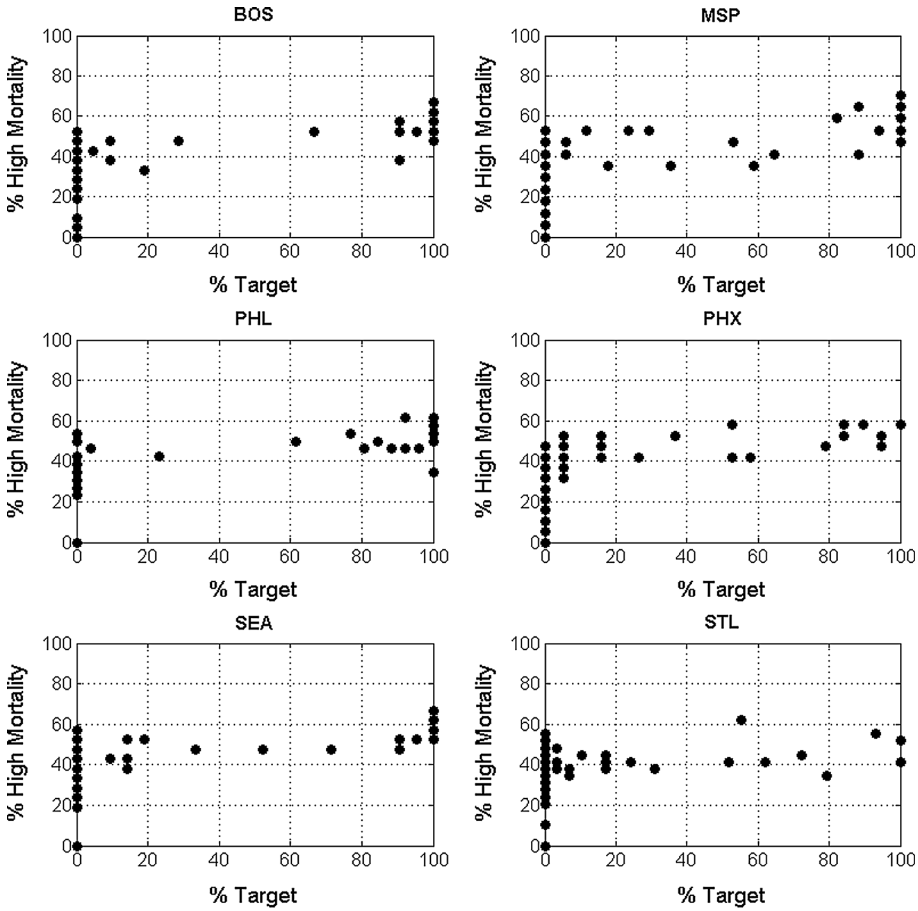

Fig. 5 Scatterplots comparing the frequency that each postal code (each represented by one *black circle*) was identified as a target zone during a model training iteration and identified as a high mortality zone during a model testing evaluation. A *circle* in the *bottom left hand corner* of each panel indicates places that were never identified as targets based on training data and were never associated with high mortality in testing data; the *top right hand corner* indicates places that were always identified as targets and always had high mortality

To provide a point of comparison for the percentage of excess mortality found in target zones, the total number of summertime deaths and total number of summertime deaths in target zone years is also shown in Table 3. If the percentage of excess mortality found in target zones mirrored the percentage of total mortality found in target zones, the excess mortality result might simply be reflecting different population sizes or overall mortality rates at the postal code scale. Although target postal codes represented 49.9 % of excess deaths on above-threshold days, they accounted for only 25.0 % of total summertime mortality. In five of the six cities, target zones accounted for a disproportionate fraction of the excess mortality on above-threshold days. In Minneapolis, which showed the greatest contrast, target zones were associated with 103.4 % of excess mortality but only 18.6 % of total mortality. In Boston and St. Louis, the difference was approximately a factor of two. Philadelphia and Seattle were associated with slight increases over total mortality, and in

Table 3 Descriptive statistics of total summer mortality and estimated excess heat-related mortality during the study period for each city examined

	BOS	MSP	PHL	PHX	SEA	STL	All cities
Total summer deaths	63,996	54,204	96,423	84,536	53,912	93,317	446,388
Total summer deaths (targets only)	19,330	10,090	47,268	13,020	10,287	11,800	111,795
Total excess mortality	770.55	301.88	1,198.80	456.05	533.29	229.08	3,489.65
Total excess mortality (targets only)	439.79	312.06	766.54	50.85	113.12	58.94	1,741.31
Percent of total deaths in targets	30.21	18.62	49.02	15.40	19.08	12.65	25.04
Percent of excess deaths in targets	57.08	103.37	63.94	11.15	21.21	25.73	49.90
Number of years	21	17	26	19	21	29	

The column to the right represents summation across all six cities

Phoenix target zones accounted for 11.2 % of excess mortality and 15.4 % of total mortality.

5 Discussion

A common theme in the heat-health literature is that heat-related mortality should be preventable. Our goal was to investigate whether prevention of such deaths could be facilitated by a geographically targeted approach, and our results indicated that there are locations within the cities we examined where the mortality rate was consistently high when extreme heat occurred. Our main findings were that postal codes identified as targets using a sample of training data were associated with higher mortality rates and were more likely to have high mortality in testing periods. There was statistically significant predictive ability in identifying high mortality postal codes in four of the study cities. Overall, predictability was strongest in Minneapolis and weakest in St. Louis and Phoenix. We believe this study is the first of its kind to explore the predictability of spatial units (e.g., postal codes) with high incidence of heat-related mortality. The results support the notion that geographically targeted heat intervention activities within cities may lead to greater benefits than a uniform citywide approach. If heat-related mortality could be prevented in the target zones only—a relatively small number of postal codes in most of the study cities—the benefits for public health could be quite substantial.

There were considerable differences in the study results between cities. Regardless of the metric used to quantify predictability, Minneapolis consistently exhibited the greatest difference in heat impacts between target and non-target zones. Depending on the specific model iteration, only 10–15 of the 101 postal codes in the Minneapolis study area were identified as targets. On above-threshold days, these locations were associated with 20 % higher mortality rates than non-target zones. More strikingly, however, these locations were associated with 103 % of the *excess* mortality on above-threshold days. As the target postal codes only account for 18.6 % of the total summertime mortality during the study period, the results provide strong evidence that there are specific localities within the Minneapolis area that are especially sensitive to heat.

In Phoenix and St. Louis, target postal codes were the least different from non-target zones. Target zones in both cities were associated with higher mortality rates and a higher RR of being associated with high mortality in any given year than non-targets, but the RR was not significantly different than random. The fact that the response is more homogeneous across Phoenix may not be surprising, as the hot desert climate has forced Phoenix residents to adopt a range of adaptation measures to cope with high temperatures throughout much of the year. We did not include enough cities in this study to determine whether there is systematic variability in the predictability of high-risk zones related to geography and climate, but our results are consistent with such a hypothesis (generally higher predictability in Boston, Seattle, and Minneapolis, intermediate in Philadelphia, and lower in St. Louis and Phoenix). With respect to excess mortality, we found that the percentage of excess deaths within target postal codes in Phoenix (11.2 %) was lower than the percentage of total deaths (15.4 %) that occurred within those regions. This result seems to contrast the fact that we found overall higher mortality rates on above-threshold days within target postal codes in Phoenix on hot days; the difference arises because the excess mortality calculations are weighted by the population of each postal code and the number of above-threshold days each year. The finding is consistent with the fact that there was not statistically significant model skill in identifying the places where heat mortality was more likely to occur in withheld years.

The use of a prediction-based framework for analysis using separate training and testing data sets is especially relevant and important for problems where there may be significant costs associated with a missed forecast (e.g., wasted resources or underserved vulnerable populations). Heat-health researchers have only recently begun to use this type of study design (Hajat et al. 2010). Although there is a very strong statistical association between heat and human mortality in many locations, the correlation between temperature and mortality is imperfect. Not all hot days are associated with excess mortality, and some individuals are likely to be unaffected by heat (Hajat et al. 2010; Zhang et al. 2012; Pascal et al. 2013). This inherent noisiness should be accounted for when designing intervention and mitigation activities.

An additional strength of this study is the extension of an intra-city, outcome-based approach to identifying high-risk locations for heat-related mortality to six cities that have not been previously examined in such a manner. In an era of finite public resources, the identification of zones where elevated mortality rates can reliably be anticipated in advance offers a useful tool for planning and response. Practical measures that could benefit from this work include public warning systems, the opening of cooling centers, activation of phone lines and buddy networks, and allocation of additional medical resources and personnel (Sheridan and Kalkstein 2004). There are also longer-term strategies employed to reduce the impacts of extreme heat that could benefit from identification of areas with temporally predictable risk; municipalities are adopting educational campaigns and building weatherization and modification programs aimed at reducing urban heat island effects (e.g., Solecki et al. 2005; Smith and Levermore 2008).

This research concerns the *likelihood* that a particular postal code will experience heat-related mortality in a given year; this complements information about the magnitude of the response within each postal code aggregated over longer time periods (e.g., Vaneckova et al. 2010; Hondula et al. 2012). Geographic zones where the heat-related risk is consistently high are ideal candidates for focused intervention measures. Although the temperature–mortality relationship is indeed noisy as discussed above, in these particular locations, the response is more consistent from one year to another. The methodology we presented in this manuscript could readily be applied to many other locations where heat may pose a threat to public health. In our experience, the limiting factor in the feasibility of conducting such a study is access to geographically referenced health outcome data; we advocate for the accessibility of such data to the research community. Furthermore, the results of studies like this one merit re-evaluation on a regular basis, as new temporal and spatial patterns in risk may emerge in the coming years.

We acknowledge that this study design creates a disconnect from the proper chronological sequencing of training and testing data; in practice, data from 2007 could never be used to generate a model to predict outcomes in 1995. As surveillance data continue to be accumulated, we recommend that a different subsetting model be applied to determine whether a model generated in one time period can be used to predict outcomes in a subsequent period. This is particularly important for the problem at hand because of the long-term time trends in the response that could be impacted by acclimatization and climate change. We identified an absolute temperature threshold (i.e., one that does not vary across months or years), but research efforts continue be directed to improve the identification and definition of high-risk extreme heat days (e.g., Barnett et al. 2010; Zhang et al. 2012; Antics et al. 2013). There may be both theoretical and model performance advantages to using more complex metrics for quantifying the relationship between human health and heat, and the optimal metric could vary from one city to another. The thresholds we found were based on the model-predicted mortality (and error) at given temperatures relative to the mean summer temperature, and for some cities this resulted in the threshold temperature falling several degrees or more below those linked to activation of various warnings and advisories by weather forecast offices (National Weather Service 2006) and temperatures commonly perceived as being dangerous (Sheridan 2007). This is not necessarily a weakness of our study, as we likely included all days where heat could have been a contributing factor to elevated summertime mortality rates. We also included *all* days where the temperature exceeded a city-specific threshold, but it would be worthwhile to examine whether the response is more predictable when limiting the study to heat waves where dangerous temperatures are experienced over several consecutive days (Anderson and Bell 2011). We did not consider the effects of mortality displacement (Hajat et al. 2005; Saha et al. 2013) or potential confounding by air pollutants (e.g., Roberts 2004).

6 Conclusion

Using multi-decadal high-resolution mortality records from cities spanning the climate zones of the United States, we have found evidence that geographic zones associated with higher heat-related mortality are predictable. The RR of a given postal code being associated with high mortality in one withheld year if identified as a target zone from a training model was statistically better than random in Boston, Minneapolis, Philadelphia, and Seattle. In these cities, there was a 32 % greater chance that target zones for intervention activities had high heat-related mortality in a data set unseen by the predictive model. Phoenix and St. Louis were associated with greater temporal heterogeneity in the spatial mortality pattern on hot summer days and did not exhibit statistically significant predictability. Across all six cities, 50 % of excess mortality on above-threshold temperature days was confined to postal codes identified as targets. As these locations represented only 25 % of total mortality during the study period, they were disproportionately impacted by extreme heat. Study results indicate significant reductions in heat-related mortality could be achieved with proper intervention programs aimed at specific localities within major metropolitan areas, particularly in Boston, Minneapolis, Philadelphia, and Seattle.

References

Anderson GB, Bell ML (2011) Heat waves in the United States: mortality risk during heat waves and effect modification by heat wave characteristics in 43 US communities. Environ Health Perspect 119(2):210

Antics A, Pascal M, Laaidi K, Wagner V, Corso M, Declercq C, Beaudeau P (2013) A simple indicator to rapidly assess the short-term impact of heat waves on mortality within the French heat warning system. Int J Biometeorol 57:75–81

Arnfield AJ (2003) Two decades of urban climate research: a review of turbulence, exchanges of energy and water, and the urban heat island. Int J Climatol 23:1–26

Barnett AG, Tong S, Clements ACA (2010) What measure of temperature is the best predictor of mortality? Environ Res 110(6):604–611

Davis RE, Knappenberger PC, Michaels PJ, Novicoff WM (2003) Changing heat-related mortality in the United States. Environ Health Perspect 111(14):1712

Gardner MJ, Altman DG (1989) Statistics with confidence: confidence intervals and statistical guidelines. BMJ Books, London

Guo Y, Barnett AG, Tong S (2012). High temperatures-related elderly mortality varied greatly from year to year: important information for heat-warning systems. Scientific Reports 2

Hajat S, Armstrong BG, Gouveia N, Wilkinson P (2005) Mortality displacement of heat-related deaths: a comparison of Delhi, Sao Paulo, and London. Epidemiology 16(5):613–620

Hajat S, Sheridan SC, Allen MJ, Pascal M, Laaidi K, Yagouti A et al (2010) Heat-health warning systems: a comparison of the predictive capacity of different approaches to identifying dangerously hot days. Am J Public Health 100(6):1137–1144

Hondula DM, Davis RE, Leisten MJ, Saha MV, Veazey LM, Wegner CR (2012) Fine-scale spatial variability of heat-related mortality in Philadelphia County, USA, from 1983–2008: a case-series analysis. Environmental Health 11(1):1–11

Hondula DM, Davis RE, Rocklöv J, Saha MV (2013a) A time series approach for evaluating intra-city heat-related mortality. J Epidemiol Community Health. doi:10.1136/jech-2012-202157

Hondula DM, Vanos JK, Gosling SN (2013b) The SSC: a decade of climate–health research and future directions. Int J Biometeorol. doi:10.1007/s00484-012-0619-6

Laaidi K, Zeghnoun A, Dousset B, Bretin P, Vandentorren S, Giraudet E, Beaudeau P (2012) The impact of heat islands on mortality in Paris during the August 2003 heat wave. Environ Health Perspect 120(2):254

Minnesota Population Center (2011) National Historical Geographic Information System: Version 2.0. University of Minnesota, Minneapolis

National Weather Service (2006) Heat wave: a major summer killer. http://www.nws.noaa.gov/om/brochures/heat_wave.shtml Accessed 1 July 2013

Pascal M, Wagner V, Le Tertre A, Laaidi K, Honoré C, Bénichou F, Beaudeau P (2013) Definition of temperature thresholds: the example of the French heat wave warning system. Int J Biometeorol 57(1):21–29

Reid CE, O'Neill MS, Gronlund CJ, Brines SJ, Brown DG, Diez-Roux AV, Schwartz J (2009) Mapping community determinants of heat vulnerability. Environ Health Perspect 117(11):1730

Reid CE, Mann JK, Alfasso R, English PB, King GC, Lincoln RA et al (2012) Evaluation of a heat vulnerability index on abnormally hot days: an environmental public health tracking study. Environ Health Perspect 120(5):715

Roberts S (2004) Interactions between particulate air pollution and temperature in air pollution mortality time series studies. Environ Res 96(3):328–337

Rocklöv J, Forsberg B, Meister K (2009) Winter mortality modifies the heat-mortality association the following summer. Eur Respir J 33(2):245–251

Saha MV, Davis RE, Hondula DM (2013) Mortality displacement as a function of event strength in seven U.S. cities. Am J Epidemiol 179(4):467–474

Sheridan SC (2007) A survey of public perception and response to heat warnings across four North American cities: an evaluation of municipal effectiveness. Int J Biometeorol 52(1):3–15

Sheridan SC, Dolney TJ (2003) Heat, mortality, and level of urbanization: measuring vulnerability across Ohio, USA. Clim Res 24(3):255–265

Sheridan SC, Kalkstein LS (2004) Progress in heat watch-warning system technology. Bull Am Meteorol Soc 85(12):1931–1941

Sheridan SC, Kalkstein AJ, Kalkstein LS (2009) Trends in heat-related mortality in the United States, 1975–2004. Nat Hazards 50(1):145–160

Smargiassi A, Goldberg MS, Plante C, Fournier M, Baudouin Y, Kosatsky T (2009) Variation of daily warm season mortality as a function of micro-urban heat islands. J Epidemiol Community Health 63(8):659–664

Smith C, Levermore G (2008) Designing urban spaces and buildings to improve sustainability and quality of life in a warmer world. Energy Policy 36(12):4558–4562

Solecki WD, Rosenzweig C, Parshall L, Pope G, Clark M, Cox J, Weincke M (2005) Mitigation of the heat island effect in urban New Jersey. Glob Environ Change Part B Environ Hazards 6(1):39–49

Vaneckova P, Beggs PJ, Jacobson CR (2010) Spatial analysis of heat-related mortality among the elderly between 1993 and 2004 in Sydney, Australia. Soc Sci Med 70(2):293–304

Wilby RL (2003) Past and projected trends in London's urban heat island. Weather 58(7):251–260

Wood SN (2003) Thin plate regression splines. J R Stat Soc Ser B (Stat Methodol) 65(1):95–114

Wood S (2006) Generalized additive models: an introduction with R. CRC Press, Boca Raton

Zhang K, Rood RB, Michailidis G, Oswald EM, Schwartz JD, Zanobetti A et al (2012) Comparing exposure metrics for classifying 'dangerous heat' in heat wave and health warning systems. Environ Int 46:23–29