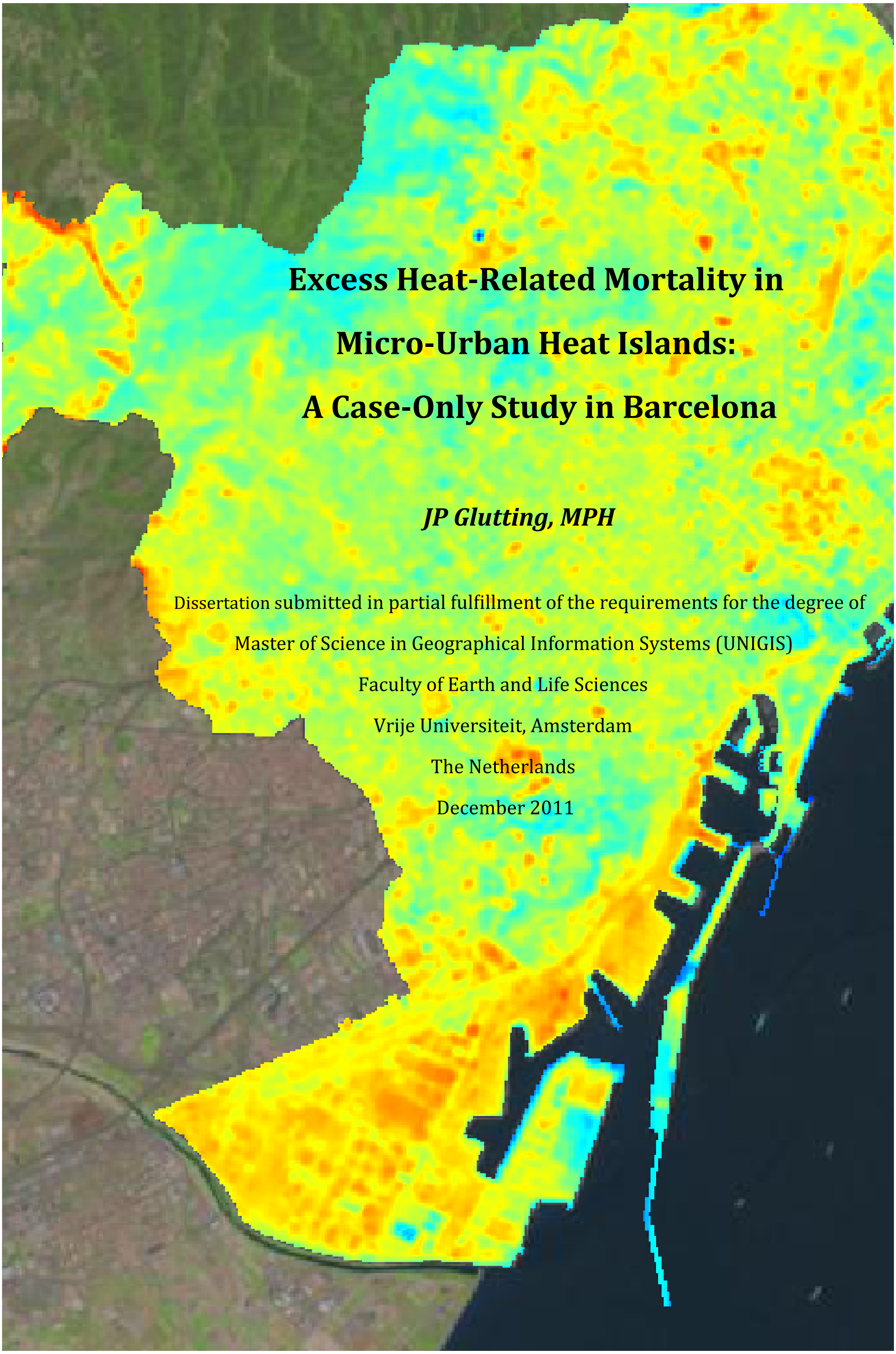


Excess Heat-Related Mortality in Micro-Urban Heat Islands: A Case-Only Study in Barcelona

JP Glutting, MPH

Dissertation submitted in partial fulfillment of the requirements for the degree of

Master of Science in Geographical Information Systems (UNIGIS)

Faculty of Earth and Life Sciences

Vrije Universiteit, Amsterdam

The Netherlands

December 2011

Thesis Committee

Prof. Dr. Henk Scholten

Dr. Jasper Dekkers

Drs. Chris Jacobs

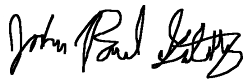
Drs. Ronnie Lassche

Drs. Nils de Reus

Declaration of Originality

I certify that this work is entirely my own, unless explicitly acknowledged (including citation of published and unpublished sources). This work has not been previously submitted in any form to the Vrije Universiteit or to any other institution for assessment or for any other purpose.

Signed:



Date: 15/12/2011

Table of Contents

1 Introduction 8

2 Literature Review 9

2.1 Heat Effects 9

2.2 Identifying Dangerous Heat 10

2.2.1 Findings 11

2.3 Vulnerability Mapping 13

3 Research Methodology 15

3.1 Overview 15

3.2 Study Area 15

3.3 Data Sets 17

3.3.1 Overview 17

3.3.2 Mortality data 17

3.3.3 Weather data 18

3.3.4 Census Data 18

3.3.5 Thermal infrared Images 18

3.3.6 Data Selection 19

3.3.6.1 Data Sets 19

3.3.6.2 Selection Criteria 19

3.4 Data Processing 20

3.4.1 Overview 20

3.4.2 Mortality data 20

3.4.3 Weather data 24

3.4.4 Census data 26

3.4.5 Thermal Infrared Images 26

3.5 Data Analysis 28

3.5.1 Overview 28

3.5.2 Descriptive Analysis of Variables 29

3.5.3 Case-Only Analysis 31

3.5.3.1 Special Confounding Issues 33

3.5.3.2 Spatial Issues 34

3.5.3.3 Summary 34

3.5.4 Time-Lagged Mortality 35

3.5.5 Regression 35
3.5.5.1 Akaike Information Criterion 36
3.5.5.2 HRR Threshold 37
3.5.5.3 HAT temperature types and measures 42
3.5.6 Modeling Time Lag 45
3.5.7 Cofactor Selection 50
3.5.8 Controlling for Spatial Autocorrelation 51
3.5.9 Addressing Confounding Variables 52
3.5.9.1 Air-conditioning 52
3.5.9.2 Air Pollution 53
3.5.9.3 Ethnicity 53
3.5.9.4 Population Density 53
3.5.10 Network Selection and Weighting 54
3.5.11 Final Model 54
3.5.12 Further Testing 55
3.6 Weaknesses of the Study Design 55
4 Results 57
4.1.1 Other Model Choices 57
5 Discussion 59
6 Conclusion 60
6.1 Practical Implications 60
6.2 Caveats 61
6.3 Potential improvements and further study 61
References 63
Appendix 1: Initial Regression Results 67
Appendix 2: Details of Methods 70
6.4 Data Management 70
6.5 Regression 71
6.5.1 Model 71
6.5.2 Moran eigenvector GLM filtering 71
6.6 Modeling Time Lag 72
6.6.1 How it works 72
6.6.2 Methods 73
6.7 Network Selection and Weighting 73

6.7.1 Network Selection........ 73
6.7.1.1 Network Types........ 74
6.7.1.2 Network Weighting........ 75
6.7.1.3 Network Selected........ 75
6.7.2 The effect of distance thresholds on the model........ 76
6.7.3 Other Network Types........ 76
Appendix 3: Moran's I Values........ 78
Appendix 4: Variable Correlations........ 82
Appendix 5: Spatially-Adjusted Regression Results........ 83

Index of Tables

Table 1: Weather Stations in Barcelona.......... 24

Table 2: Landsat 5 thermal infrared images meeting study criteria.......... 28

Table 3: Landsat 7 thermal infrared images meeting study criteria.......... 28

Table 4: Summary Statistics for Study Variables.......... 29

Table 5: A 2x2 table of a case-only design with the exposure (High Air Temperature) in rows and the modifier (High Relative Radiance) in columns.......... 31

Table 6: A 2x2 table of a case-only design with exposure (High Air Temperature) in rows and modifier (High Relative Radiance) in columns, substituting formulas into cells.......... 33

Table 7: Percentage Distribution of Deaths by RRZ (Z-Score) 42

Table 8: 2x2 Table for HRR threshold of 0.4.......... 42

Table 9: 2x2 Table for HRR Threshold of 0.2.......... 42

Table 10: 2x2 Table for HRR Threshold of 0.0.......... 42

Table 11: Absolute Case Numbers and Global Percentages by Temperature Measure and Percentile..........43

Table 12: Selected regression results for 90th percentile temperatures, before adjustment for spatial autocorrelation.......... 44

Table 13: Temperature cut-points for 90th percentile of values (1996-2008).......... 45

Table 14: Moran's I values for selected regression model variables.......... 51

Table 15: Selected regression results, after adjustment for spatial autocorrelation.......... 54

Table 16: Selected results of sensitivity testing of the regression model.......... 57

Table 17: Regression results by Weighting Scheme.......... 58

Table 18: Regression results before adjustment for spatial autocorrelation.......... 67

Table 19: Results of distance threshold testing with best model (tmin90_l2)..........76

Table 20: Results of a 36% (2000 point) sample, by network type and distance..........77

Table 21: Moran's I values for study variables, by Network and Weighting Type.......... 78

Table 22: Correlations between study variables.......... 82

Table 23: Regression results after adjustment for spatial autocorrelation.......... 83

Illustration Index

Illustration 1: Heat mortality curves for 11 US cities (Curriero 2002) ... 12

Illustration 2: Temperature-Mortality curve and 95% confidence bands in Barcelona (Baccini, et al. 2008) ... 12

Illustration 3: Barcelona, Spain, city limits. Census data, 2001. [overleaf] ... 15

Illustration 4: Data processing work flow. ... 21

Illustration 5: Barcelona weather stations and warm-season, at-home deaths, 2000-2003.[overleaf] ... 24

Illustration 6: Population density in persons per square kilometer, by census tract. Census data 2001. [overleaf] ... 29

Illustration 7: Psuedocolor representation of average radiance (z-score) from Landsat 7 thermal infrared images (Barcelona, 2000-2002)[overleaf] ... 37

Illustration 8: At-home deaths April 1 – September 30, 2000-2003 (red). HRR0 areas in yellow. [overleaf] ... 39

Illustration 9: Histogram of Deaths by RRZ (Z-Score). Thresholds: 0.0 [red], 0.2 [blue], 0.4 [green]. ... 41

Illustration 10: Minimum Temperature Effects Across 10-day Lag (3D Plot). Data 2000-2003, April-September ... 46

Illustration 11: Minimum Temperature Effects Across 5-day Lag (Contour Plot). Data 2000-2003, April-September, tmin90=20.94, RR=relative risk ... 48

Illustration 12: ... 48

Illustration 13: Apparent Mean Temperature Effects Across 5-day lag. Data 2000-2003, April-September. ... 49

Illustration 14: Mean Temperature Effects Across 5-day lag.Data 2000-2003, April-September. ... 49

Illustration 15: Apparent Minimum Temperature Effects Across 5-day lag.Data 2000-2003, April-September ... 49

Illustration 16: Minimum Temperature Effects Across 5-day lag.Data 2000-2003, April-September. ... 49

Illustration 17: Apparent Maximum Temperature Effects Across 5-day lag.Data 2000-2003, April-September ... 50

Illustration 18: Maximum Temperature Effects Across 5-day lag.Data 2000-2003, April-September. ... 50

Illustration 19: Schematic of PostGIS 2.0 database used for project analysis ... 70

1 Introduction

Extreme heat events (EHE), some of which are commonly known as “heat waves”, are an issue of increasing concern and research in recent years, especially after the high mortality experienced during a series of European EHE in 2003 and 2006 (*Fouillet et al. 2008; Kovats and Hajat, 2008*). Recently published climate models predict increases in the frequency, duration and, especially, the intensity of EHE in the latter half of this century (*Gosling et al. 2008*). Meehl and Tebaldi (*2004*) have noted that these changes are expected to affect the Mediterranean area disproportionately.

Urban Heat Islands refer to the phenomenon in which urban areas are warmer than surrounding, non-urban areas. Over time, using more sophisticated instruments, it has become clear that heat is not distributed uniformly throughout urban areas, and the heterogenous nature of this distribution gives rise to micro-urban heat islands (MUHI), areas of more intense heat within an urban zone.

The research question addressed in this study is the following: Are micro-urban heat islands associated with an increased risk of mortality during extreme heat events? A secondary questions is: Which temperature measures are the best predictors of this association, if any?

The data used for exploring the effect of MUHI on mortality were images from satellite remote thermal infrared sensing, weather station readings, a mortality registry for the city of Barcelona, and the Spanish census. The methods used were a case-only study design, distributed lag non-linear models to determine the lag in heat-associated mortality, logistic regression, and eigenvector filtering to partially adjust for spatial autocorrelation in the regression model. The tools used in the study were a PostGIS 2.0 spatially-enabled PostgreSQL database, the R statistical language, and a selection of other Open Source tools (*gdalwarp*, *MCElite*).

This document first reviews what is known about heat-associated mortality and micro-urban heat islands, in the *Literature Review* section. The study design and methods are presented in the *Research Methodology* section, which includes a short presentation of the study area, followed by the three main sections: *Data Sets* which describes the data used in the study, *Data Processing* which describes the way the data sets were handled and new data sets derived for the analysis, and the *Data Analysis* which focuses on the the fitting of regression models to the processed data.

The final portion of the paper consists of the presentation of the *Results*, the *Discussion* of the results and the implications of the choices made, and the *Conclusion*. After the *References*, the *Appendices* include the full results set, including results not selected for the main text, and technical details about software and regression formulas used.

2 Literature Review

There is an extensive and growing literature addressing many different aspects of the relationship between heat and mortality. These include identifying heat thresholds that create important changes in mortality, and risk factors that increase mortality generally or in specific populations.

2.1 *Heat Effects*

Extreme heat stresses a living organism by increasing core body temperature, cholesterol levels, blood viscosity and circulation of blood to the skin (reducing circulation to the heart and lungs) (*Curriero 2002*). Increased sweating can lead to dehydration. Prolonged heat can cause heat cramps, heat syncope, heat exhaustion, heat stroke, hospitalization and death (*Kovats and Hajat 2008; Luber and McGeehin 2008*). Heat may kill so quickly that hospital admissions do not reflect the full scale of mortality (*Kovats and Hajat 2008*); some authors report that most victims are found dead in their residences (*Bouchama et al. 2007*).

Heat effects are particularly detrimental to certain sub-populations: the elderly (*Iñiguez 2010; O'Neil et al. 2009*), the poor (*Basu 2009; O'Neil et al. 2009*), minorities (*Basu 2009; Schwartz 2005b*), the mentally ill (*Hansen et al. 2008*), outdoor workers (*Kinney et al. 2008*), the bedridden (*Bouchama et al. 2007*) and people who take certain types of medications that interfere with electrolytes and water balance (e.g., diuretics, anticholergic agents, and tranquilizers that impair sweating) (*Johnson et al. 2009; Martin-Latry 2007; Luber and McGeehin 2008; Hansen et al. 2008*).

While the relationship between ethnicity and increased heat-related mortality may not be obvious initially, it has been borne out in several studies (*Basu 2009; Kinney et al, 2008; O'Neill et al. 2003; Schwartz 2005b*), despite the formidable complexity of the interactions between race/ethnicity, class, social status and health. These findings are not overly surprising, as minority status has been associated with disparities in health care provision and health outcomes in various areas, even when controlling for social and economic factors (*Brulle and Pellow, 2006; Klonoff 2009*), partly because minority populations tend to benefit less economically from the same level of education, and there are large disparities in wealth at the same income levels (*Whaley 2003*). Although health-care is free in Spain, there continue to be health-care disparities in immigrant and minority populations, as recent immigrants tend to work longer hours and have more inflexible work schedules which make it more difficult for them to receive preventive care and lead to more emergency visits. They also face language and cultural barriers (*Rué et al. 2008*).

Multiple (not necessarily incompatible) explanations have been offered for these differences. Disparities in health and health outcomes between White populations and non-White populations have been extensively studied in the US epidemiological literature - close to half of all studies published in the American Journal of Epidemiology between 1921 and 1990 included measures of race/ethnicity (*Whaley 2003*). Research on racial disparities in health outcomes indicate that some of these differences are likely due to cumulative factors related to life experiences and discrimination in the diagnosis and treatment of illnesses (*Klonoff 2009*), although the underlying mechanisms are not clear. The subjective experience of discrimination itself has been associated with poorer mental and physical health outcomes (*Williams et al. 2008*). To add a spatial dimension to this issue, areas where minority populations live tend to be subjected to disproportionate levels of environmental health impacts (*Brulle and Pellow 2006*). Clearly, measures of race/ethnicity are likely to be multiply confounded and difficult to examine.

Because heat and mortality do not have a clear "dose-response" relationship with identifiable thresholds, a wide-ranging set of topics are being explored to:

1. identify dangerous heat and important exacerbating factors (humidity, heat duration, high nighttime temperatures, housing, etc.) (*Basu 2009*)
2. identify vulnerable populations (the elderly, mentally ill, chronically ill) (*Bouchama et al. 2007; O'Neill 2009*)
3. create heat health watch warning systems (HHWWS) to protect vulnerable individuals (*Hajat et al. 2010; Matthies and Menne 2009*), and
4. intervene to better manage vulnerable groups and improve urban conditions, including improving housing and urban planning (*Kinney et al. 2008; Kovats and Hajat 2008; Matthies and Menne 2009*).

2.2 Identifying Dangerous Heat

There are several principal epidemiological approaches to identifying dangerous heat in a given location (*Hajat et al. 2010; Kovat and Hajat 2008*). These include two broad types of analysis: general additive models (GAM) (time-series) (*Basu 2009; Basu and Samet 2002; Bayentin et al. 2010; Curriero 2002; Gosling et al. 2008; Iñiguez et al. 2010; Kovats and Hajat 2008; Metzger et al. 2010*), and case-only studies (*Armstrong 2003; Medina-Ramón et al. 2005; Schwartz 2005b*). A GAM is a complex statistical model used to analyze time-series data. This analysis requires a number of sophisticated techniques to establish baseline values for variables and compensate for health effects that have seasonal periodicities (flu, allergies, etc.) or day-of-the-week periodicities (increased travel-related accidents on weekdays, alcohol-related accidents on weekends, transportation-related air pollution, etc.). Case-only studies do not use data as a time-series, but divide cases by time-fixed and time-variable characteristics. A more detailed explanation of the case-only methodology can be found in the *Case-Only Analysis* portion of the *Data Analysis* section.

Important variables analyzed include:

- **maximum, minimum and mean air temperatures** and their relationship with mortality over time. Minimum temperatures conceptually are important because of their physiological effect of not allowing nocturnal respite from the effects of heat,
- **temperature-humidity measures** (humidex, heat index or "apparent temperature" and the thermo-hygrometric index), as humidity has important impacts on the physiological effects of heat at a given temperature (*Basu 2009; Metzger et al. 2010*), these measures can be better

predictors of mortality than regular air temperatures,

- **heat duration measures** for minimum, maximum and synthetic (heat index, apparent temperature) temperatures. Moving averages of the current day and a previous time-period attempt to estimate the cumulative or lagged effects of heat (*Kinney et al. 2008; Metzger et al. 2010*),
- **spatial synoptic classification** of weather conditions, also used in localized epidemiological studies. These attempt to categorize local weather patterns into a small number of groups (hot-dry, hot-humid, etc.) for easier analysis (*Metzger et al. 2010*).

For practical purposes, humidity has been found to be a less useful measure for interventions such as developing HHWWS, as it is difficult to predict accurately (*Hajat et al. 2010*).

Confounding factors such as physical activity, exposure to sunlight or air-conditioning can exacerbate or mitigate the effects of local heat conditions. There are several approaches to dealing with these factors, which fall into two main categories: exposure assessment for specific populations such as the ill and elderly (*population-based*), or assessing the stress inherent in being in a specific place at a specific time (*place-based*) (*Basu and Samet, 2002*). The latter would address issues of exposure to sunlight or shade.

Air-conditioning is an extremely important factor that overlaps the two categories. It is place-based in that it only applies to people who spend the hot days in certain settings, but it also has a population-based aspect, in that it potentially can be linked to income, social class or even ethnicity. O'Neill, Zanobetti and Schwartz (*2005*) studied the prevalence of air-conditioning by ethnic groups in urban areas in the US. They found that mortality among Blacks was more prevalent than among Whites and that White households were twice as likely to have central air-conditioning. It concluded that part of the explanation for racial disparities in EHE-related mortality was due to unequal access to air-conditioning. Functioning air-conditioning decreased the probability of death in Chicago's 1995 heat wave by 70% (*O'Neill et al. 2009*).

In addition, statistical techniques such as distributed lag models (*Armstrong 2006; Schwartz 2000*) are used to examine the relationships between exposure and associated mortality over time. Studies have shown mixed evidence of lag periods between peak heat and peak mortality, but the majority show short-term effects over one to three days (*Gosling et al. 2008*).

2.2.1 Findings

A common epidemiological approach is to plot different types of regression curves for the temperature-mortality relationship. These curves are normally U, V or J shaped, and can be used to identify a minimum mortality temperature (MMT), above and below which mortality increases.

There are several key findings. First, almost all attempts to identify dangerous heat have shown that threshold mortality levels, and MMT, vary from one city to the next, generally around average temperatures in a location, making a “universal” measure of dangerous heat more difficult to identify.

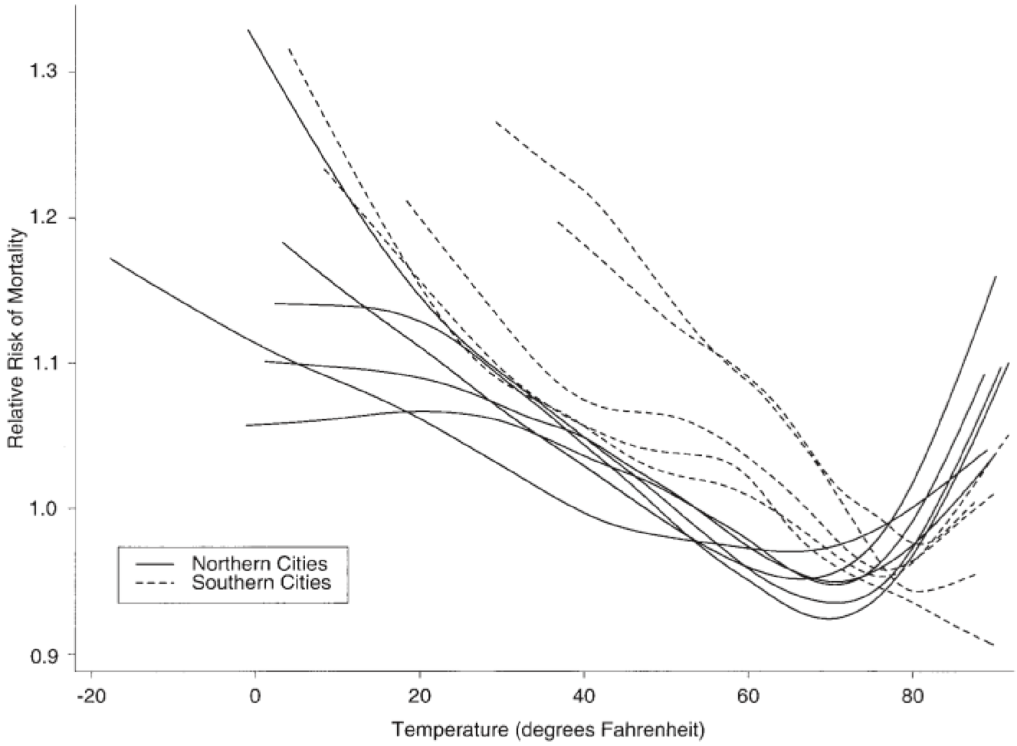


Illustration 1: Heat mortality curves for 11 US cities (Curriero 2002)

Hotter areas have higher MMTs than colder areas (*Curriero et al. 2002; Iñiguez et al. 2010*). Illustration 1 shows heat-mortality curves for 11 US cities, grouped by their relative latitude (North or South). Unfortunately, most later research has shown these types of curves to be somewhat plastic, changing over time in the same locations.

Temperature-mortality curves for Barcelona show similar patterns (see Illustration 2), with a minimum mortality temperature of 22.4°C (95% CI: 20.7 – 24.2) maximum apparent temperature (*Baccini et al. 2008*).

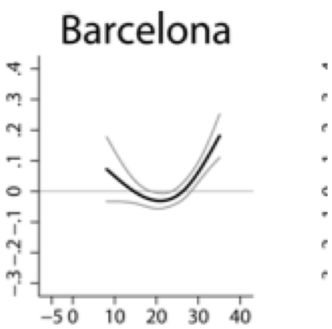


Illustration 2: Temperature-Mortality curve and 95% confidence bands in Barcelona (Baccini, et al. 2008)

Other findings:

- Air conditioning has a strong impact on reducing or eliminating mortality (*O'Neill et al. 2009*), by up to 70%, as mentioned above.
- Mortality increases in a non-linear association with temperature and the duration of an EHE (*Metzger et al. 2010; Saez et al. 2000*).
- Air pollution, especially ozone, may increase with heat and may have synergistic effects on mortality, although the evidence is unclear and varies depending on the type of pollution (*Basu 2009*). According to Luber and McGeehin (*2008, page 30*), “A positive association has been found between temperatures > 32°C (> 90°F) and ground-level ozone production, and increasing evidence suggests that ozone and high temperature affect mortality synergistically.” Other studies have shown associations between heat and ozone, but no clear association with mortality (*Schwartz 2005a, Smargiassi et al. 2009*).
- There can be a lag period between peak heat and peak mortality. Bayentin et al. (*2010*) found evidence of a lag in ischemic heart disease specific mortality based on a study of year-round temperatures. However, Kinney et al. (*2008*) found “Heat effects on mortality usually have been observed to occur immediately (i.e., heat today affects deaths today) and/or with a 1-day lag, whereas cold effects may occur with a multiple-day lag”. Metzger et al. (*2010*) found that average, minimum and maximum temperatures and averages for the same or up to three previous days (lags 0-3) “performed equally well”. In a literature review, Gosling et al. (*2008*) found that most studies reported the main effects of heat with a 0- or 1-day lag, and only two reported lagged main effects for a period of longer than three days.
- Food poisoning cases (e.g. salmonella) increase during EHEs (*Matthies and Menne 2009*), creating additional morbidity and mortality.

There are some additional caveats to these findings, however. Examining historical data from too far in the past is problematic, as overall mortality is decreasing (*Metzger et al. 2010*) and the use of air conditioning is increasing (*Kinney et al. 2008*). Heat effects may vary, and there is some evidence of an adaptive effect, from one hot year to another (*Fouillet et al. 2008, Kinney et al. 2008*) or between EHEs in the same season. Some suggest that the latter is due to a “harvesting” effect, leaving less vulnerable individuals later in the season (*Metzger et al. 2010*), while others attribute these to changes in behavior (restricted activity, acquisition of air-conditioning, or the implementation of HHWWS). Gosling et al. (*2008*) found evidence of mortality displacement in the literature, but the variable definitions of heat and displacement periods mean that the results are very mixed.

2.3 Vulnerability Mapping

One technique that is helpful for the development and operation of HHWWS is the creation of vulnerability maps, identifying vulnerable populations and other potential risk factors (*Johnson, Wilson and Luber 2009; Kovats and Hajat, 2008; O'Neill et al. 2009*). These maps can use a combination of *population-based* or *place-based* (spatial) criteria. Vulnerable populations include those mentioned above: the elderly, the poor, racial/ethnic minorities, the mentally ill, people taking some types of medications, outdoor workers and the bedridden.

One spatial factor out of many that affects exposure to heat is the urban heat island (UHI) effect, which causes cities to have higher temperatures than their less urbanized surroundings, due to hard surfaces that absorb heat (steel, cement, asphalt), building density, wind and the the effect of the built environment on wind direction and speed, and lack of vegetation. Urban areas also show high variability in surface temperatures over short distances (*Nichol et al. 2009*), variations commonly referred to as "micro-urban heat islands" (MUHI). MUHI have been associated with increased mortality risks in previous publications (*Johnson, Wilson and Luber, 2009; Kestens et al. 2011; Smargiassi et al. 2009*). MUHI are important because they affect densely populated areas, thus putting large numbers of people at risk.

3 Research Methodology

3.1 Overview

The question that the project sought to answer based on this map and supporting data was the following: Are micro-urban heat islands associated with an increased risk of mortality during extreme heat events? A secondary question was: Which temperature measures are the best predictors of this association, if any? The general method for answering these questions was to create and test a place-based heat vulnerability map for the city of Barcelona.

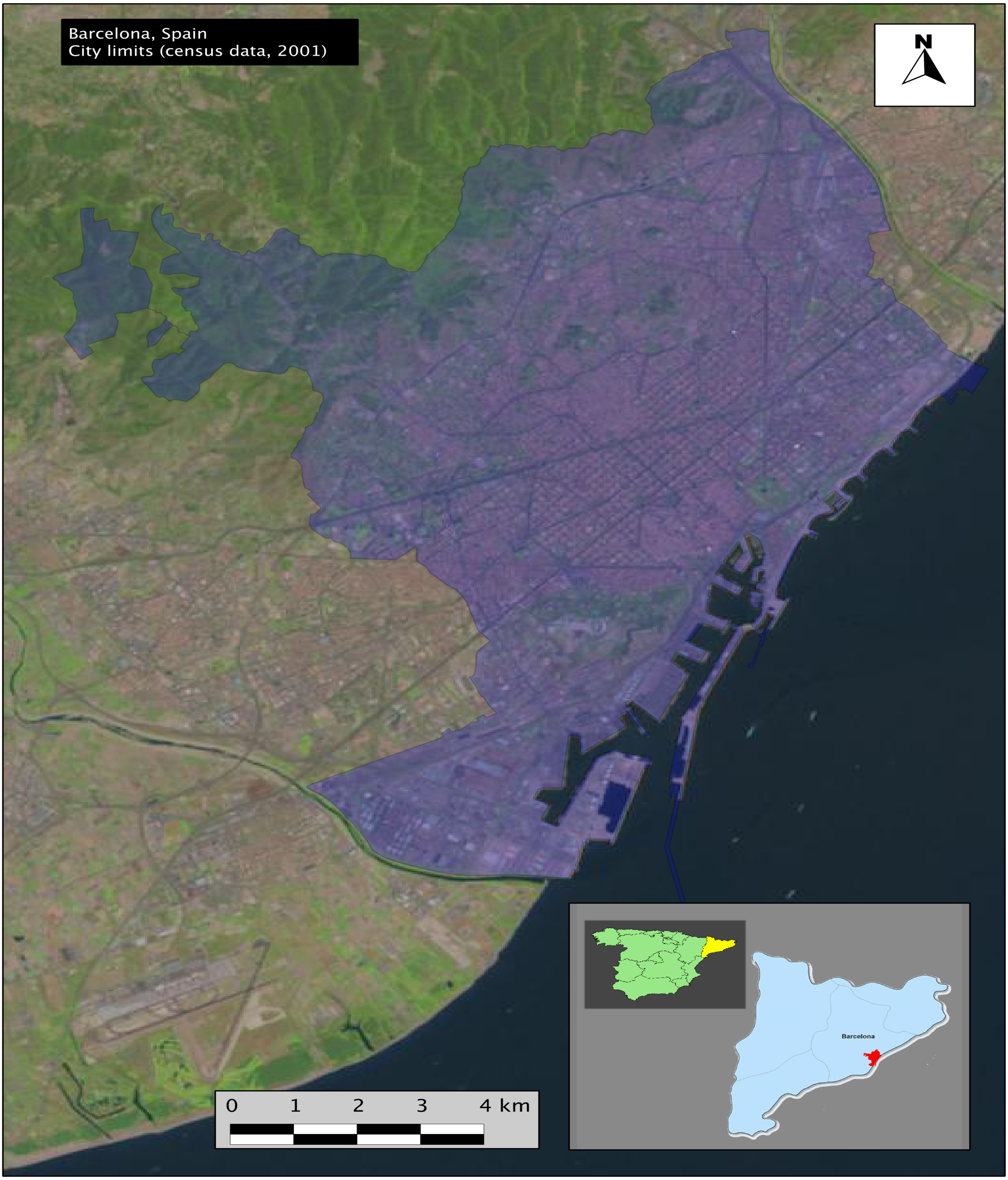
In order to identify MUHI and relate them to real effects on human mortality, several steps were required, as described below:

1. Identify daily average air temperatures over the study period.
2. Create a composite, stable measure approximating fixed intra-urban *variability* (not absolute values) in land surface temperatures (LST) in the study area. LST tend to vary more over short distances than air temperatures (*Nichol et al. 2009*), and a stable composite measure would identify fixed "hot spots" of interest, or MUHI.
3. Test whether the MUHI identified were associated with relative increases in mortality during EHE. This testing involved:
    1. Mapping georeferenced MUHI in the study area.
    2. Locating (geocoding) the addresses of deaths in the study period.
    3. Identifying whether the deceased lived within a defined MUHI.
    4. Identifying time lag in heat-related mortality in the data set.
    5. Identifying and adjusting for confounding factors (to the extent possible).
    6. Identifying and adjusting for spatial autocorrelation.

3.2 Study Area

The area examined in the study was the urban area of Barcelona, Spain. Barcelona is the capital of Catalonia, an autonomous region in the northeast of Spain, shown in Illustration 3, below. The study area itself was defined by the municipal boundaries of the city. The map shows that the city covers a coastal area, part of a mountain, and includes both dense urban and wooded areas.

Illustration 3: Barcelona, Spain, city limits. Census data, 2001. [overleaf]


Barcelona, Spain
City limits (census data, 2001)
N
0 1 2 3 4 km
Barcelona

3.3 *Data Sets*

3.3.1 Overview

To carry out this analysis, information was gathered about global air temperatures, the distribution of persistent MUHI in the study area and the dates and locations of all deaths. All data was kept and, to a large extent processed, in a PostGIS database (the PostgreSQL relational database management system, with spatial extensions installed) and covered the period from 2000-2003, and the months of April-September (inclusive), as these are the months in which EHE occur in the study area (*Baccini et al. 2011; Baccini et al. 2008; Iñiguez et al. 2009; Saez et al. 2000*).

Most data sets used were available for the years 2000-2008, but the study period was reduced to the years 2000-2003, for reasons that will be explained in detail below. The standard datum and coordinate system for the study were the World Geodetic System (WGS84) datum and the Universal Transverse Mercator geographical coordinate system, zone 31° North (UTM 31N). The EPSG code for the spatial coordinate system used is 32631.

The following data were obtained:

- A registry of all deaths in Barcelona by date and address, between 2000 and 2009.
- Temperature data from multiple weather stations in and around Barcelona, from 1996 to 2008.
- Population counts by census tract for the year 2001.
- Two sets of satellite thermal infrared (TIR) images of the study area between the years 2000 and 2009, one from Landsat 5 and another from Landsat 7.

The specifics of each data set are described below.

3.3.2 Mortality data

Mortality data was collected from mortuary registries, and available by address of residence and address of death. It also included the sex, age (an integer, not date of birth), and day of death of the deceased. Cause of death was unavailable because the data source was not health-related. In any case, health data that provided more information would need to be aggregated into areal counts to address data privacy concerns.

In the case that the deceased died more than one day before being found, it is unclear whether the date of death registered was the date the body was collected by the authorities or the estimated date of death based on forensic investigation, although the latter would be more appropriate.

The use of all-cause deaths (with the exception of accidents) is common in the study of heat-related mortality, because there is currently no systematic definition of heat-related mortality (*Basu 2009*). This study used only deaths at the address of residence, so traffic and other accidents outside the home were largely excluded.

Some advantages and limitations of the mortality data set are described below:

Advantages

1. The dataset included all deaths in the city of Barcelona during the study period.
2. All deaths were identified by residential address and address of death, a level of detail unavailable in other types of mortality data.

Limitations

1. Cause of death was not available.
2. Other variables were limited (age and sex only).
3. No information was available on health status (chronic illness, medications, hospital admissions, etc.).

Appropriateness

1. Cause of death (other than to exclude accidents) is often not used in general heat-related mortality studies.
2. Lack of other variables can simplify analytical methods (see *Data Analysis*, below).
3. The ability to match residential and death addresses made it possible to pinpoint exposure on at least part of the day of death.

3.3.3 Weather data

The weather data set included daily air temperatures (minimum, maximum and mean temperatures), atmospheric pressure and relative humidity from weather stations in and around Barcelona. The data set also included wind speed, wind direction, precipitation and other meteorological variables which were not be used in this study, as the complexity of modeling these interactions were beyond the scope of this project. The positions of the weather stations were geolocated (Datum: ED50, GCS: UTM 31N). These coordinates were converted to the standard coordinates for the study (Datum: WGS84, GCS: UTM 31N; EPSG:32631).

3.3.4 Census Data

Census data were used to calculate population density per square kilometer as a reference value, based on census data collected in the year 2001. Population density per square meter could also have been used, but values by square kilometer are easier to interpret intuitively. No information about air-conditioning (AC) in the home was collected in the 2001 census. There was a limited survey of AC use done in 2008, and questions about AC were eventually included in the 2011 census, which was not available for this study.

3.3.5 Thermal infrared Images

TIR images (band 6: 10.4-12.5 μm) taken by the Landsat 7 instrument (60 meter resolution, resampled to 30m resolution) were used for the study. All images were screened to select only those images with less than 10% cloud cover and were inspected visually by the author to insure that clouds did not obscure any of the study area.

An alternative set of images from Landsat 5 (120 meter resolution) were collected, but not used. The resolution of Landsat 7 is significantly higher, although its temporal coverage was limited by the failure of its scan line corrector in May of 2003. After assessment (see *Thermal Infrared Images* in the *Data Processing* section), the Landsat 7 images were chosen.

3.3.6 Data Selection

3.3.6.1 Data Sets

The final data sets used were the following:

- Mortality data for 2000-2003, April-September (inclusive)
- Meteorological data for 2000-2003, April-September (inclusive)
- Landsat 7 images from the years 2000-2002 (nine total, listed in Table 3)
- Census data for the year 2001 (to calculate population density)

3.3.6.2 Selection Criteria

The decision was made to limit the study period to 2000-2003, despite having data for a larger period of time, in order to adjust the study period to the availability of the best quality data. The best TIR images publicly available were from Landsat 7, which were not available after 2003 due to equipment failure. Not only was the resolution of the Landsat 7 instrument better than Landsat 5, but the density of images over time (nine in three years) was much better than for Landsat 5.

As for meteorological data, apparent temperatures could not be calculated after 2003 from the data set used, due to a change in the availability of meteorological data after that year. Also, population density measures, a key parameter in the regression model, were calculated for the year 2001, based on the most recent census, and change significantly over the next ten years.

Furthermore, the use of air-conditioning (and associated confounding of results) was less likely to take place earlier in time, especially before the widespread and widely publicized mortality caused by the heat waves of 2003. It is has been suggested that AC and other types of interventions have reduced mortality in later heat waves (*Fouillet et al. 2008)*.

There is a final issue of the computational resources needed to process the data to adjust for spatial autocorrelation (see *Controlling for Spatial Autocorrelation*). The eigenvector filtering approach used is a brute-force technique that involves selecting key eigenvectors out of a matrix developed from networks built on the coordinates of cases. Preliminary tests indicated that the memory and computing power required increased with the square of the number of cases included, although in practice the computing time needed was unpredictable and often much larger, possibly due to scale-driven changes in the matrices from which the eigenvectors were derived.

3.4 Data Processing

3.4.1 Overview

The fundamental dynamic was a process of enriching the mortality data by linking it with other types of data in a PostGIS database. By geocoding the addresses of the deceased, these locations could be mapped over other data layers containing meteorological information, processed TIR imagery, as well as census tracts and associated information (population density). All data were transformed to the WGS84 datum and the UTM 31N geographical coordinate system (EPSG: 32631).

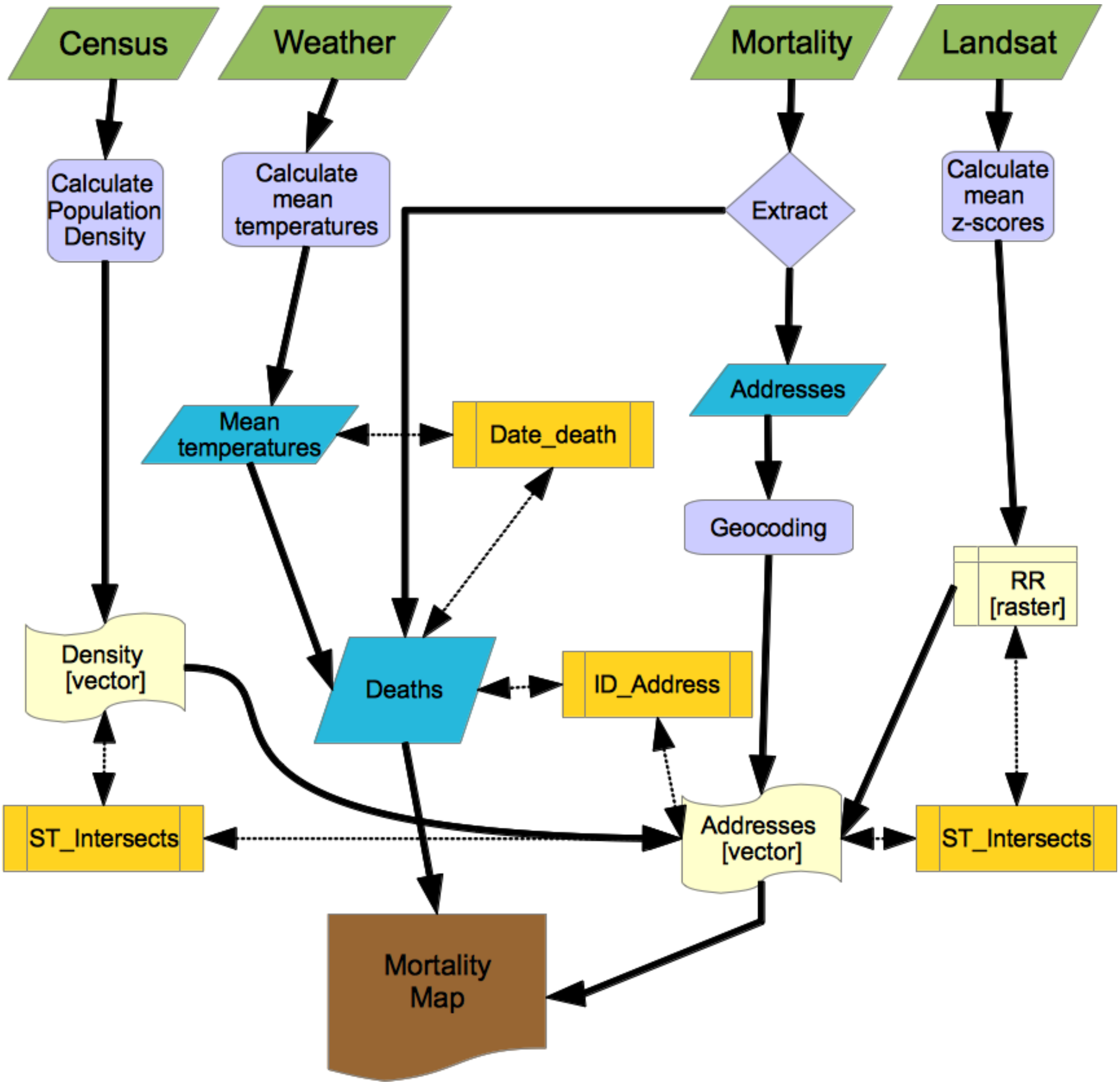
For the most part, data was stored and processed in a PostGIS database. Most of the basic processing and data enrichment was done using the functionality provided by PostGIS version 2.0 (distance, spatial awareness, point-in-polygon calculations, etc.). Additional analytical functions were performed using the R statistical package, which can be compiled directly into the PostgreSQL executable as the PL/R language, although processed data were exported to R for regression analysis.

Raster functionality has been integrated into the PostGIS project starting at version 2.0, which is currently a pre-release version. The raster functionality in PostGIS 2.0 was not complete at the time of the study, so some analyses were done outside the database with *gdalwarp* and a modified version of *MCElite* (software previously developed by the author). A schematic for the data processing workflow is shown in Illustration 4, at the end of this section.

3.4.2 Mortality data

The mortality dataset consisted of one large Excel file per year (2000-2009) with uncoded deaths registered by date, including sex, age, residential address, address of death and no other information. There were 165,744 deaths registered for residents of Barcelona during this ten year period, 150,574 with addresses. Deceased persons who were not in Barcelona at the time of their death, or who either had no known address or whose address was not registered in the database, were not included in the analytical data. This criterion may have excluded the deaths of homeless or undocumented persons, who would be particularly vulnerable to the effects of heat.

This curated data was transferred into two tables, one containing unique street addresses, and a second containing deaths identified with a unique id, including the date of death, other personal information, and two address id fields for the residential address and death address. The address id in both cases was an external key from the address table (see schematic, Illustration 4). Mean temperature data was linked to the id code for deaths by the date of death and relative radiance and census tract id (thus density) were linked to the address table using the ST_Intersects spatial function of PostGIS (see *Appendix 2: Details of Methods* for more details).



Green: raw data; Purple: data processing; Blue: processed data; Yellow: spatial data; Orange: data links; Brown: final map.

Illustration 4: Data processing work flow.

The address table contained "clean" street addresses, with apartment numbers and other "vertical" information removed. While apartment height certainly would have an effect on the subjective experience of heat, the specific effect is unclear, and not necessarily uniform. Increased height could result in increased airflow, but also increased solar gain, and both issues depend on the presence, distance and orientation of nearby buildings. Also, position relative to the absolute height of a building, especially whether a subject lived on the top floor or not, and the age and level of insulation of the building, could have important effects on the heat inside the dwelling. Modeling at this level of detail was beyond the scope of the project.

There were roughly 50,000 distinct street addresses in the data set, with multiple death events at the same street address. This fact is important when developing different types of networks, as commented on below, as classical Gabriel and Relative Neighbor networks do not allow multiple nodes in the same position, so these networks were slightly modified to make them symmetric.

The addresses from the address table were cleaned and then geocoded using a Python script to extract addresses from the database, geocode them using the Google and Yahoo geocoding APIs, and write the results back to the database, along with information about the quality of the geocoded address match. These results were filtered for quality using a cross-checking method: by matching geocoded addresses returned by the respective APIs to parts of the submitted address (street number, street name, postal code, city, country). The Google-coded addresses were chosen for the study. The Google API used (Version 3) takes an address as text, determines the closest match from its database, and returns that matched address with its WGS84 coordinates, along with an assessment of the quality of the match and the geocoding. Google provides the following codes for the match as the variable *location_type*, in decreasing order of quality:

1. **Rooftop** (exact coordinates of the address)
2. **Range Interpolated** (coordinates interpolated from street name and street number)
3. **Geometric Center** (the returned coordinates are for the centroid of the matched area, such as city, postal code, etc.)
4. **Approximate** (results with unknown precision)

The quality criteria defined for Google-geocoded addresses to be used in the study were the following:

1. All five elements of the original address (street number, street name, postal code, city, country) had to match the geocoded address returned by the Google API.
2. The geocoded address had to have a *location_type* of "Rooftop" or "Range Interpolated".

Geocoding was an iterative process. Street names were included in the mortality records in an extremely abbreviated form, which in many cases needed to be modified so that Google's software could interpret them properly. There were also systematic errors in the addresses ("MARQUES" as both "MQES" and "MQUES"), and these errors needed to be corrected one at a time and then the failed addresses processed again. In addition, Google places a limit of 2,500 addresses that can be geocoded in a single 24-hour period, requiring multiple geocoding sessions.

Of the 50,335 street addresses, 41,032 met the defined geocoding quality criteria (address matching on street number, street name, postal code, city and country), representing a total of 122,481 deaths. Of these deaths, 28,494 had the same address for both residence and death – that is, they were deaths in the home. These deaths were fairly evenly distributed during the study period (between 2,300 and 2,800 deaths per year from 2000-2009), and the years 2000-2003 were not substantially different from the others. The study period (April to September) between 2000-2003 included 5,554 geocoded, at-home deaths.

In-home deaths were used for several reasons. First, heat-related mortality tends to take place among the elderly and in the home (*Bouchama et al. 2007*). Secondly, it is difficult to determine if persons dying outside the home or in hospitals were exposed to recent heat in their area of residence (they may have been hospitalized for a significant period of time or may have been staying at another address), whereas in-home deaths mean the deceased necessarily experienced local conditions at the time of death (with the exception of those found more than one day after death, but proper death certificates should include probable time of death).

An additional issue is the cause of death. Because EHE can affect human physiology in multiple and not entirely understood ways (interactions with medications, food poisoning, etc.) (*Bouchama et al. 2007; O'Neill et al. 2010*), there are two principle types of study samples used with reference to cause of death: specific cause (cardiovascular or asthma, for example) or all-cause, excluding accidents. Currently, there is no systematic definition of heat-related mortality (*Basu 2009*), making overspecification of causes of death potentially counter-productive, resulting in missing some effects of EHE on mortality. Many researchers use all-cause mortality due to potential misclassification; very few deaths are classified as heat-related, and these definitions vary by location (*Basu and Samet, 2002*). This study used all-cause mortality, including accidents, as there was no way to exclude them. However, at-home deaths excluded traffic accidents, one of the largest causes of accidental mortality.

3.4.3 Weather data

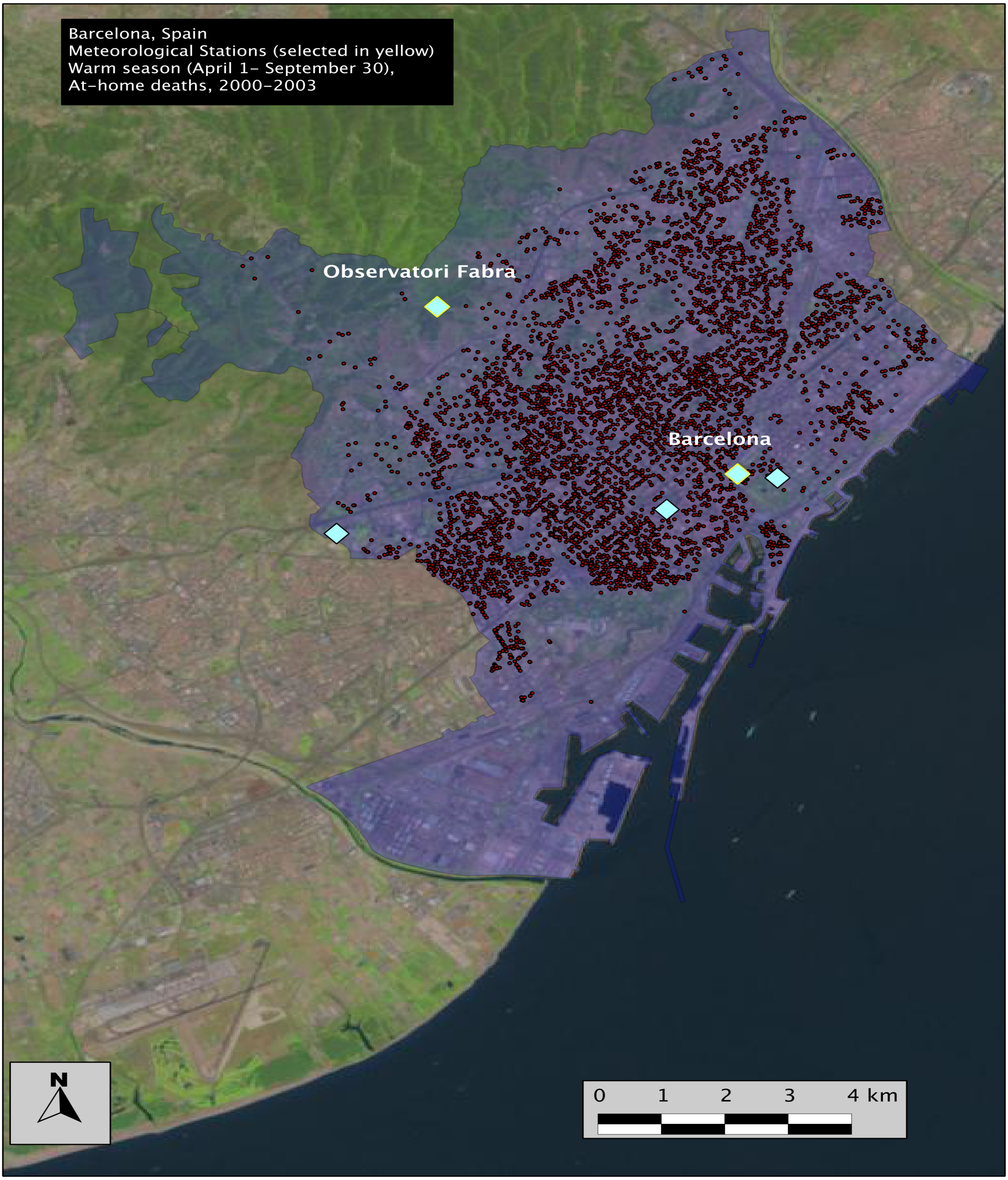
Original weather data was obtained from 28 weather stations in Catalonia, Spain, five of which were within the metropolitan limits of Barcelona. These data included daily temperature measures, and some other measures, for the periods in which the weather stations were in operation. Table 1 provides information about each of the weather stations from which data was collected, while Illustration 5 shows their location in the city, along with geocoded deaths during the warm season (April 1 – September 30, 2000-2003).

Table 1: Weather Stations in Barcelona

Station	Codes	Operational	Temperature	Humidity
Barcelona	**4N / 27**	**1994-2002**	**Yes**	**Yes**
Observatori Fabra	**D5 / 3**	**1996-2008**	**Yes**	**Yes**
Raval	X4 / 21	2006-2008	Yes	No
Zona Universitària	X8 / 22	2008	Yes	No
Zoo	X2 / 20	2006-2008	Yes	No

Of these six weather stations, only two (highlighted in yellow) collected humidity data and were operational during the study period (the *Observatori Fabra* and *Barcelona* weather stations), and only the former for the entire study period. Both stations collected information about relative humidity, needed for the calculation of apparent temperature. The *Observatori Fabra* is located above most of the populated areas of Barcelona, partially up the Collserola mountain in a wooded area. The Barcelona station was located in the city itself, near the *Arc de Triomf*. Between the two, they cover a range of conditions in the city, as well as a vertical distance of 403.7 meters – the Barcelona weather station is 7.5 meters above sea level and the Observatori Fabra station is located 411.2 meters above sea level. Illustration 5 shows the weather stations with respect to the cases in the study. Temperatures were recorded as the minimum, maximum and mean for each day (there were no hourly measures).

Illustration 5: Barcelona weather stations and warm-season, at-home deaths, 2000-2003. [overleaf]


Barcelona, Spain
Meteorological Stations (selected in yellow)
Warm season (April 1- September 30),
At-home deaths, 2000-2003
Observatori Fabra
Barcelona
N
0 1 2 3 4 km

A database table was created with averaged daily temperature values (minimum, maximum, mean) for the study area during the period for which temperatures were recorded (1996-2008). These were averages of the temperatures taken from weather stations in Barcelona, a technique which was at least as accurate as temperature assessments used in many published epidemiological studies, which often use measures at the nearest airport to determine temperatures for the study area (*Baccini et al. 2011; Baccini et al. 2008; Smargiassi 2009*). The Barcelona airport is visible in the southwest corner of Illustration 5. These means were later used to calculate the percentiles of temperatures in the study area.

Apparent temperatures (AT) and their averages were also calculated, using the formula:

$$AT = -2.653 + 0.994 * temperature + 0.0153 * dewpoint^2$$ (*Baccini et al. 2008*)

3.4.4 Census data

The Spanish census, managed by the INE (*Instituto Nacional de Estadistica*), provides demographic information about the Spanish population, organized by census tracts. The most important information for this study were population density and the use of air conditioners. Population density was calculated based on population counts and the size of census tracts for the year 2001, based on local population registries. The population density of Barcelona changed from 14,850.26 to 15,912.42/km^2 between 2001 and 2008, although the population distribution did not change dramatically during this time period. A strict comparison is somewhat difficult, as the census areas are changed yearly, and sometimes twice in a single year.

The official census did not incorporate questions about the use or availability of air-conditioning until the most recent version (being carried out this year), although there was a specific survey done in 2008 to determine the use of air-conditioning. The results of that survey showed that some 35% of households in Spain had air-conditioning in 2008. The average for Catalonia was 36.1%, slightly higher than the national average. The size of the survey meant that detailed information by census tract was not available. The responses were broken down by income categories as well, ranging from 22.3% of households with household incomes of less than 1,000 euros per month with air-conditioning, while for households with monthly incomes greater than 2,700 euros that percentage rose to 46.2%.

3.4.5 Thermal Infrared Images

TIR images taken by the Landsat satellites can be converted to an estimated LST value (*Smargiassi et al. 2009*). However, this study worked directly with the radiance values recorded by the Landsat 7 TIR images, not temperature estimates. The goal was to determine areas where radiance was consistently higher in comparison to the rest of the study area, not to determine specific temperature thresholds or mortality increments attributable to a specific (e.g. one degree) change in temperature. Smargiassi et al. (*2009*) used a similar technique to analyze mortality associated with MUHI in Montreal, by stratifying their temperature measures. The *Data Analysis* section contains more details.

Standardized scores were generated for each pixel position in the study area based on Landsat 7 TIR images that fulfill the inclusion criteria (< 10% clouds), as follows:

1. Radiance values (0-255) are provided for each pixel in each Landsat TIR image.
2. Mean and standard deviation (SD) values were calculated for pixel values in each image.
3. Z-score values $z = \frac{x - \mu}{\sigma}$ were calculated for each pixel based on the population parameters (mean and SD) of each TIR image. The results were rasters of z-scores indicating the relative radiance (RRZ) over the study area for each Landsat TIR image used in the study.
4. Z-score rasters were averaged to produce a single raster with mean RRZ scores, to indicate the averaged RRZ for each pixel area.

This value represented the RRZ of that location as compared to the rest of the study area, using z-scores from TIR image pixel values, averaged over multiple images for reliability. The z-scores served as a proxy for relative differences in LST across the study area at any given temperature or time. The estimated LST itself was not calculated, because, while LST based on TIR images are correlated with real temperatures, there are several factors that limit the precision of the estimate, and the temporal coverage of the TIR images was limited. The goal was not to produce numeric values for temperatures on specific days, but to identify areas that were relatively hotter than others within the city on most days. No analysis of land use or ground cover was done, although this might have helped screen industrial and forested areas out of the study data set, for example. This could have been useful for limiting the range of z-scores to examine, but it would have artificially limited the study area, as there are residences inside the Collserola park, and in other locations that are technically illegal. While this introduces some confusion, it also includes a wider range of radiance values than purely residential areas might.

Because the raster manipulation functionality in PostGIS was not yet feature complete, the raster images had to be processed outside the database. A modified version of the MCElite library (*http://pyqgis.org/contributed/mcelite.zip*) was used for the calculations and the GeoTiff results were imported using the *raster2pgsql.py* program included with the PostGIS source code.

It is important to note that Barcelona underwent changes with the construction of the Diagonal Mar section of the city during the first decade of the 21st century, converting largely industrial areas to more mixed urban space (commercial, housing). Thus, averaging RRZ across the entire study period could miss important relationships in areas where the built environment underwent significant changes, and the available Landsat images were examined with this in mind.

Landsat 5 images (Table 2) were divided into two groups, four from the year 2003 and three from the year 2009. Landsat 7 images were only available for the years 2000-2002, but there were more of them available in this period (nine, from 2000-05-13 to 2002-09-24) (Table 3), and the resolution of the Landsat 7 instrument was much better.

Table 2: Landsat 5 thermal infrared images meeting study criteria

Date	Quality
2003-07-10	Excellent
2003-07-26	Excellent
2003-08-02	Excellent
2003-08-11	Excellent
2009-07-01	Excellent
2009-08-11	Excellent
2009-08-18	Excellent

All images with less than 10% cloud cover and no clouds obscuring study area.

Table 3: Landsat 7 thermal infrared images meeting study criteria

Date	Quality
2000-05-13	Good
2000-08-10	Excellent
2000-08-17	Excellent
2000-09-11	Excellent
2001-04-14	Excellent
2001-06-26	Excellent
2001-08-13	Excellent
2002-05-19	Good
2002-09-24	Good

All images with less than 10% cloud cover and no clouds obscuring study area.

The z-scores were calculated and averaged for the selected images, and the resulting raster file imported to the PostGIS database. The values of the rasters at each address were extracted and added to the coordinate data for each death. These coordinates were later used by R to create spatial layers for analysis.

3.5 Data Analysis

3.5.1 Overview

A descriptive analysis was done of the study variables. Then a more detailed evaluation of the data was done using a case-only study design, distributed lag non-linear models (DLNM) to estimate lags in mortality, logistic regression analysis and eigenvector filtering to partially compensate for spatial autocorrelation in the regression model.

3.5.2 Descriptive Analysis of Variables

A preliminary descriptive analysis of the study variables is shown in Table 4.

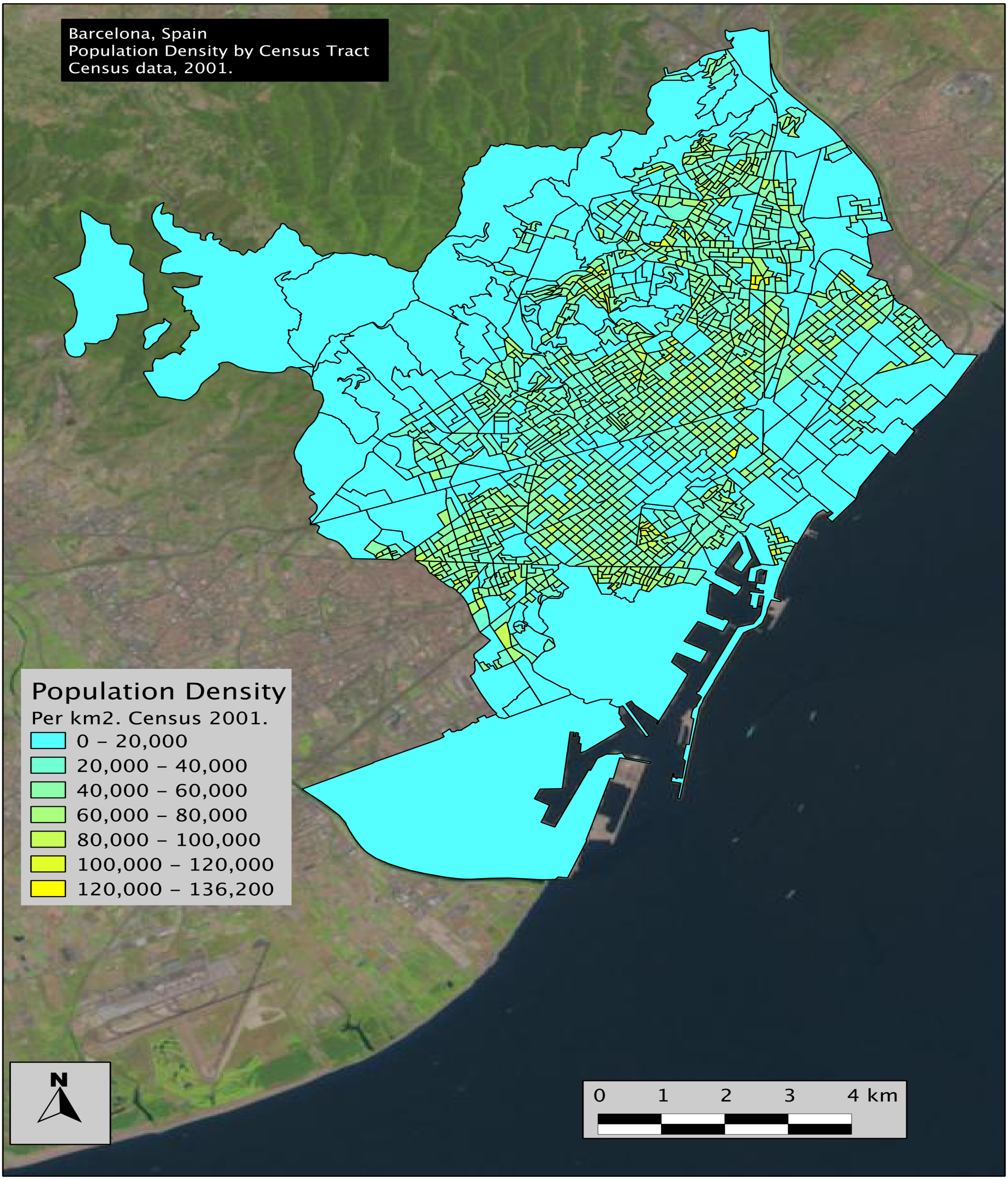
Table 4: Summary Statistics for Study Variables

Variable	Minimum	Median	Mean	Maximum	Standard deviation
Age	0	84	81.8	108	11.869
Mean temperature	8.65	21.35	20.75	31.60	4.789
Apparent mean temperature	5.95	18.59	17.99	28.77	4.754
Minimum temperature	4.20	18.15	17.37	27.20	4.588
Apparent minimum temperature	1.52	15.41	14.63	24.39	4.546
Maximum temperatures	11.70	25.90	25.39	39.20	5.406
Apparent maximum temperature	8.99	23.12	22.60	36.34	5.391
Population density (km^2)	76.63	37,000.00	38,720.00	136,200.00	22,407.800

The variable age was provided as an integer, so the minimum age of zero would correspond to infant mortality. The average and median ages of death, however, are elevated. This is not an unexpected result for at-home deaths (or deaths in the first world). For several of the variables (age and all the temperature measures), the mean values are lower than the median values, indicating that the distribution of these variables is skewed to the left. It is worth noting that, in all cases, the effect of humidity measures was to reduce the extremes of subjective temperatures rather than increase them. This is possibly due to the dry Mediterranean climate of the area, resulting in a low dew-point with respect to the current temperature.

Barcelona is also one of the most densely populated cities in the world, with one census tract achieving a population density of 136,200 persons per square kilometer (the real values for this area are 2096 persons living in an area of 15,389 square meters, in an oddly shaped tract in downtown Barcelona). The population density by census tract is shown in Illustration 6. The highest density areas are located around the city center, with the coastal and mountain areas having lower population densities.

Illustration 6: Population density in persons per square kilometer, by census tract. Census data 2001. [overleaf]


Barcelona, Spain
Population Density by Census Tract
Census data, 2001.
Population Density
Per km2. Census 2001.
0 – 20,000
20,000 – 40,000
40,000 – 60,000
60,000 – 80,000
80,000 – 100,000
100,000 – 120,000
120,000 – 136,200
N
0 1 2 3 4 km

3.5.3 Case-Only Analysis

A case-only analysis was used in this project. The case-only methodology is used to study risk factors that modify the effects of a given type of exposure. Case-only methods can be used to analyze a single, time-fixed modifier of a given time-variable exposure in a set of cases over time. In this study, the exposure variable was high air temperature (HAT+/-) based on a percentile threshold of historical temperature data, and the modifier variable was high relative radiance (HRR+/-) based on the averaged z-scores from the raster data. Cases were deaths registered in the mortality registry and the time units analyzed were days during the study period.

Case-only designs were developed for use in environmental epidemiology, specifically genetic epidemiology, to address gene-environment interactions, but they have also been used to study heat- and pollution-associated mortality in time-series data (*Armstrong 2003; Medina-Ramón et al. 2006; Schwartz 2005b*).

In genetic epidemiology the genotype of the cases is time-fixed, while an exposure (radiation, pollution) is time-variant. One advantage of this type of study design in genetic epidemiology is that controls are not needed, as it is difficult to identify and genotype sufficient controls. The primary advantages of a case-only design in studies of EHE is that the technique is much simpler than a standard time-series analysis and has greater power, as long as certain conditions are met.

For an environmental case-only analysis, a dichotomous time-varying exposure that applies to all study subjects (in this study a dichotomous variable representing high or low air temperature) and a dichotomous modifier that is time-fixed but varies by subject (these can be individual factors, such as sex, or other fixed factors such as housing type – in this study it was a dichotomous variable representing high or low RRZ at the location where the case died) are analyzed by comparing the proportion of cases that are +/- for the time-fixed modifier (HRR) that die on days when the exposure variable is positive (HAT+) or negative (HAT-).

The fundamental dynamic is the following: if a time-fixed factor makes an individual more susceptible to the effects of a time-varying exposure, then the proportion of cases with the fixed factor will be greater when exposed, as compared to periods without the exposure. If HRR has a detrimental effect on mortality on hot days (HAT+), then the proportion of cases with a positive modifier (HRR+) will be higher when a day is HAT+ than when a day is HAT-.

A 2x2 table can be constructed to show the logic in this calculation (Table 5).

Table 5: A 2x2 table of a case-only design with the exposure (High Air Temperature) in rows and the modifier (High Relative Radiance) in columns

		Modifier		***Proportion***	***Odds***
		HRR = 1	***HRR = 0***	***HRR+ deaths***	
Exposure	***HAT = 1***	A	B	A/(A+B)	A/(A+B) / B/(A+B)
	HAT = 0	C	D	C/(C+D)	C / (C+D) / D/(C+D)

If the proportion A/(A+B) is greater than C/(C+D), then this indicates that HRR is more often seen among deaths with high air temperatures (HAT=1) than without high air temperatures (HAT=0). The Odds Ratio is the ratio of the exposed and unexposed odds, or

$$\frac{A/(A+B)/B/(A+B)}{C/(C+D)/D/(C+D)}=\frac{AD}{CB}$$

The null hypothesis is H_0: AD / CB = 1.0. Under the alternative hypothesis, that HRR is a risk factor for death during EHE, the odds ratio would be greater than one. An odds ratio less than one would indicate that HRR had a protective effect on deaths during EHE.

The Odds Ratio can be modeled as a logistic regression (*Schwartz 2005b*). A logistic regression calculates the odds that the dependent variable is equal to one, using a dichotomous variable rather than event counts.

The initial formula would be (*Armstrong 2003*):

$$\log(E(Y_t))=\beta_0+\beta_1 HAT+\beta_2 HRR+\beta_3 HAT\cdot HRR+\beta_4 Tv+\beta_5 S$$

Where Y_t represents the daily counts of deaths, *HAT* is a dichotomous variable indicating whether (1) or not (0) an averaged air temperature measure (minimum, maximum, mean) for a given day is above the 90^{th} percentile of observed temperatures during the study period (April-September, 2000-2003), *HRR* is a dichotomous variable based on the value of the averaged z-score raster at the position of each death, *Tv* is a vector of time-varying confounders and *S* represents any smoothing functions applied. *HRR* = 1 if the relative radiance score (mean z-score) is greater than or equal to zero, and *HRR* = 0 otherwise.

Armstrong (*2003*) describes a series of equations showing how most of the secondary variables in the analysis cancel out. They will be reproduced here with reference to the equation presented above, where *HAT* is the exposure and *HRR* is the modifier. First, we assume that *HRR* = 0 and *HAT* = 0 in subpopulation *i* on day *j* (the terms $ß_1$, $ß_2$ and $ß_3$ drop out):

$$k_0=E(\sum Y_{ij}|HRR_i=0,HAT_j=0)=\sum\left[\exp(\beta_0+\beta_4 Tv+\beta_5 S)\right]$$

then, the case when *HRR* = 0 and *HAT* = 1 (the terms $ß_2$ and $ß_3$ drop out):

$$k_1=E(\sum Y_{ij}|HRR_i=0,HAT_j=1)=\sum\left[\exp(\beta_0+\beta_1+\beta_4 Tv+\beta_5 S)\right]$$

then, the case when *HRR* = 1 but *HAT* = 0 :

$$E(\sum Y_{ij}|HRR_i=1,HAT_j=0)=k_0\exp(\beta_2)$$

and finally, when both *HRR* = 1 and *HAT* = 1:

$$E(\sum Y_{ij}|HRR_i=1,HAT_j=1)=k_1\exp(\beta_2+\beta_3)$$

These formulas are displayed graphically in a 2x2 table in Table 6:

Table 6: A 2x2 table of a case-only design with exposure (High Air Temperature) in rows and modifier (High Relative Radiance) in columns, substituting formulas into cells

		Modifier	
		***HRR* = 1**	***HRR* = 0**
Exposure	***HAT* = 1**	$k_1 exp(ß_2 + ß_3)$	k_1
	***HAT* = 0**	$k_0 exp(ß_2)$	k_0

The odds ratio (AD/CB) for the association between *HAT* and *HRR* in this example is then:

$$OR_{HRR|HAT} = OR_{HAT|HRR} = \frac{k_0 k_1 \exp(\beta_2 + \beta_3)}{k_0 k_1 \exp(\beta_2)} = \exp(\beta_3)$$

The terms for the main effect ($ß_1$), the time-varying confounders ($ß_4$) and other terms ($ß_0$ and $ß_5$) cancel out, leaving only the interaction term $ß_3$. The logistic regression takes the form:

logit(*HAT*=1) = $\alpha + \beta 3 * HRR$

This simplifies the analysis, but means that the technique cannot be used to estimate the effect of the exposure variable, in this case *HAT*, so case-only analysis is generally used as a complementary analytical method. The appropriateness of using this technique in the present study will be discussed below.

An important consideration in a case-only analysis is confounding. Most importantly, the exposure and the modifier variables need to be independent, although this confounding can be controlled for under certain circumstances (*Gatto et al. 2004*). The differential temporal characteristics of the variables (time-fixed vs. time-variable) are usually strong indicators of independence.

In addition, there are two other interactions that can cause important confounding:

1. Interaction of the time-fixed modifier variable with time-variant variables other than exposure.
2. Interaction of the time-variant exposure variable with time-fixed variables other than the modifier of interest.

3.5.3.1 Special Confounding Issues

There are three specific types of confounding that could affect this type of study. In decreasing order of importance, they are: interactions between the exposure and the modifier (in this study these were air temperature and relative surface radiance at specific points in the city); interaction between the modifier and another time-variable parameter (in this study relative surface radiance and some other time-variant variable that would confound air temperature such as flu epidemics, air pollution, etc.); and interaction between the exposure and one or more of the time-fixed variables (in this case air temperatures and one or more of: age, sex, chronic disease, socio-economic status, etc).

In the first case, it seems true that the radiance of an urban area would interact with air temperatures locally, but the relative radiance is taken as a differential from the mean values for the study area, and thus changing the mean air temperatures over the study area would not confound because of the scale differences – the daily mean temperature is calculated as a constant across the whole study area (spatial dimension), while the HRR score varies across the study area. If the differences between relative radiance values from one area to another changed based on different air temperatures, this would be problematic. It would have been interesting to compare the patterns of differential radiance in daytime versus nighttime TIR images, but there were none available for the study area.

As for the second type of confounding, relative radiance could influence some time-variant variables such as cloud formation, which could affect mean temperatures, but again, the scale of the changes wouldn't allow changes in radiance to change average temperatures. In the opposite direction, weather or air temperature could certainly affect overall radiance, but not the differential between areas (unless there were an association between the radiance differential and absolute air temperature, as mentioned above). It is known that ozone production increases with heat (*Schwartz 2005a*), and ozone increases mortality, so this is a potential confounding factor in areas with higher LST.

With regard to the third type of confounding - interactions between the exposure variable and other time-fixed modifiers of interest – it seems unlikely that daily temperatures across the study area could either modify or be modified by the time-fixed variables such as population density, sex, age at death or other unidentified time-fixed factors.

3.5.3.2 Spatial Issues

This study adds a spatial dimension to the concept of orthogonality in study variables. While the case-only study design works because of the orthogonality of the variables across the time dimension, the spatial dimension adds some potential confounding, especially through spatial autocorrelation.

One issue to note from Armstrong (*2003*) is that the time-fixed variable does not have to be strictly time-fixed, but must be "effectively time-invariant", or change over time at a rate that is so much slower than the exposure variable that the modifier is fixed relative to the exposure. Armstrong (*2003*) cites examples such as age, chronic disease and housing type. Spatial scale mismatches of variables should provide a similar level of orthogonality and avoid spatial autocorrelation that may limit the power of the study. Spatial "invariability" is effectively impossible, but important differences in scale are possible.

3.5.3.3 Summary

A summary of the key points involved in the use of a case-only study design:

<u>Advantages</u>

1. Can be used without controls or population denominators.
2. Statistical analysis is less complex than standard time-series methods.
3. Has greater power.
4. Cofactors cancel out under appropriate conditions.

<u>Limitations</u>

1. Exposure (time-variable) and modifier (time-fixed) must be independent, but it is possible to adjust for interaction.

2. Cannot estimate effect of primary exposure variable (*HAT*).

Appropriateness

1. Heat effects in Barcelona have been extensively modeled in other studies, there is no need to repeat this work (*Baccini et al. 2011; Baccini et al. 2008; Matthies and Menne 2006; Saez et al. 2000*).
2. Because the available, spatially located mortality data have only two cofactors, modeling interactions would be limited even using a more complex analytical methodology.
3. The cancellation of cofactors tends to limit the effect of confounding variables.

The analytical method chosen is a good fit for the project, as it is well adapted to the data available (only cases, limited cofactors) and an estimate of the exposure effect is not needed.

A different regression was done for each temperature type (apparent and normal) and measure (minimum, maximum, mean), with the *HAT* variable set to one for days in which the value is above the 90th percentile for the specific temperature measures available. *HRR* will be defined as one only when the z-score registered for the location of death is greater than zero.

3.5.4 Time-Lagged Mortality

It has been proposed, and demonstrated, that heat-related mortality has various distributed effects over time - lagged mortality, the "harvesting" effect (imminent deaths brought forward by heat stress) – and that the shape of these lags is neither linear nor constant along the range of exposures.

Distributed lag non-linear models (DLNM) (*Gasparrini and Armstrong 2010*) were fitted to the data to identify the lagged effects of heat on mortality. Lags of up to thirty days were examined, but the analysis was focused on the first five days of lag.

Distributed lag models were proposed by Schwartz (*2000*) to address the delayed effect of atmospheric pollutants on subsequent mortality and further developed by Armstrong (*2006*) and Gasparrini et al. (*2010*). The advantage of distributed lag models is they can identify both the distribution of mortality over a period of time as well as along the scale of the exposure (temperature, for example), instead of just a single lag point or range for a single exposure.

3.5.5 Regression

Logistic regression in a general linear model (GLM) was used to explore the level of association between air temperatures, areas of high daytime infrared emissions, and mortality. All regression analysis was done using the R statistical package (version 2.13.1) after exporting the processed study data from the PostGIS database to a comma separated file format.

As mentioned above, the dependent variable in a case-only regression model is the time-fixed factor (the risk factor). The basic generic template formula for the regression model is:

HRR ~ HAT (1)

where HRR, the dependent variable, is a dichotomous variable representing whether the death took place in an area with a relative radiance (z-score) above a defined threshold, or not, and HAT is a dichotomous variable representing whether the death took place on a day in which the air temperature measure was above the HAT threshold, defined as a percentile of historic temperatures.

From this basic formula, there were a series of decisions to be made as the final model (or models) was/were built.

1. The threshold for HRR
2. The percentile threshold for HAT
3. The type of temperature to use for defining HAT: Normal or Apparent
4. The temperature measure to use for defining HAT: Minimum, Maximum or Mean
5. The best adjustment for time-lagged mortality
6. Additional variables included in the formula as cofactors
7. Adjustment for spatial autocorrelation

A more detailed template for the regression formula is the following:

HRR_xx ~ *HAT_ttype_yy_lzz* + γ +fitted(*spatial_lag_model*) (2)

where **HRR_xx** is a dichotomous variable representing whether a death occurred in a high radiance (high z-score) area, with values of one being for deaths in areas above a defined z-score (*xx*) threshold and zero otherwise, **HAT_ttype_yy_lzz** is a dichotomous variable representing whether the death took place on a day in which the air temperature measure of a certain type (*ttype*) was above a certain percentile (*yy*) at a certain lag (*zz*), **γ** is an array of explanatory variables or cofactors, and **spatial_lag_model** is a model used to adjust for spatial autocorrelation.

The development of each part of the model will be discussed below, and the final model described.

3.5.5.1 Akaike Information Criterion

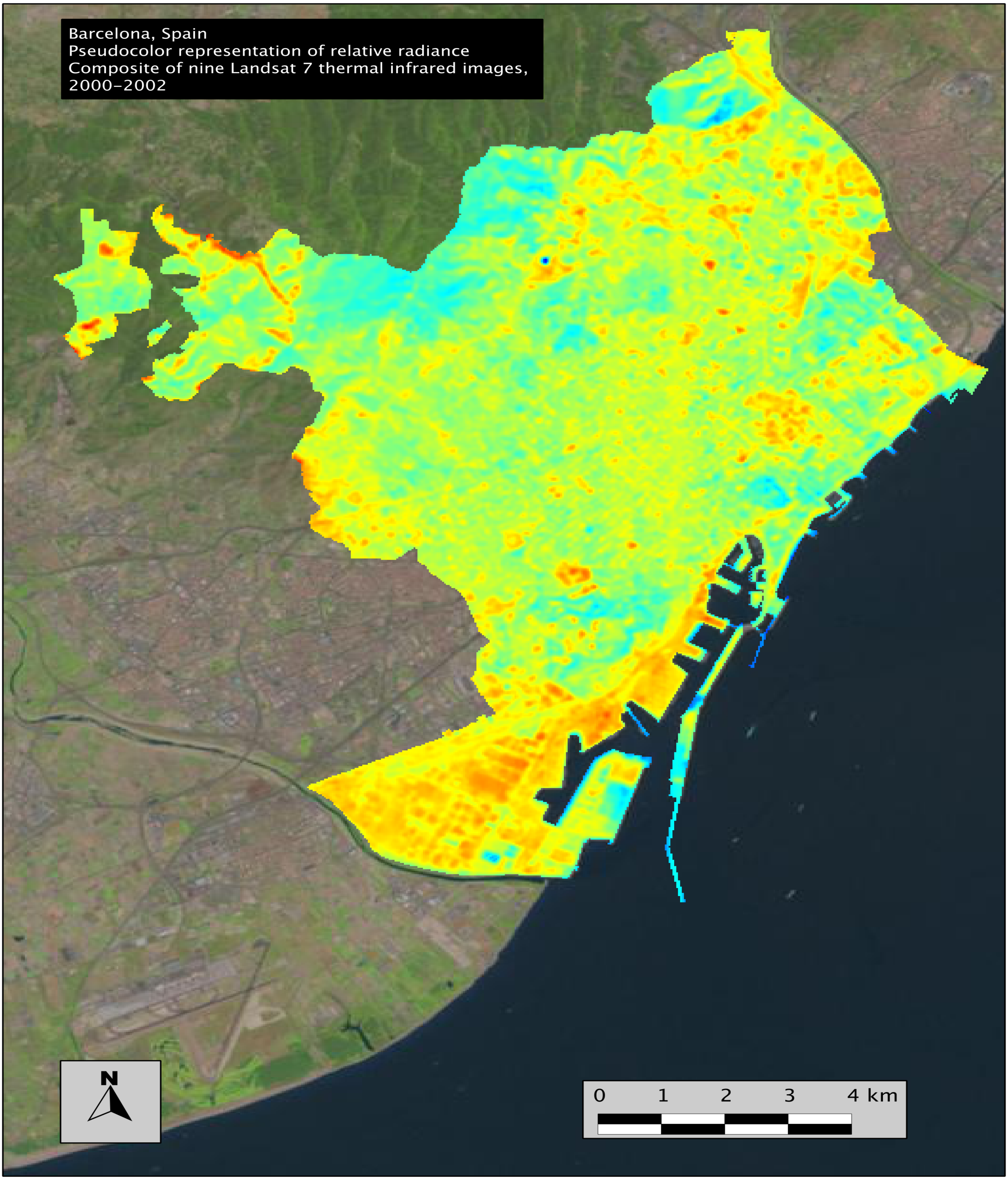
The Akaike Information Criterion (AIC) is used to compare the fit of different regression models against a given set of data. The lower the AIC score, the better the fit of the model to that specific set of data. The corrected AIC (AICc) can be used in cases where the models have a different number of parameters, and so is a preferred method in a study like this, as the model adjusting for spatial lag is included in the regression model as a variable number of individual factors (eigenvectors, see *Controlling for Spatial Autocorrelation*, below). AIC (and AICc) scores cannot be used to compare models fitted to different data sets or using different dependent variables.

When selecting among a set of models, the relative probability that model m_i minimizes estimated information loss when compared to model m_{min} (which has the lowest AIC value of the set) is the following: $exp((AIC_{min} - AIC_i)/2)$. This calculation will be referred to in this paper as ***RelAIC*** **-** the relative probability that the model minimizes information loss. The AICc was used to guide model selection in this study whenever possible (*Logan 2010*).

3.5.5.2 HRR Threshold

A graphic of the averaged RRZ (z-score) for the study area, based on Landsat 7 images, is shown in Illustration 7. It is notable that the individual images (not shown) were remarkably stable, with only minor changes between them. Evidence of this similarity can be seen in the southern part of the image, where the outlines of streets and buildings are clearly visible in this composite, averaging values from nine separate images.

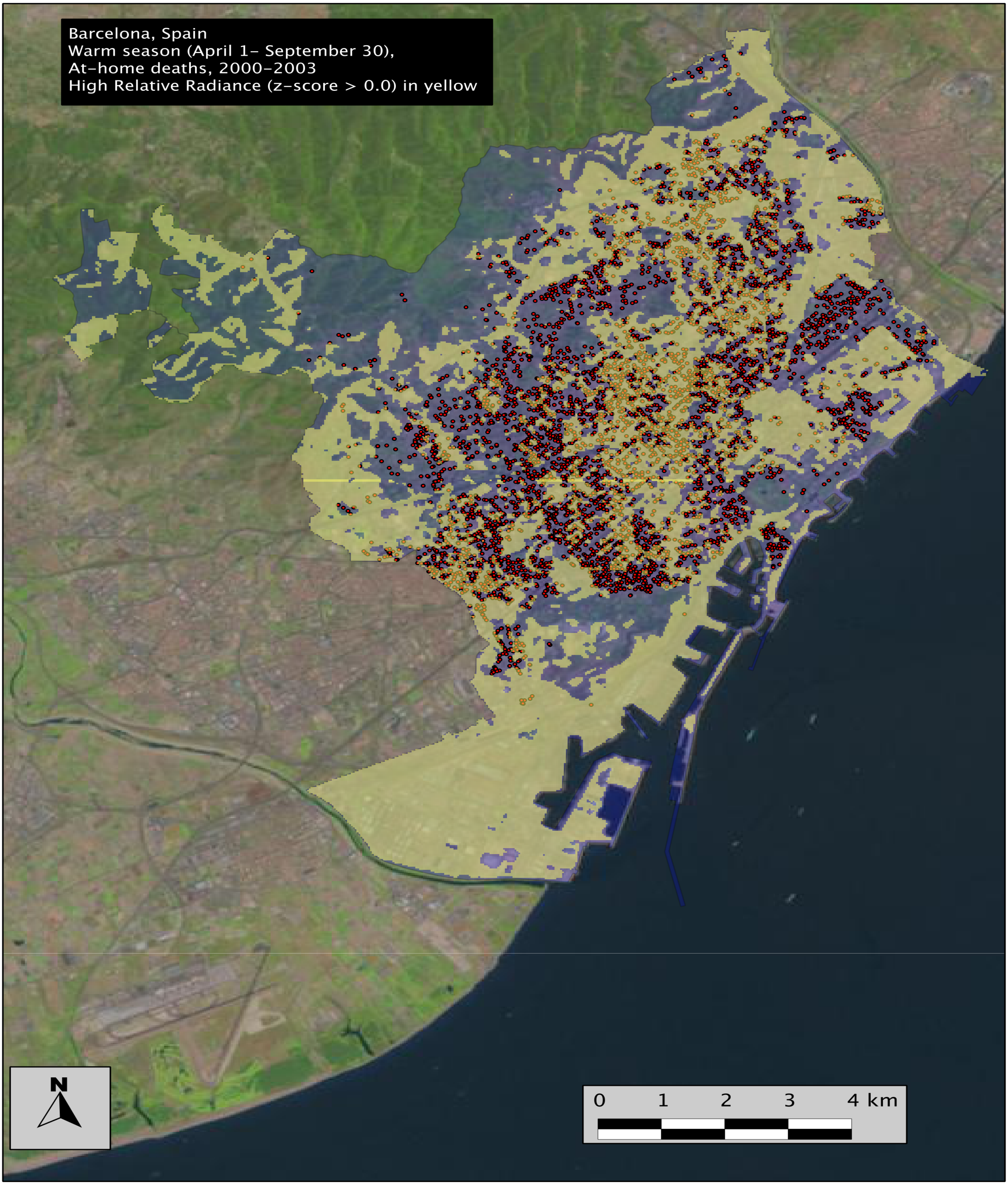
Illustration 7: Psuedocolor representation of average radiance (z-score) from Landsat 7 thermal infrared images (Barcelona, 2000-2002)[overleaf]


Barcelona, Spain
Pseudocolor representation of relative radiance
Composite of nine Landsat 7 thermal infrared images,
2000–2002
N
0 1 2 3 4 km

Three HRR thresholds were considered for this study: z-scores of 0.0, 0.2 and 0.4. At first glance, these may appear to be low thresholds for a type of study designed to use extreme values to examine an association. However, there are several reasons for this choice.

First, the hottest areas of Barcelona tended not to be residential but industrial (warehouses and factories, whose flat roofs heat up quickly), so the majority of the population lived in lower HRR areas. See Illustration 8, showing deaths in the study period and areas with z-scores above zero in red. Second, even those HRR areas which were inhabited had less people.

Illustration 8: At-home deaths April 1 – September 30, 2000-2003 (red). HRR0 areas in yellow. [overleaf]


Barcelona, Spain
Warm season (April 1- September 30),
At-home deaths, 2000-2003
High Relative Radiance (z-score > 0.0) in yellow
N
0 1 2 3 4 km

These two factors skew the distribution curve of RRZ scores across mortality to the left, and reduce case numbers for study when seeking extreme values, as seen in the histogram in Illustration 9 and in Table 7. The red, blue and green lines on the histogram represent HRR scores of 0.0, 0.2 and 0.4, respectively. The respective percentages of deaths above these values are 40%, 23% and 10%. Ninety-nine percent of all cases were below the a z-score value of 1.0. Tables 8-10 show the absolute values and global percentages (in parentheses) of each cell in 2x2 tables for each threshold, based on HAT = 1 when mean temperature is above the 90th percentile of measured temperatures (*tmean90*).

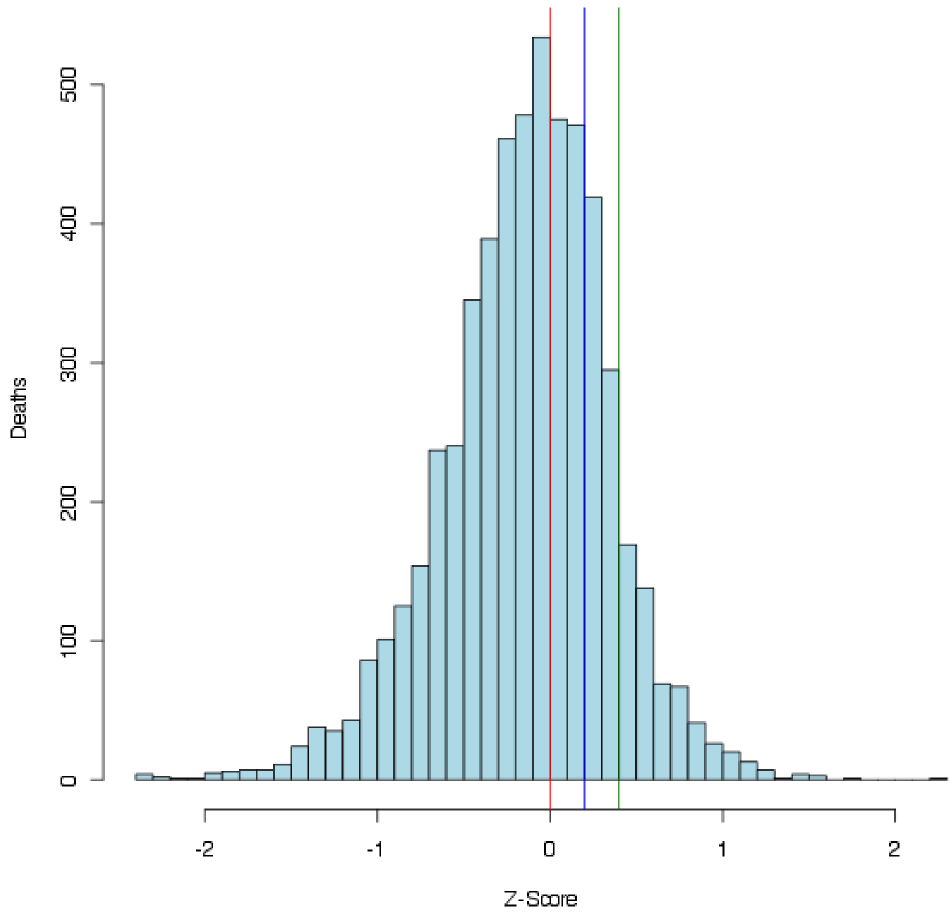


Illustration 9: Histogram of Deaths by RRZ (Z-Score). Thresholds: 0.0 [red], 0.2 [blue], 0.4 [green].

It is also notable that the cooler areas are not heavily inhabited either, as demonstrated in Table 7, where the minimum values are z-scores below -2.3, but 90% of the population lives in areas with z-scores above -0.75. Many of the lowest radiance areas are wooded.

Regression testing and AICc values could not be used to frame this decision because HRR is the dependent variable, and models with different independent variables (essentially different underlying datasets) cannot be compared using the AICc. While the choice of the HRR = 0.0 threshold could potentially limit the results of the study if the effect of MUHI on mortality only exists in the most extreme range of values, it also means that the results of the study are more generalizable, as the threshold chosen includes 40% of at-home deaths in the study period.

Table 7: Percentage Distribution of Deaths by RRZ (Z-Score)

%	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
HRR	-2.38	-0.75	-0.5	-0.35	-0.22	-0.11	0	0.11	0.24	0.4	2.28

Table 8: 2x2 Table for HRR threshold of 0.4

	HAT 1	HAT 0
HRR 1	135 (2.43%)	425 (7.65%)
HRR 0	1217 (21.91%)	3777 (68.01%)

Table 9: 2x2 Table for HRR Threshold of 0.2

	HAT 1	HAT 0
HRR 1	304 (5.47%)	970 (17.46%)
HRR 0	1048 (18.87%)	3232 (58.19%)

Table 10: 2x2 Table for HRR Threshold of 0.0

	HAT 1	HAT 0
HRR 1	2533 (45.61%)	1669 (30.05%)
HRR 0	801 (14.42%)	2533 (45.61%)

Thus, the updated reference regression formula, with the dependent variable as HRR0, is:

HRR_0 ~ *HAT_ttype_yy_lzz +γ +fitted(spatial_lag_model)* (3)

3.5.5.3 HAT temperature types and measures

In selecting reference values for the *HAT_ttypeyy_lzz* portion of the regression formula, there are several different parameters involved. Here *ttype* stands in for the type and measure of temperature (apparent or real temperature; minimum, maximum or mean measures). The *yy* part of the variable name represents the percentile temperature threshold (99^{th}, 95^{th} and 90^{th} percentiles), and the *lzz* portion represents the days of lag used (0, 1, 2, 3 or 5 day lags).

Table 11 shows the number of cases above the various thresholds for each temperature type. These numbers are then divided again into HRR+ and HRR- cells, as seen above in the *HRR Threshold* section, and the counts can become small very quickly.

Table 11: Absolute Case Numbers and Global Percentages by Temperature Measure and Percentile

Temperature Type	90^{th}	95^{th}	99^{th}
Mean (normal)	1352 (24%)	756 (14%)	306 (6%)
Mean (apparent)	1352 (24%)	766 (14%)	260 (5%)
Minimum (normal)	1362 (25%)	660 (12%)	219 (4%)
Minimum (apparent)	1264 (23%)	535 (10%)	189 (3%)
Maximum (normal)	1406 (25%)	727 (13%)	180 (3%)
Maximum (apparent	1629 (29%)	844 (15%)	210 (4%)

Multiple regressions were done to compare models based on the different temperature cutoff points. Selected results are shown in Table 12, and the entire table is in *Appendix 1: Initial Regression Results*. All temperature measures (mean, minimum and maximum) were regressed for both temperature types (normal and apparent) for the three percentage thresholds. The best model was for the 90^{th} percentile of minimum apparent temperatures with two days of lag (see *Modeling Time Lag*, below), with an AIC score of 7470.69. This model is highlighted in yellow in Table 12. Correlations for both the selected model and minimum normal temperature were significant at the 0.05 level.

A rule of thumb for using AIC scores is that models within two points of the minimum have substantial support for their results, otherwise the lower value is significantly better (*Logan 2010*). This study established a threshold of two points or less of AIC difference from the minimum for model consideration, which roughly corresponds to a RelAIC of 0.36 or higher. The models in Table 12 below this threshold are greyed out, to make the results easier to visualize. There are more sophisticated methods of model selection, but they were not used in this study, as this method was considered sufficient. The OR (odds ratio) in the table is explained in the *Case-Only Analysis* section. Simply put, if the OR is greater than one, the odds of death in the portion of the exposed group with the risk factor (HRR) is greater than the odds of death in the exposed group without the risk factor. If the OR is equal to one, then there are no differences between the two groups. If the OR is less than one, the risk factor appears to have a protective effect on the population who have it.

Table 12: Selected regression results for 90th percentile temperatures, before adjustment for spatial autocorrelation

Temperature		Lag	OR	95% CI	P-value	AICc	RelAICc
Type	*Measure*	*days*	*odds ratio*	*confidence*			
Mean	Apparent	0	1.038	0.915-1.176	0.562	7474.634	0.14
Mean	Normal	0	1.038	0.826-1.321	0.562	7474.634	0.14
Minimum	Apparent	0	1.088	0.957-1.236	0.199	7473.322	0.27
Minimum	Normal	0	1.103	0.974-1.25	0.122	7472.588	0.39
Maximum	Apparent	0	1.020	0.906-1.148	0.739	7474.858	0.12
Maximum	Normal	0	0.995	0.879-1.126	0.934	7474.963	0.11
Mean	Apparent	1	1.087	0.959-1.231	0.192	7473.269	0.28
Mean	Normal	1	1.087	0.792-1.273	0.192	7473.269	0.28
Minimum	Apparent	1	1.133	0.998-1.285	0.052	7471.22	0.77
Minimum	Normal	1	1.120	0.989-1.268	0.073	7471.758	0.59
Maximum	Apparent	1	1.051	0.934-1.182	0.407	7474.281	0.17
Maximum	Normal	1	1.029	0.91-1.163	0.652	7474.767	0.13
Mean	Apparent	2	1.114	0.984-1.26	0.088	7472.071	0.5
Mean	Normal	2	1.114	0.803-1.283	0.088	7472.071	0.5
Minimum	Apparent	2	**1.141**	**1.007-1.293**	**0.038**	**7470.69**	**1**
Minimum	Normal	2	**1.140**	**1.006-1.29**	**0.039**	**7470.728**	**0.98**
Maximum	Apparent	2	1.068	0.949-1.202	0.274	7473.774	0.21
Maximum	Normal	2	1.041	0.918-1.179	0.531	7474.578	0.14
Mean	Apparent	3	1.094	0.965-1.238	0.159	7472.987	0.32
Mean	Normal	3	1.094	0.965-1.238	0.159	7472.987	0.32
Minimum	Apparent	3	1.132	0.998-1.283	0.054	7471.26	0.75
Minimum	Normal	3	1.132	1.00-1.281	0.051	7471.157	0.79
Maximum	Apparent	3	1.052	0.934-1.184	0.406	7474.279	0.17
Maximum	Normal	3	1.025	0.905-1.161	0.693	7474.814	0.13

Best model highlighted in yellow. RelAICc threshold for consideration of other models is 0.36. Models with gray text are below RelAICc threshold. Significant results in bold.

These results show that, before applying a correction for spatial autocorrelation, there is no one model that clearly stands out above the others, but rather a gradient centered around minimum apparent temperatures at two days of lag. There is a tendency toward positive correlations in all but one of the models (maximum normal temperature at zero days of lag), and a tendency for the more significant results to have lower AICc scores. The best models were significant at the 0.05 level.

In general, the worst performing models (with AICc differences above two) were those based on average maximum temperatures and higher percentiles (95^{th} and 99^{th}, not shown). In addition, these poorer models had unstable coefficients. The higher AICc values and lower p-values seen in the 90^{th} percentile models made them better suited for the study and more useful for drawing inferences about the data. All p-values below 0.1 were found in the models for minimum temperatures (both real and apparent) above the 90th percentile, with the exception of 90th percentile mean temperatures at two days of lag.

The 90th percentile temperature threshold was selected, again in order to have sufficient numbers for statistical analysis, and because initial regression testing of a series of models showed this to be a more stable data set. At this point, the highlighted model was used as a reference, but only to the extent that it is representative of the set of models shown that had similar AICc scores. These models may not all respond similarly to adjustment for spatial autocorrelation.

Although the two-day lag had some indications in its favor based on AICc scores, the importance of time lags on heat-associated mortality mean that time lags needed to be studied further, using more sophisticated tools, before a decision is made. This analysis is extended in the *Modeling Time Lag* section, below.

The choice between apparent and normal temperatures is also one in which AIC scores are not clear indicators of a good model because the values are so similar. Apparent temperature has the advantage of an underlying physiological explanation, which is that relative humidity is an important factor in physiological stress as it makes temperature regulation through sweating less effective. As such, it potentially models actual heat stress better than a simple measure of air temperature. This measure was also used Baccini et al (*2010; 2008*) in their study of heat effects in 15 european cities (including Barcelona).

The disadvantage of apparent temperature is not directly germane to this current work, which is that it is difficult to predict humidity. This limits the practical applications of applying any results to specific activities, such as for developing HHWWS.

Based on the results of this phase of the decision process, the reference regression formula becomes:

HRR_0 ~ HAT_atmin_90_*lzz +γ +fitted(spatial_lag_model)* (4)

Table 13 shows the cut-points for HAT+ using the 90th percentile of temperatures, in centigrade degrees.

Table 13: Temperature cut-points for 90th percentile of values (1996-2008)

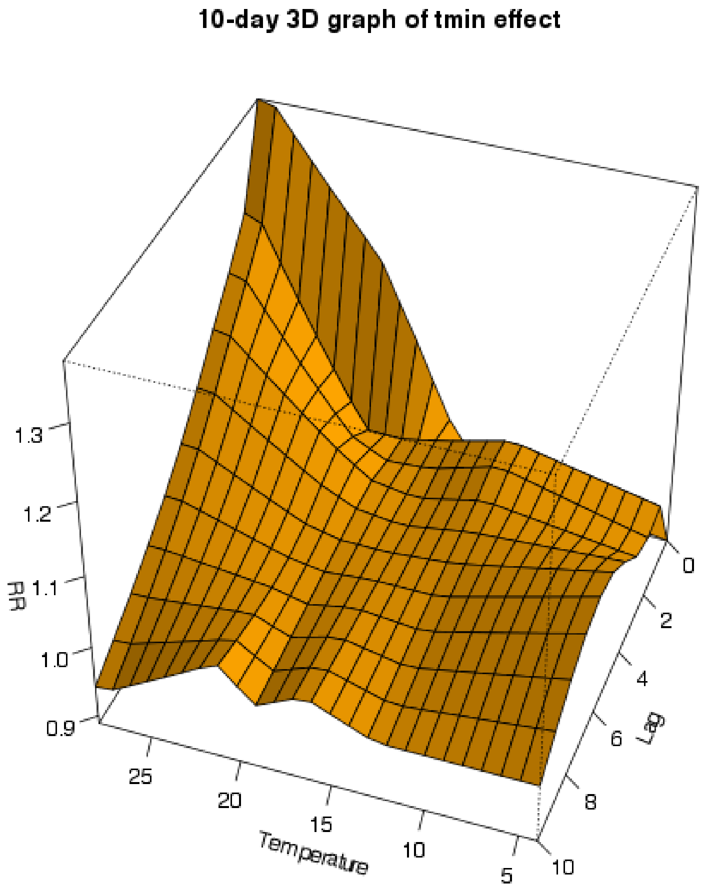
Temperature Type	Mean-90	Min-90	Max-90
Normal	24.40	20.95	29.20
Apparent	21.64	18.33	25.86

3.5.6 Modeling Time Lag

It has been proposed for many years that heat-associated mortality shows non-linear effects over time, affecting mortality days after EHE, both through delayed mortality (*Gosling et al. 2008; Bayentin et al. 2010; Kinney et al. 2008; Metzger et al. 2010*) and the "harvesting" effect of bringing forward imminent deaths, thus reducing subsequent mortality for an undetermined period (*Gosling et al. 2008; Metzger et al. 2010*). Most studies showed lags of 0-3 days (*Gosling et al. 2008*).

The R module DLNM (*Distributed Lag Non-linear Models*) was used to profile the effects of heat across a range of temperatures and lag times, using the study data reorganized into a time series. Details of the method and the code used can be found in *Appendix 2: Details of Methods* , in the *Modeling Time Lag* section. The results indicated that peak mortality risk was roughly one to two days after exposure, but that mortality risk ratios were elevated for several days.

Illustration 10 shows the predicted effects of one of the better performing models from Table 12 (minimum normal temperature above 90th percentile - tmin90) over 10 days of lag.



Axes: Relative risk (RR) of mortality - vertical; Temperature - horizontal; Days of lag - depth

Illustration 10: Minimum Temperature Effects Across 10-day Lag (3D Plot). Data 2000-2003, April-September.

In this image it is clear that peak risk is on the day of initial exposure, and the relative risk (RR) at the highest analyzed temperatures remains elevated for most of the 10-day lag period, but at lower temperatures this is not the case.

The plot also shows only a very rough form of the J-shape typical of mortality graphed by temperature, and seen in the Baccini et al. (*2008*) plot from Illustration 2. This may be because this data was only from the warm season, and because only temperature and daily deaths were used in the regression formula, so it is a much more simple model.

Although this type of three dimensional plot is very good for getting a general idea of how heat impacts mortality across a range of temperatures and lags, it is somewhat difficult to interpret this graphic in detail (such as where the 1.0 RR threshold is crossed along the risk surface).

A more useful graphic is a contour plot of the same model, with different hue intensities of the colors red and blue representing elevated and decreased risk, respectively, as seen in Illustration 11, a contour plot for relative risk over a range of temperatures and five days of lag. The green line in the figure represents the cutoff point for the 90^{th} percentile of minimum temperatures (tmin90=20.95), over five days of lag. In general terms, it can be seen that the lagged effects of heat extend over a greater period of time as the heat exposure increases. It is also apparent that the most intense risk above the tmin90 threshold is at one day of lag. There is a peak at one day of lag, but it also shows that the entire period of highest risk extends from zero to two days of lag, which is consistent with a two-day lag showing the best fit for this data, as the temperatures regressed in this model are an average of the previous two-day period.

Illustrations 13 through 18 show plots for atmean90, tmean90, atmin90, tmin90, atmax90 and tmax90, respectively, for the study period (2000-2003, April-September). There is substantial variation in the distribution of high and low risk areas, but they all show peak risk at one day of lag followed by a period of lower risk. The minimum temperatures both show more risk density across the lower part of the graph, while the two mean temperature plots show a more defined peak of mortality at one day of lag. This may be due to the fact that minimum temperatures take place in the very early morning, and thus mortality could be moved forward with reference to that day's minimum temperatures, as there are probably 20 hours remaining in a day after the minimum temperature is reached.

Illustration 11:

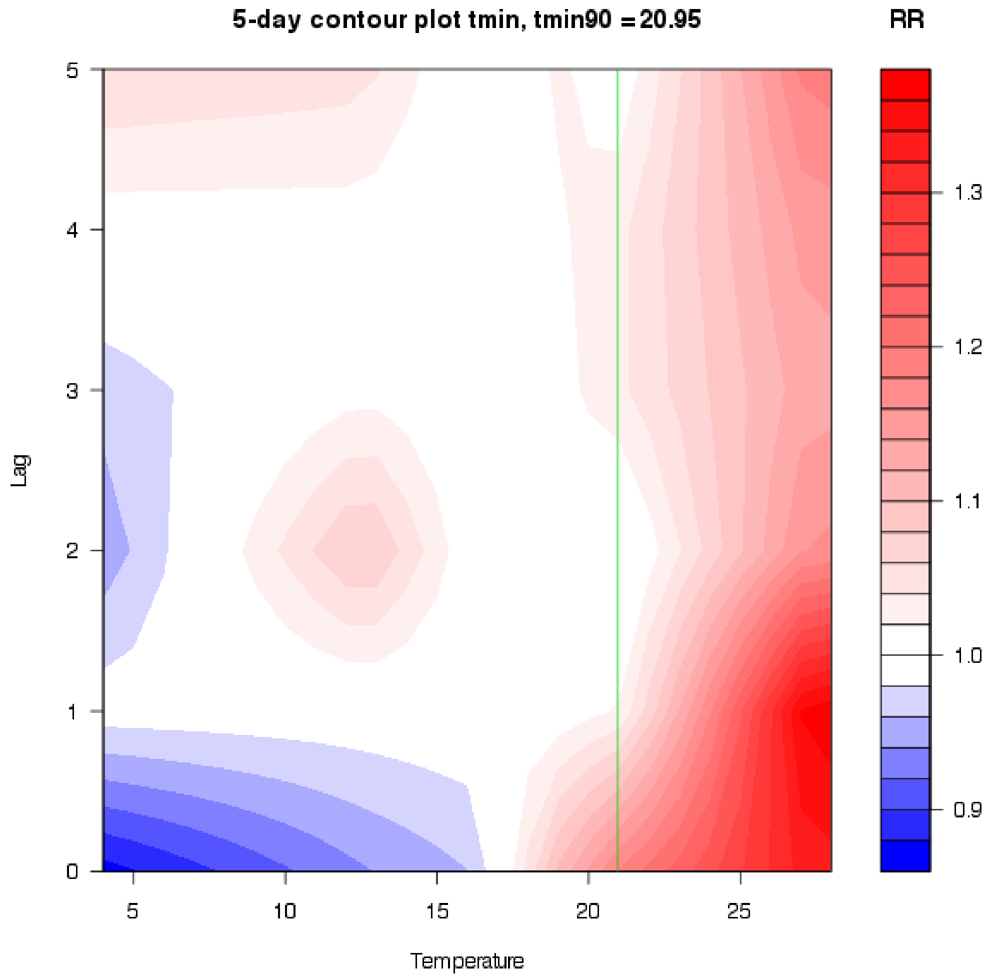


Illustration 12: Minimum Temperature Effects Across 5-day Lag (Contour Plot). Data 2000-2003, April-September, tmin90=20.94, RR=relative risk

Minimum temperatures represent a measure that has been highlighted in earlier research (*Baccini et al, 2010; Baccini et al. 2008*). The theoretical explanation for the impact of high minimum temperatures on health is that this temperature represents the level of physiological respite that an organism gets at night. With high minimum temperatures, it is difficult to recover from daytime heat stress (in the absence of air conditioning), and thus it represents an absolute threshold of stress on the population in a 24-hour period. Thus, the good performance of minimum temperatures measures has a sound theoretical basis for use in the reference model.

The apparent and normal maximum temperatures in Illustrations 17 and 18 show blue areas around the threshold line, indicating a period of decreased mortality risk around two days of lag. One explanation for this decrease in risk would be the harvesting effect, in which the death of extremely vulnerable individuals would be brought forward 1-3 days (in this case), leading to decreased mortality in the period immediately following the highest mortality impact of the EHE.

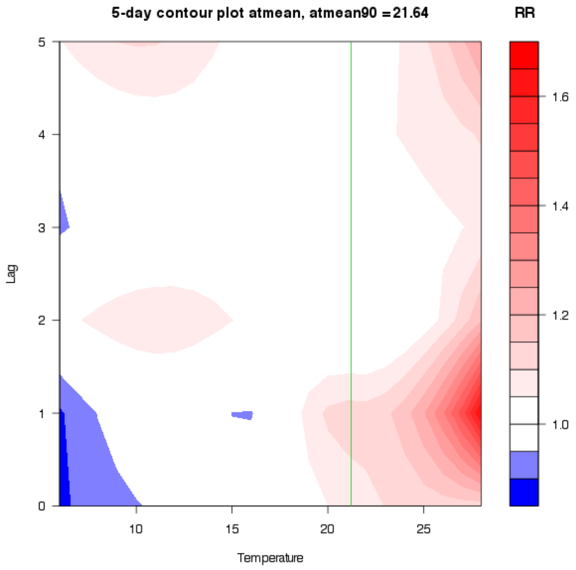


Illustration 13: Apparent Mean Temperature Effects Across 5-day lag. Data 2000-2003, April-September.

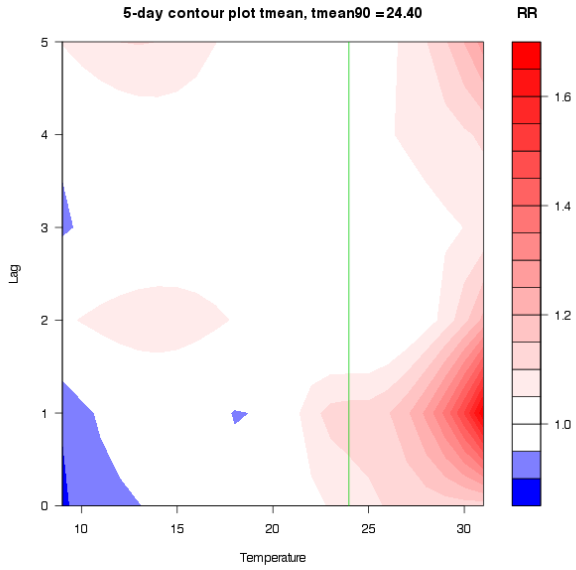


Illustration 14: Mean Temperature Effects Across 5-day lag. Data 2000-2003, April-September.

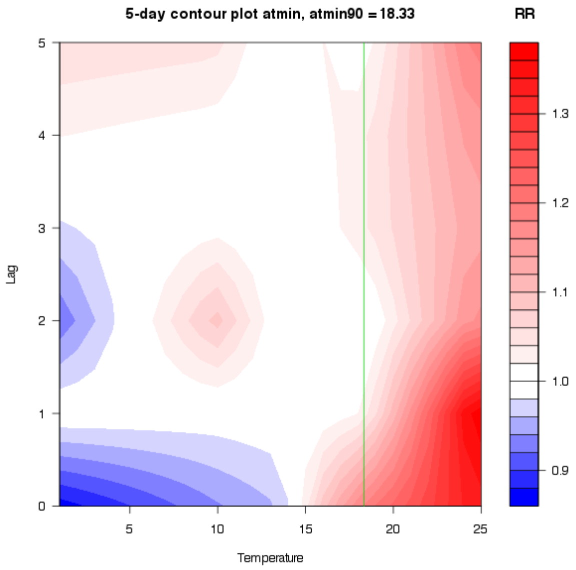


Illustration 15: Apparent Minimum Temperature Effects Across 5-day lag. Data 2000-2003, April-September.

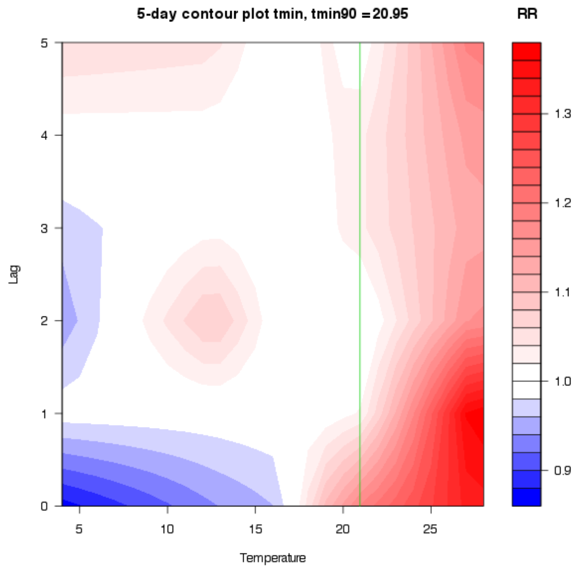


Illustration 16: Minimum Temperature Effects Across 5-day lag. Data 2000-2003, April-September.

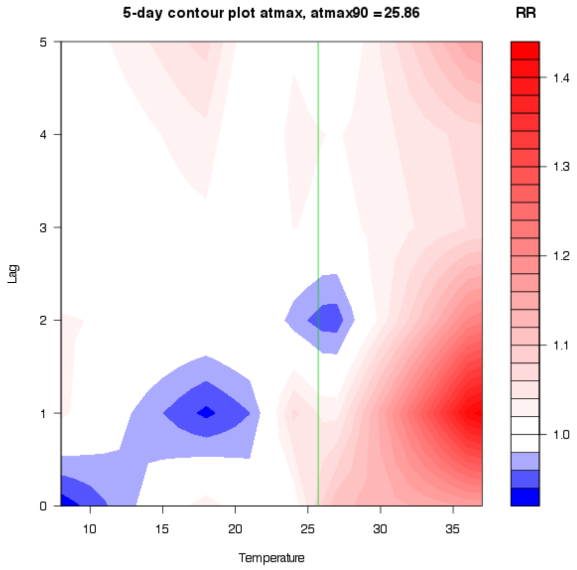


Illustration 17: Apparent Maximum Temperature Effects Across 5-day lag. Data 2000-2003, April-September.

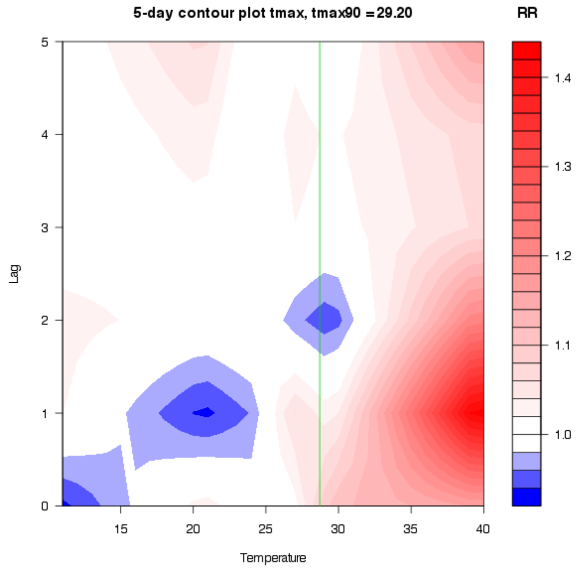


Illustration 18: Maximum Temperature Effects Across 5-day lag. Data 2000-2003, April-September.

These results are consistent with the highest mortality risk being between 0-2 days of lag. The lag adjustment technique used in this study was fairly simple, using measures based on the average temperatures of the *N* previous days. This avoids the complexity of some more sophisticated techniques and still allows for good adjustment of models. Multiple lag measures were not included in a single model to avoid multicollinearity in the regressions.

Thus, the choice of a model with two days of lag has the most support and can be added as the lag value in the reference regression formula:

HRR_0 ~ HAT_atmin_90_l2 $+\gamma+$fitted*(spatial_lag_model)* (5)

3.5.7 Cofactor Selection

Some regression testing was done to select significant factors, but several factors were included in the model due to their real-world relevance. Population density was significant in all of the models tested, as expected, and is an important factor because it can partially represent other data that were not included in this study. Social status can be loosely associated with population density, as crowded conditions are likely to be seen with lower socio-economic levels. Population density can also affect solar gain. High population density means high-rise buildings and limited green space. High-rise buildings can provide shade, but also are made of materials that absorb solar heat, and do so when lower buildings may be protected from direct sunlight.

However, sex and age were not significant in many of the models tested, or lost their significance after adjusting for spatial autocorrelation. Nevertheless, these factors were included in all models because of the important real-world relevance of age – most studies show the highest heat-related mortality in older age groups (*Baccini et al. 2008; Medina Ramon et al. 2006; O'Neill et al. 2009)* - and some studies have found differences by gender (*Basu 2009)* on the impact of heat on health and mortality.

Adding cofactors produces the reference formula:

HRR_0 ~ HAT_atmin_90_l2 +age +sex +pdens_km +fitted*(spatial_lag_model)* (6)

3.5.8 Controlling for Spatial Autocorrelation

Once the basic regression model in (6) was defined, there remained one important element to take into account, which is the effect of spatial autocorrelation. The dependent variable (HRR) and one of the independent variables included in this model (population density) are spatially defined, and others may be spatially autocorrelated (age). There are likely to be important confounding effects that are spatially distributed, such as the effect of increased mortality in areas that are more marginalized or exposed to environmental stressors.

All variables used in the regression models were analyzed for spatial autocorrelation using a calculation of the Moran's I (MI) statistic based on W-weighting of different types of networks (see *Network Selection and Weighting*). The selected results for a distance network with a threshold of 750 meters are shown in Table 14. The 750 meter threshold was chosen for its real-world relevance and ease of calculation. The complete table can be found in *Appendix 3: Moran's I Values*.

Table 14: Moran's I values for selected regression model variables

Network	Distance	Wt	Variable	MI	P-Value
Distance	750	W	Tmin	-0.002	0.892
Distance	750	W	Tmax	-0.002	0.907
Distance	750	W	Tmean	-0.001	0.647
Distance	750	W	Age	0.023	7.88E-050
Distance	750	W	Sex	0.010	1.22E-010

The Moran's I (MI) statistic is a measure of spatial autocorrelation derived from values of a single variable by comparing values of the variable at one location (either point or areal) to values of the same variable in a neighboring location – autocorrelation, or correlation with itself. The property of being a neighbor can be defined in a number of ways, often by setting a distance threshold and identifying all locations within a radius equal to that distance as neighbors. Consistent with Tobler's Law (nearer locations are more similar than less near locations), weights are applied to the correlations with neighboring areas based on their distance. Using a binary coding scheme, neighbors would be assigned a weight of 1 and non-neighbors a weight of 0. There are many different types of networks and weighting schemes that can be applied to them (see the *Network Selection and Weighting* section*).*

Moran's I is similar to a Pearson's correlation coefficient (PCC) in that it produces a statistical measure of correlation for the factor of interest. It is different in that a PCC compares two independent variables, while the MI provides a measure of spatial-autocorrelation of a single variable. A positive MI indicates that the values in the locations examined are more clustered than would be expected in a random distribution, and a negative MI indicates that the values in those locations are more dispersed than would be expected in a random distribution. Because p-values can be calculated for MI, the level of certainty to which a spatial distribution departs from spatial randomness can be quantified and treated as any other one- or two-tailed test (*Paradis 2006; Tiefelsdorf, Griffith and Boots 1999*).

The variables that were significantly spatially autocorrelated were age and sex. The spatial clustering of age seems reasonable - neighborhoods can develop different demographic profiles as they evolve over time. The spatial autocorrelation of sex, though, while not high, is not as easy to explain. One cause could be confounding, as women tend to live longer than men.

When age is regressed against sex, there is a strong correlation (6.302; p-value < 0.0001), indicating that women in the dataset die with an average age that is 6.3 years greater than that of men. The full set of correlations between variables can be found in *Appendix 4: Variable Correlations*.

Overall spatial autocorrelation was controlled for in the regression model using the Moran eigenvector GLM filtering (ME) function (Dray, Legendre and Peres-Neto 2005). This software uses brute force eigenvector selection to eliminate part of the spatial autocorrelation in a general linear model, in this case using a logistic (logit) link. Logistic regression is used to model relationships in which the outcome (dependent) variable is dichotomous.

The ME function adjusts for spatial autocorrelation by identifying the degree to which connected nodes (i.e., the elements in a spatial weighting matrix) are related to each other, and producing additional attributes for the subsequent regression model, which uses those spatial interrelationships to compensate for bias in the model. The downside of this technique is that the brute-force method is resource intensive (several hours or days are required for a single spatial adjustment), and it is difficult to determine the time needed to complete an adjustment. In general, the time needed increases as the square of the number of nodes, but there is a great deal of variability in real world calculations. Multiple models needed to be tested, and it was clear that determining the best model to use was extremely important before committing extensive resources to the final analysis. The arguments to the ME function and the parameters used for the creation and weighting of network objects can be found in *Appendix 2: Details of Methods* , *Network Selection and Weighting.*

3.5.9 Addressing Confounding Variables

Specific factors which can be identified as important mortality confounders or cofactors are: the use of air conditioning (*O'Neill et al. 2005*), air pollution (*Medina-Ramón et al. 2006*), ethnicity (*O'Neill et al. 2009*) and population density (*Johnson et al. 2009*).

3.5.9.1 Air-conditioning

Air-conditioning, as mentioned above, has an enormous effect on heat-related mortality rates. A 2008 energy-usage study by the Spanish Statistics Institute (*Instituto Nacional de Estadistica*), which conducts the census, reported that AC ownership in Catalonia was 36.1% in 2008, slightly above the national average of 35.5%. (*Instituto Nacional de Estadística, 2008*). The Spanish figures were also broken out by income level, with AC ownership ranging from 22.3 in households with income below 1,100 euros a month and reaching 46.2% in households with income greater than 2,700 euros a month.

Information on household AC availability by census tract was not collected in the 2001 census, but has been added to the 2011 census questionnaire. A reasonable expectation is that AC acquisition increases over time, especially after the EHE of 2003. Unfortunately, it was impossible to adjust the data analysis for this factor, and it certainly had some effect on the results.

3.5.9.2 Air Pollution

Most types of air pollution are not highly correlated with temperature, with the exception of ozone. Ozone has a high positive correlation with temperatures, which can lead to confounding if higher ozone levels increase mortality significantly, and it has therefore been a subject of study in heat-associated deaths.

Medina-Ramón et al. (*2006*) adjusted models with and without ozone in their case-only study and found similar results in both. Schwartz (*2005a*) studied the relationship between ozone and mortality risk across 14 US cities, and found that temperature was unlikely to be a confounding factor. Smargiassi et al. (*2009*) found no evidence of association with ozone. Ozone and air pollution data were not available for this study. Nevertheless, collecting ozone data and controlling for daily ozone concentrations is a potential area of further investigation.

3.5.9.3 Ethnicity

The Spanish census data does not include data on ethnicity, only country of birth and immigration status. This study could not control for ethnicity, since the data were not available, but confounding may have affected the robustness of the conclusions. There have been several studies of differential health outcomes in immigrant populations, but this study did not examine immigration status as the immigration status of the cases was unknown.

3.5.9.4 Population Density

Population density has been shown to be a factor in heat-related mortality (*Johnson et al. 2009*). At the same time, Zhang and Wang (*2008*) found a high correlation between population density and the UHI effect, with an R^2-value of 0.9438. In addition, high population densities in urban areas are almost always achieved with high-rise buildings, which necessarily means limited green space, extensive concrete surfaces, and the potential for extra heat gain on the vertical surfaces exposed to the sun. Whether tall buildings increase or decrease air movement depends on local weather conditions. High population density can also stand in for some measures of socioeconomic class, as smaller, more crowded dwellings are less likely to be homes of the wealthy. All of these factors make population density a variable that is likely to be correlated with several of the other study variables.

3.5.10 Network Selection and Weighting

The type of spatial network used and weight coding scheme used are fundamental to the process of adjusting for spatial autocorrelation (Dray, Legendre and Peres-Neto *2005*; Tiefelsdorf 1999). This study used distance networks (DNN) with a neighbor threshold of 750 meters and W-weighting for nodes. The selection process is detailed in *Appendix 2: Details of Methods* , in the *Network Selection and Weighting* section.

3.5.11 Final Model

The reference model developed in the *Regression* section used minimum apparent temperatures above the 90th percentile threshold at two days of lag, with adjustment for spatial autocorrelation being done with the *ME* function, using W-weighted DNN networks. Because the effects of adjustment for spatial autocorrelation were unknown, the entire set of DNN models with a distance threshold of 750 meters were fitted, including lags from lag0 to lag3. Selected results are shown in Table 15, including all models with AICc scores within the range of consideration (up to two points of AIC above the minimum value, or a RelAIC of 0.36 or greater). Full results are shown in *Appendix 5: Spatially-Adjusted Regression Results*.

The best performing model based on AICc scores used normal minimum temperatures at two days of lag, closely followed by the model based on minimum apparent temperatures at two days of lag. Both were significant at the 0.05 level, as was the model for minimum apparent temperature at one day of lag.

Table 15: Selected regression results, after adjustment for spatial autocorrelation

Temperature		Lag	Threshold	OR	95% CI	P-value	AICc	RelAIC
Type	**Measure**	***days***	***meters***					
Apparent	Minimum	1	750	**1.149**	**1.001-1.32**	**0.048**	**6517.011**	**0.899**
Normal	Minimum	1	750	1.131	0.987-1.296	0.076	6517.750	0.621
Apparent	Mean	2	750	1.131	0.987-1.295	0.076	6517.765	0.616
Normal	Mean	2	750	1.131	0.987-1.295	0.076	6517.765	0.616
Apparent	Minimum	2	750	**1.150**	**1.003-1.319**	**0.046**	**6516.910**	**0.945**
Normal	Minimum	2	750	**1.151**	**1.005-1.319**	**0.043**	**6516.797**	**1.000**
Apparent	Minimum	3	750	1.143	0.996-1.312	0.057	6517.295	0.780
Normal	Minimum	3	750	1.137	0.992-1.302	0.065	6517.488	0.708

90th percentile temperatures used. All eigenvector filtering used W weighting with alpha=0.05. Best model highlighted in yellow. The RelAICc threshold for consideration of other models is 0.36. Correlations significant at the 0.05 level bolded.

Both types of minimum scores produced results significant at the 0.05 level, as did the model for minimum apparent temperature at one day of lag. These results are consistent with the DLNM results shown in the *Modeling Time Lag* section, indicating that the highest mortality risk is between one and two days of lag.

There are several characteristics of these results that are important for model selection. First, all coefficients were positive, and they tended to be higher the better the fit of the model (lower AICc score). Second, p-values were more significant with a better model fit as well. Third, all models within the range of consideration (RelAICc > 0.36) are significant or nearly significant (below p=0.10, with the highest being 0.076). What we see is a gradient centered around minimum temperatures and two days of lag, which is entirely consistent with the DLNM results as well.

Based on these spatially-adjusted regression tests, the reference regression model can be finalized, as:

HRR_0 ~ HAT_tmin90_l2 +age +sex +pdens_km +fitted(spatial_lag_model) (7)

While average minimum temperature above the 90^{th} percentile at two days of lag (**tmin90_l2**) is the best fitting model, it is representative of a set of models including minimum apparent temperature above the 90^{th} percentile at two days of lag (**atmin90_l2**), minimum apparent temperature at one day of lag (**atmin90_l1**) and normal and apparent minimum temperatures at three days of lag (**tmin90_l3** and **atmin90_l3**).

3.5.12 Further Testing

Taking this best performing model identified in the *Final Model* section, a series of explorations were done to see how different networks (GAB and RNN) and different distance thresholds for DNN networks affect the strength of the association and the quality of the fit to the data. To this end, the following models were tested.

1. Distance Networks (DNN)
   1. 1000m and 500m thresholds for DNN
   2. The selected model (*tmin90_l2*) against 95^{th} and 99^{th} percentile temperature values
   3. The selected model (*tmin90_l2*) against HRR02 and HRR04 areas
2. Other networks (using a 36% sample [2000 nodes])
   1. Relative Neighbor Network (RNN)
   2. Gabriel Network (GAB)
   3. DNN thresholds for comparison (1000m, 750m, 500m)

3.6 Weaknesses of the Study Design

There were several potential weaknesses in the study as designed. Three important confounding factors (income/socio-economic status, ozone air pollution and the use of air-conditioning) were not used due to the difficulty of obtaining them in the study period.

Adjustment of temperature estimates based on the physical environment at each death, such as interpolation between weather stations based on distance and elevation, or including factors such as solar exposure or wind speed and direction, was not done due to the complexity of the modeling and the extensive data sets required for this type of analysis. The HRR and population density variables may partially compensate for these factors.

More sociodemographic data, such as building socio-economic status (SES) indicators from census data could have helped adjust for the effects of SES on health, as well as compensate to some extent for the disparity in AC use by income level.

A very interesting study by Smargiassi et. al. (*2007*) showed the effectiveness of a method for estimating indoor air temperature based on known characteristics of the built environment and the area around the building. Either of these techniques (estimating indoor air temperatures or adjusting for the area around the building) may have helped in this study. However, the technique used by Smargiassi requires extensive knowledge of the the age and construction type of each building, information which was not available for this study. Neither was there detailed data available on the type of public and private space around buildings, especially vegetation. To an important extent the amount of vegetation and its effect on heat is reflected in the infrared imagery as cooler areas.

4 Results

As described in the previous section, a whole set of DNN models were studied at the 750 meter threshold, and the *tmin90_l2* model showed a significant relationship with HRR0 after adjustment for spatial correlation, with results presented in Table 15 (full results in *Appendix 5: Spatially-Adjusted Regression Results*). These results were consistent with the results of analyzing the dataset with distributed lag non-linear models (see *Modeling Time Lag*, above), and consistent with a 15% greater risk of death in HRR areas of the city of Barcelona during extreme heat events identified by high minimum daily temperatures. The findings are significant at the 0.05 level.

4.1.1 Other Model Choices

Some limited sensitivity testing of the basic model was done by changing selected parameters and fitting a new model with the minimum apparent temperature values at two days of lag. The results are shown in Table 16.

Table 16: Selected results of sensitivity testing of the regression model

Z-Score	HAT percentile	Lag	Coefficient	OR	95% CI	P-value	AICc
0	99th	2	-0.132	0.877	0.647-1.183	0.392	6520.16
0	95th	2	0.124	1.132	0.875-1.228	0.190	6519.18
0.2	90th	2	0.069	1.072	0.918-1.249	0.379	5324.1
0.4	90th	2	0.045	1.046	0.845-1.288	0.677	3294.32

All models used DNN networks with a distance threshold of 750 meters.

The models with HRR0 as a dependent variable and different HAT thresholds show worse AICc scores (roughly in the range of some of the other models in *Appendix 5: Spatially-Adjusted Regression Results*). While these results could be considered counter-intuitive (more extreme values would be expected to produce more pronounced results in a case-only study), the reduction in the sample size can lead to unstable coefficients and non-significant p-values. While the models with HRR=0.2 and HRR=0.4 as dependent variables have lower AICc scores, these are not comparable to the other models presented in this study because the underlying data sets are different. In any case, they are subject to the same problem as the first two models – the small cell size produced by selecting more extreme values can lead to unstable coefficients.

Different weighting schemes were also tested with the key models, using the C and S coding schemes (*Tiefelsdorf et al. 1999*). More details about weighting are in the *Network Weighting* section of *Appendix 2: Details of Methods* . Results are shown in Table 17.

Table 17: Regression results by Weighting Scheme

Model	Wt	Coefficient	OR	95% CI	P-value	AICc
atmin90_l2	C	0.147	**1.158**	**1.009-1.328**	**0.035**	6513.482
tmin90_l2	C	0.147	**1.158**	**1.01-1.327**	**0.036**	6513.555
atmin90_l2	S	0.146	**1.157**	**1.007-1.326**	**0.037**	6506.981
tmin90_l2	S	0.144	**1.155**	**1.009-1.326**	**0.040**	6507.108

The different weighting schemes do not appear to make large differences in the results, although both the S-coding scheme and the C-coding scheme produce lower p-values and AICc values, indicating a better fit to the data. The C-coding scheme clearly has a better fit than the W-coding scheme (more than three points of AICc), while the S-coding scheme is much better than the C-coding (more than six points lower). There also seems to be a gradient here as well, with the globally standardized values (which emphasize the number of connections) showing the most significance, the intermediate S-coding showing somewhat less significance, and the W-coding showing the least. Nevertheless, all results are significant, and very similar.

It seems intuitively important that outlier values not have a disproportionate influence on the results of the analysis, as Tiefelsdorf (*1999*) indicates is the case with W-weighting, but it does not seem to be an important differential factor in the results for this data set (at the 750 meter threshold).

5 Discussion

Other studies have also examined mortality associated with MUHI, primarily using areal data, as most health data sets do not include specific addresses. The one study with the best resolution was done in Montreal, by Smargiassi, et al. (*2009*) using the geographical centroids of postal codes. Postal codes in Montreal cover an area roughly equivalent to one side of a city block, with approximately 50 residents each. Their study also used Landsat 7 images, in their case used to derive estimated LST directly, and these estimated temperatures were stratified into groups. The highest-LST group (above the 75^{th} percentile) showed increased mortality risk. Due to cloud cover, they were only able to identify two adequate Landsat images from a 13 year period.

Tomlinson et al (*2011*) have published a detailed method for including MUHI data in a health risk assessment, but the method was not tested against specific mortality data.

This study is the first to use a mortality registry identifying deaths by address, and due to the relative lack of clouds in Barcelona, had nine high quality TIR images of the study area over a three-year period. It is also the first study of MUHI to use eigenvector filtering to partially compensate for spatial autocorrelation.

The results that were obtained are consistent with a real effect – that mortality increases in areas that appear hotter on infrared imagery during extreme heat events. Given all that is known about the effects of heat on health and mortality, it is only reasonable to assume that areas that become hotter during the day, as demonstrated by their radiance in the infrared range, would be associated with greater mortality.

The relative lack of humidity in Barcelona may have been an important factor in the findings. Dry air transmits heat less easily, potentially making radiant heat a more important factor in the subjective experience of cases than it might be in a more humid area, such as Montreal. The findings of this study may not be directly applicable to other areas with different atmospheric conditions.

There remains the possibility that the significant results identified could be confounded, especially by a combination of social class and air-conditioning use. The former could be associated with poorer health, older and less well maintained housing stock, and poorer neighborhood conditions (less trees, for example). This factor could also have a direct influence on the availability of air-conditioning. The net effect is that people living in these areas would be more exposed to the effects of heat, as compared to residents in higher-income areas, who would be better protected from the effects of heat by better insulated buildings, more green space, and air-conditioning.

6 Conclusion

This study attempted to answer the question: Are micro-urban heat islands associated with an increased risk of mortality during extreme heat events? It also posed a sub-question: Which temperature measures are the best predictors of this association, if any?

Based on the results of the statistical analysis and the data used, the answer to the first question is yes. The null hypothesis that the odds ratio for the study data is equal to one can be rejected with significant confidence, as the 95% confidence intervals of the significant odds ratio do not overlap the value of one. These significant differences are robust across different network weighting schemes and a small range of distance thresholds.

In response to the second question, two-day moving averages of minimum temperatures above the 90th percentile of normal temperatures showed the best fit to the data, based on AICc values. They were associated with a significant, 15% increase in the odds of mortality during extreme heat events in micro-urban heat islands (MUHI) identified on satellite thermal infrared (TIR) images. Both apparent and normal minimum temperatures performed similarly. The two-day averages show that lagged mortality measures are a better estimate of true heat-related mortality than a simple association between temperatures and mortality on the same day.

Furthermore, the sensitivity testing done of different temperatures and radiance thresholds indicates that the parameter choices made in the initial regression model were good ones, and that these results were robust across different distance thresholds and weighting schemes. The testing of samples across different network types was inconclusive, although it did indicate that GAB and RNN network results could provide an interesting perspective, if they could be completed in a timely manner. More testing might better profiled the behavior of the dataset. Unfortunately, the computationally intensive nature of the eigenvector filtering approach means that it was difficult to explore the data further, especially at lower distance thresholds and with other network types. An additional methodological observation is that the use of eigenvector filtering in large data sets (> 10,000 node networks) may not be possible with the types of computing hardware available today.

6.1 Practical Implications

Extreme heat events are complex, difficult to identify clearly, and increase mortality in populations that are exposed to them directly. Furthermore, global climatic changes indicate that these events will appear with increasing frequency and intensity (*Baccini et al. 2010; Gosling et al. 2008; Meehl and Tibaldi 2004; O'Neill et al. 2009*).

The results of this study have implications for interventions at two levels: urban planning and public health interventions to target vulnerable individuals. Urban planning interventions to create “cooler” neighborhoods (reflective rooftops and pavement, more trees and green space, development that takes advantage of wind patterns, etc.) can reduce the heat associated with MUHIs.

Public health interventions, which already exist to address heat, can add areas identified as having HRR to their vulnerability maps. Heat emergency response efforts can redirect extra resources to these areas. Workers on the ground can also help identify buildings that may be more vulnerable to heat because they are older, exposed to sunlight, poorly insulated, etc. Existing HHWWS may want to consider how minimum temperatures are weighted in their existing risk formulas, as this study indicates that they are the best measures.

While Landsat 7 is no longer provides coverage of Barcelona, there are many other options for infrared imaging (aerial overflights, other satellite systems, etc.), and the cost of remote sensing would be easily justifiable in densely populated urban areas.

6.2 Caveats

While these results are certainly indicative, there are a number of issues that could limit their practical applications. The most important would be a confounding effect by a variable such as socio-economic status, either linked directly to greater mortality in HRR areas, or indirectly by influencing the differential use of air conditioning by income level (which the INE survey has shown is the case in Spain). If this could explain a large portion of the association between HRR and increased mortality on hot days, the consequent policy implications are different.

Second, this association was identified in a period before and during an important series of extreme heat events in 2003. Both increased air-conditioning and public health initiatives have been implemented in the intervening time, and it is not clear that this mortality association would still be found today.

6.3 Potential improvements and further study

There are a number of aspects of the study that could be improved, and further work in this area could provide more insight into the spatial dynamics of heat and mortality, at least in Barcelona. Of course, specific processes could be improved (such as geocoding) or new data sets could be sought (such as other TIR satellite or aerial images).

The most important issue would be to address the potential confounding effect of socio-economic status (SES) and/or income by using them to adjust for confounding in the regression analysis. The use of air conditioning can't be adjusted for directly until the 2011 census results are available, but income levels could provide a rough gradient of potential use.

Geocoding errors could significantly reduce the ability to detect mortality clustering, and association with MUHI and demographic information. Since standard commercial sources are used (Google and Yahoo), there may be other alternatives that would provide better results, such as address databases maintained by the *Ajuntament* of Barcelona and the *Generalitat* of Catalonia.

Another important improvement would be an increase in the sample size. Completing the geocoding of the remaining 9,000 addresses of the 50,000 “clean” addresses could help with this. A more useful proposition might be incorporating mortality data from Badalona and El Hospitalet, two cities on either side of Barcelona, to the northeast and west, forming a single, contiguous urban area. These two cities have similar profiles to some of the HRR areas in the city of Barcelona: they are more industrial, have higher radiance values, and tend to have a lower socioeconomic profile (although this factor has not been shown to be the case in the data set used in this study, it is merely speculation at this point). On the other hand, a much larger set of cases could create problems with the eigenvector filtering technique.

In a very practical sense, it is also important to identify new sources of high-resolution thermal infrared images in the study area, now that Landsat 7 is unavailable, to follow up on the findings of this study and develop a more up-to-date vulnerability map for the area.

References

Armstrong, B (2003) "Fixed Factors That Modify the Effects of Time-Varying Factors: Applying the Case-Only Approach." *Epidemiology* 14 (4) (July 1): 467-472

Armstrong, B (2006) "Models for the Relationship Between Ambient Temperature and Daily Mortality." *Epidemiology* 17 (6) (November): 624-631.

Baccini, M, Kosatsky, T, Analitis, A, Anderson, HR, D'Ovidio, M, Menne, B, Michelozzi, P, Biggeri, A (2011) "Impact of heat on mortality in 15 European cities: attributable deaths under different weather scenarios." *Journal of Epidemiology and Community Health*, vol. 65, no. 1, pp. 64-70.

Baccini, M, Biggeri, A, Accetta, G, Kosatsky, T, Katsouyanni, K, Analitis, A, Anderson, HR (2008) "Heat effects on mortality in 15 European cities." *Epidemiology (Cambridge, Mass.)* 19 (5) (September): 711-719.

Basu, R (2009) "High ambient temperature and mortality: a review of epidemiologic studies from 2001 to 2008." *Environmental Health*, vol. 8, no. 1, p. 40.

Basu, R and Samet JM (2002) "Relation between Elevated Ambient Temperature and Mortality: A Review of the Epidemiologic Evidence." *Epidemiologic Reviews* 24 (2) (December 1): 190 -202.

Bayentin, L, El Adlouni, S, Ouarda, T, Gosselin, P, Doyon, B and Chebana, F (2010) "Spatial variability of climate effects on ischemic heart disease hospitalization rates for the period 1989-2006 in Quebec, Canada." *International Journal of Health Geographics* 9 (1): 5.

Bouchama, A, Dehbi, M, Mohamed, G, Matthies, F et al. (2007) "Prognostic factors in heat wave related deaths: a meta-analysis." *Archives of Internal Medicine*, vol. 167, no. 20, pp. 2170-2176.

Brulle, RJ and Pellow, DN (2006) "Environmental justice: human health and environmental inequalities." *Annual Review of Public Health* 27: 103-124.

Curriero, FC (2002) "Temperature and Mortality in 11 Cities of the Eastern United States." *American Journal of Epidemiology* 155 (1) (January): 80-87.

Dray, S, Legendre, P and Peres-Neto, PR (2006) "Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM)." *Ecological Modelling* 196 (3-4) (July 25): 483-493.

Fouillet, A, Rey, G, Wagner, V, Laaidi, K et al. (2008) "Has the impact of heat waves on mortality changed in France since the European heat wave of summer 2003? A study of the 2006 heat wave." International Journal of Epidemiology, vol. 37, no. 2, pp. 309 -317.

Gasparrini A and Armstrong B (2010) dlnm: Distributed Lag Non-linear Models. R package version 1.3.2.

Gasparrini, A, Armstrong, B and Kenward, MG (2010) "Distributed lag non-linear models." *Statistics in Medicine* 29 (21) (September 20): 2224-2234.

Gatto, NM, Campbell, UB, Rundle, AG and Ahsan, H (2004) "Further development of the case-only design for assessing gene–environment interaction: evaluation of and adjustment for bias." *International*

Journal of Epidemiology 33 (5) (September 1): 1014 -1024.

Gosling, SN, Lowe, JA, McGregor, GR, Pelling, M, and Malamud, BD (2008) "Associations between elevated atmospheric temperature and human mortality: a critical review of the literature." *Climatic Change* 92 (3-4) (August): 299-341.

Hajat, S, Sheridan, SC, Allen, MJ, Pascal, M et al. (2010) "Heat-health warning systems: a comparison of the predictive capacity of different approaches to identifying dangerously hot days." *American Journal of Public Health*, vol. 100, no. 6, pp. 1137-1144.

Hansen, A, Bi, P, Nitschke, M, Ryan, P, Pisaniello, D and Tucker, G (2008) "The effect of heat waves on mental health in a temperate Australian city." *Environmental Health Perspectives* 116 (10) (October): 1369-1375.

Instituto Nacional de Estadística (2008) *Encuesta de Hogares y Medio Ambiente 2008.* http://www.ine.es/jaxi/tabla.do?path=/t25/p500/2008/p01/l0/&file=01013.px&type=pcaxis&L=0.

Accessed: September 15, 2011

Iñiguez, C, Ballester, F, Ferrandiz, J, Pérez-Hoyos, S, Sáez, M, López, A (2010) "Relation between temperature and mortality in thirteen Spanish cities." *International Journal of Environmental Research and Public Health*, vol. 7, no. 8, pp. 3196-3210.

Johnson, DP, Wilson, JS and Luber, GC (2009) "Socioeconomic indicators of heat-related health risk supplemented with remotely sensed data." *International Journal of Health Geographics*, vol. 8, no. 1, p. 57.

Kestens, Y, Brand, A, Fournier, M, Goudreau, S, Kosatsky, T, Maloley, M and Smargiassi, A (2011) "Modelling the variation of land surface temperature as determinant of risk of heat-related health events." *International Journal of Health Geographics* 10 (1): 7.

Kinney, PL, O'Neill, MS, Bell, ML and Schwartz, J (2008) "Approaches for estimating effects of climate change on heat-related deaths: challenges and opportunities." *Environmental Science & Policy*, vol. 11, no. 1, pp. 87-96.

Klonoff, EA (2009) "Disparities in the provision of medical care: an outcome in search of an explanation." *Journal of Behavioral Medicine* 32 (1) (January): 48-63.

Kovats, RS and Hajat, S (2008) "Heat Stress and Public Health: A Critical Review." *Annual Review of Public Health*, vol. 29, no. 1, pp. 41-55.

Logan, Murray (2010) *Biostatistical design and analysis using R: A practical guide*. Chichester, UK ;Hoboken, NJ: Wiley-Blackwell.

Luber, G and McGeehin, M (2008) "Climate Change and Extreme Heat Events." *American Journal of Preventive Medicine*, vol. 35, no. 5, pp. 429-435.

Martin-Latry, K, Goumy, MP, Latry, P, Gabinski, C, Bégaud, B, Faure, I and Verdoux, H (2007) "Psychotropic drugs use and risk of heat-related hospitalisation." European Psychiatry 22 (6) (September): 335-338.

Matthies, F, and Menne, B (2009) "Prevention and management of health hazards related to heatwaves." *International Journal of Circumpolar Health* 68 (1) (February): 8-22.

Medina-Ramón, M, Zanobetti, A, Cavanagh, DP and Schwartz, J (2006) "Extreme temperatures and mortality: assessing effect modification by personal characteristics and specific cause of death in a multi-city case-only analysis." Environmental Health Perspectives, vol. 114, no. 9, pp. 1331-1336.

Meehl, GA and Tebaldi C (2004) "More Intense, More Frequent, and Longer Lasting Heat Waves in the 21st Century." *Science* 305 (5686): 994 -997.

Metzger, KB, Ito, K and Matte, TD (2010) "Summer heat and mortality in New York City: how hot is too hot?" *Environmental Health Perspectives* 118 (1) (January): 80-86.

Nichol, JE, Wing YF, Ka-se L, and Man SW (2009) "Urban heat island diagnosis using ASTER satellite images and 'in situ' air temperature." *Atmospheric Research* 94 (2) (October): 276-284.

O'Neill, MS, Carter, R, Kish, JK, Gronlund, CJ et al. (2009) "Preventing heat-related morbidity and mortality: new approaches in a changing climate." Maturitas, vol. 64, no. 2, pp. 98-103.

O'Neill, MS, Zanobetti, A and Schwartz, J (2005) "Disparities by race in heat-related mortality in four US cities: the role of air conditioning prevalence." *Journal of Urban Health: Bulletin of the New York Academy of Medicine* 82 (2) (June): 191-197.

Paradis, E (2006) *Analysis of phylogenetics and evolution with R*. Springer.

Rué, M, Cabré, X, Soler-González, J, Bosch, A Almirall, M and Serna, M (2008) "Emergency hospital services utilization in Lleida (Spain): A cross-sectional study of immigrant and Spanish-born populations." *BMC Health Services Research* 8 (1): 81.

Saez, M, Sunyer, A, Tobias, A, Ballester, F and Antó, JM (2000) "Ischaemic heart disease mortality and weather temperature in Barcelona, Spain." *The European Journal of Public Health* 10 (1) (March 1): 58 -63.

Schwartz, J (2000) "The distributed lag between air pollution and daily deaths." *Epidemiology (Cambridge, Mass.)* 11 (3) (May): 320-326.

Schwartz, J (2005a) "How Sensitive Is the Association between Ozone and Daily Deaths to Control for Temperature?" *Am. J. Respir. Crit. Care Med.* 171 (6) (March 15): 627-631.

Schwartz, J (2005b) "Who is sensitive to extremes of temperature?: A case-only analysis." *Epidemiology (Cambridge, Mass.)* 16 (1): 67-72

Smargiassi, A, Fournier, M, Griot, C, Baudouin, Y, and Kosatsky, T (2007) "Prediction of the indoor temperatures of an urban area with an in-time regression mapping approach." *J Expos Sci Environ Epidemiol* 18 (3) (June 20): 282-288.

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

Tiefelsdorf, M, Griffith, DA and Boots, B (1999) "A variance-stabilizing coding scheme for spatial link matrices." *Environment and Planning A* 31 (1). Environment and Planning A: 165-180.

Waller, LA and Gotway, CA (2004) *Applied spatial statistics for public health data*. Wiley-IEEE, July 1.

Whaley, AL (2003) "Ethnicity/race, ethics, and epidemiology." *Journal of the National Medical*

Association 95 (8) (August): 736-742.

Williams, DR, Neighbors, HW and Jackson, JS (2008) "Racial/Ethnic Discrimination and Health: Findings From Community Studies." *Am J Public Health* 98 (Supplement 1) (September 1): S29-37.

Zhang, J and Wang Y (2008) "Study of the Relationships between the Spatial Extent of Surface Urban Heat Islands and Urban Characteristic Factors Based on Landsat ETM+ Data." *Sensors* 8 (November 20): 7453-7468.

Appendix 1: Initial Regression Results

Table 18: Regression results before adjustment for spatial autocorrelation

Temperature		Percentile	Lag	Coefficient	OR	95% CI	P-value	AICc	RelAIC
type	***measure***								
Apparent	Mean	99th	0	0.073	1.076	0.834-1.384	0.572	7474.661	0.138
Apparent	Mean	99th	1	-0.005	0.995	0.773-1.277	0.970	7474.979	0.118
Apparent	Mean	99th	2	0.015	1.015	0.79-1.3	0.904	7474.966	0.119
Apparent	Mean	99th	3	0.021	1.022	0.783-1.328	0.874	7474.955	0.119
Apparent	Mean	99th	5	0.028	1.028	0.765-1.375	0.853	7474.946	0.120
Apparent	Mean	99th	10	0.089	1.093	0.779-1.525	0.604	7474.712	0.135
Apparent	Minimum	99th	0	0.136	1.146	0.852-1.534	0.364	7474.160	0.177
Apparent	Minimum	99th	1	0.021	1.021	0.768-1.352	0.885	7474.959	0.119
Apparent	Minimum	99th	2	0.016	1.016	0.743-1.381	0.921	7474.970	0.118
Apparent	Minimum	99th	3	-0.008	0.992	0.728-1.344	0.961	7474.978	0.118
Apparent	Minimum	99th	5	-0.035	0.965	0.674-1.37	0.844	7474.942	0.120
Apparent	Minimum	99th	10	0.084	1.088	0.743-1.584	0.661	7474.789	0.130
Apparent	Maximum	99th	0	-0.014	0.986	0.742-1.305	0.924	7474.971	0.118
Apparent	Maximum	99th	1	0.025	1.026	0.787-1.332	0.849	7474.944	0.120
Apparent	Maximum	99th	2	-0.022	0.978	0.743-1.282	0.875	7474.955	0.119
Apparent	Maximum	99th	3	-0.030	0.970	0.728-1.287	0.836	7474.937	0.120
Apparent	Maximum	99th	5	0.040	1.041	0.744-1.447	0.814	7474.925	0.121
Apparent	Maximum	99th	10	0.068	1.070	0.781-1.46	0.670	7474.800	0.129
Apparent	Mean	95th	0	0.024	1.024	0.876-1.196	0.764	7474.890	0.123
Apparent	Mean	95th	1	0.026	1.027	0.88-1.197	0.739	7474.869	0.124
Apparent	Mean	95th	2	0.029	1.029	0.882-1.2	0.713	7474.845	0.126
Apparent	Mean	95th	3	0.024	1.024	0.879-1.192	0.761	7474.888	0.123
Apparent	Mean	95th	5	0.023	1.023	0.872-1.198	0.780	7474.903	0.122
Apparent	Mean	95th	10	-0.047	0.954	0.806-1.128	0.585	7474.681	0.137
Apparent	Minimum	95th	0	0.092	1.097	0.914-1.314	0.319	7473.992	0.193
Apparent	Minimum	95th	1	0.059	1.061	0.882-1.275	0.527	7474.582	0.144
Apparent	Minimum	95th	2	0.065	1.067	0.889-1.28	0.483	7474.490	0.150
Apparent	Minimum	95th	3	0.010	1.010	0.838-1.214	0.918	7474.970	0.118
Apparent	Minimum	95th	5	-0.015	0.985	0.798-1.214	0.890	7474.961	0.119
Apparent	Minimum	95th	10	0.005	1.005	0.799-1.26	0.968	7474.979	0.118
Apparent	Maximum	95th	0	-0.060	0.942	0.81-1.094	0.435	7474.369	0.160
Apparent	Maximum	95th	1	-0.028	0.972	0.833-1.133	0.718	7474.850	0.126
Apparent	Maximum	95th	2	-0.019	0.981	0.844-1.139	0.802	7474.917	0.121

Temperature		Percentile	Lag	Coefficient	OR	95% CI	P-value	AICc	RelAIC
type	*measure*								
Apparent	Maximum	95^{th}	3	-0.035	0.965	0.831-1.119	0.640	7474.762	0.131
Apparent	Maximum	95^{th}	5	-0.013	0.987	0.848-1.147	0.863	7474.950	0.119
Apparent	Maximum	95^{th}	10	-0.041	0.960	0.818-1.124	0.611	7474.721	0.134
Apparent	Mean	90^{th}	0	0.037	1.038	0.915-1.176	0.562	7474.644	0.139
Apparent	Mean	90^{th}	1	0.083	1.087	0.959-1.231	0.192	7473.280	0.275
Apparent	Mean	90^{th}	2	0.108	1.114	0.984-1.26	0.088	7472.082	0.501
Apparent	Mean	90^{th}	3	0.090	1.094	0.965-1.238	0.159	7472.998	0.317
Apparent	Mean	90^{th}	5	0.049	1.051	0.927-1.191	0.439	7474.383	0.159
Apparent	Mean	90^{th}	10	0.011	1.011	0.886-1.153	0.866	7474.952	0.119
Apparent	Minimum	90^{th}	0	0.084	1.088	0.957-1.236	0.199	7473.332	0.268
Apparent	Minimum	90^{th}	1	0.125	1.133	0.998-1.285	0.052	7471.231	0.767
Apparent	**Minimum**	**90^{th}**	**2**	**0.132**	**1.141**	**1.007-1.293**	**0.038**	**7470.701**	**1.000**
Apparent	Minimum	90^{th}	3	0.124	1.132	0.998-1.283	0.054	7471.271	0.752
Apparent	**Minimum**	**90^{th}**	**5**	**0.131**	**1.140**	**1.004-1.294**	**0.043**	**7470.888**	**0.911**
Apparent	Minimum	90^{th}	10	0.030	1.031	0.903-1.176	0.654	7474.780	0.130
Apparent	Maximum	90^{th}	0	0.020	1.020	0.906-1.148	0.739	7474.869	0.124
Apparent	Maximum	90^{th}	1	0.050	1.051	0.934-1.182	0.407	7474.292	0.166
Apparent	Maximum	90^{th}	2	0.066	1.068	0.949-1.202	0.274	7473.785	0.214
Apparent	Maximum	90^{th}	3	0.050	1.052	0.934-1.184	0.406	7474.290	0.166
Apparent	Maximum	90^{th}	5	0.047	1.049	0.93-1.182	0.438	7474.379	0.159
Apparent	Maximum	90^{th}	10	0.012	1.012	0.895-1.145	0.846	7474.943	0.120
Normal	Mean	99^{th}	0	0.045	1.046	0.826-1.321	0.708	7474.840	0.126
Normal	Mean	99^{th}	1	0.006	1.006	0.792-1.273	0.961	7474.978	0.118
Normal	Mean	99^{th}	2	0.016	1.016	0.803-1.283	0.892	7474.962	0.119
Normal	Mean	99^{th}	3	-0.032	0.968	0.747-1.251	0.807	7474.920	0.121
Normal	Mean	99^{th}	5	0.009	1.009	0.761-1.333	0.949	7474.976	0.118
Normal	Mean	99^{th}	10	0.111	1.117	0.805-1.542	0.503	7474.535	0.147
Normal	Minimum	99^{th}	0	0.137	1.146	0.871-1.505	0.327	7474.026	0.190
Normal	Minimum	99^{th}	1	0.047	1.048	0.805-1.359	0.724	7474.856	0.125
Normal	Minimum	99^{th}	2	-0.071	0.931	0.704-1.225	0.614	7474.725	0.134
Normal	Minimum	99^{th}	3	0.010	1.010	0.752-1.349	0.947	7474.976	0.118
Normal	Minimum	99^{th}	5	-0.067	0.935	0.682-1.273	0.672	7474.800	0.129
Normal	Minimum	99^{th}	10	0.029	1.030	0.718-1.466	0.871	7474.954	0.119
Normal	Maximum	99^{th}	0	-0.036	0.965	0.709-1.305	0.819	7474.928	0.121
Normal	Maximum	99^{th}	1	-0.086	0.918	0.668-1.251	0.590	7474.688	0.136
Normal	Maximum	99^{th}	2	-0.064	0.938	0.683-1.28	0.689	7474.819	0.128
Normal	Maximum	99^{th}	3	0.018	1.018	0.725-1.422	0.915	7474.969	0.118
Normal	Maximum	99^{th}	5	0.040	1.041	0.744-1.447	0.814	7474.925	0.121

Temperature		Percentile	Lag	Coefficient	OR	95% CI	P-value	AICc	RelAIC
type	***measure***								
Normal	Maximum	99th	10	0.089	1.093	0.779-1.525	0.604	7474.712	0.135
Normal	Mean	95th	0	0.030	1.031	0.881-1.205	0.705	7474.837	0.126
Normal	Mean	95th	1	0.013	1.013	0.867-1.182	0.874	7474.955	0.119
Normal	Mean	95th	2	0.029	1.029	0.882-1.2	0.713	7474.845	0.126
Normal	Mean	95th	3	0.024	1.024	0.879-1.192	0.761	7474.888	0.123
Normal	Mean	95th	5	0.023	1.023	0.872-1.198	0.780	7474.903	0.122
Normal	Mean	95th	10	-0.047	0.954	0.806-1.128	0.585	7474.681	0.137
Normal	Minimum	95th	0	0.094	1.098	0.93-1.295	0.267	7473.755	0.217
Normal	Minimum	95th	1	0.120	1.128	0.953-1.333	0.161	7473.024	0.313
Normal	Minimum	95th	2	0.053	1.055	0.891-1.247	0.534	7474.595	0.143
Normal	Minimum	95th	3	0.022	1.023	0.855-1.22	0.805	7474.920	0.121
Normal	Minimum	95th	5	0.009	1.009	0.84-1.21	0.921	7474.970	0.118
Normal	Minimum	95th	10	-0.022	0.979	0.795-1.202	0.838	7474.938	0.120
Normal	Maximum	95th	0	-0.057	0.945	0.804-1.108	0.487	7474.496	0.150
Normal	Maximum	95th	1	-0.029	0.971	0.824-1.143	0.726	7474.857	0.125
Normal	Maximum	95th	2	-0.029	0.972	0.826-1.141	0.726	7474.857	0.125
Normal	Maximum	95th	3	0.003	1.003	0.852-1.179	0.972	7474.979	0.118
Normal	Maximum	95th	5	0.025	1.025	0.874-1.2	0.758	7474.886	0.123
Normal	Maximum	95th	10	-0.076	0.927	0.782-1.097	0.381	7474.210	0.173
Normal	Mean	90th	0	0.037	1.038	0.915-1.176	0.562	7474.644	0.139
Normal	Mean	90th	1	0.083	1.087	0.959-1.231	0.192	7473.280	0.275
Normal	Mean	90th	2	0.108	1.114	0.984-1.26	0.088	7472.082	0.501
Normal	Mean	90th	3	0.090	1.094	0.965-1.238	0.159	7472.998	0.317
Normal	Mean	90th	5	0.047	1.048	0.925-1.188	0.458	7474.430	0.155
Normal	Mean	90th	10	0.017	1.017	0.892-1.159	0.803	7474.918	0.121
Normal	Minimum	90th	0	0.098	1.103	0.974-1.25	0.122	7472.598	0.387
Normal	Minimum	90th	1	0.114	1.120	0.989-1.268	0.073	7471.769	0.586
Normal	**Minimum**	**90th**	**2**	**0.131**	**1.140**	**1.006-1.29**	**0.039**	**7470.739**	**0.981**
Normal	Minimum	90th	3	0.124	1.132	1.00-1.281	0.051	7471.168	0.792
Normal	**Minimum**	**90th**	**5**	**0.128**	**1.137**	**1.003-1.288**	**0.045**	**7470.976**	**0.872**
Normal	Minimum	90th	10	0.040	1.040	0.914-1.184	0.548	7474.619	0.141
Normal	Maximum	90th	0	-0.005	0.995	0.879-1.126	0.934	7474.973	0.118
Normal	Maximum	90th	1	0.028	1.029	0.91-1.163	0.652	7474.777	0.130
Normal	Maximum	90th	2	0.040	1.041	0.918-1.179	0.531	7474.588	0.143
Normal	Maximum	90th	3	0.025	1.025	0.905-1.161	0.693	7474.825	0.127
Normal	Maximum	90th	5	0.006	1.006	0.886-1.14	0.930	7474.973	0.118
Normal	Maximum	90th	10	-0.013	0.987	0.864-1.127	0.850	7474.944	0.120

Appendix 2: Details of Methods

6.4 *Data Management*

All study data was kept and to a large extent processed in a PostGIS 2.0 database. This consisted of a PostgreSQL 9.0.5 installation with PostGIS functions installed. The version of PostGIS used was 2.0 (SVN), revision 7789. Illustration 19 is a schematic of the database used for the developing the relationships between the different raw and derived data sets. The prefixes "*geov*" and "*geor*" represent geometry tables with vector and raster values, respectively.

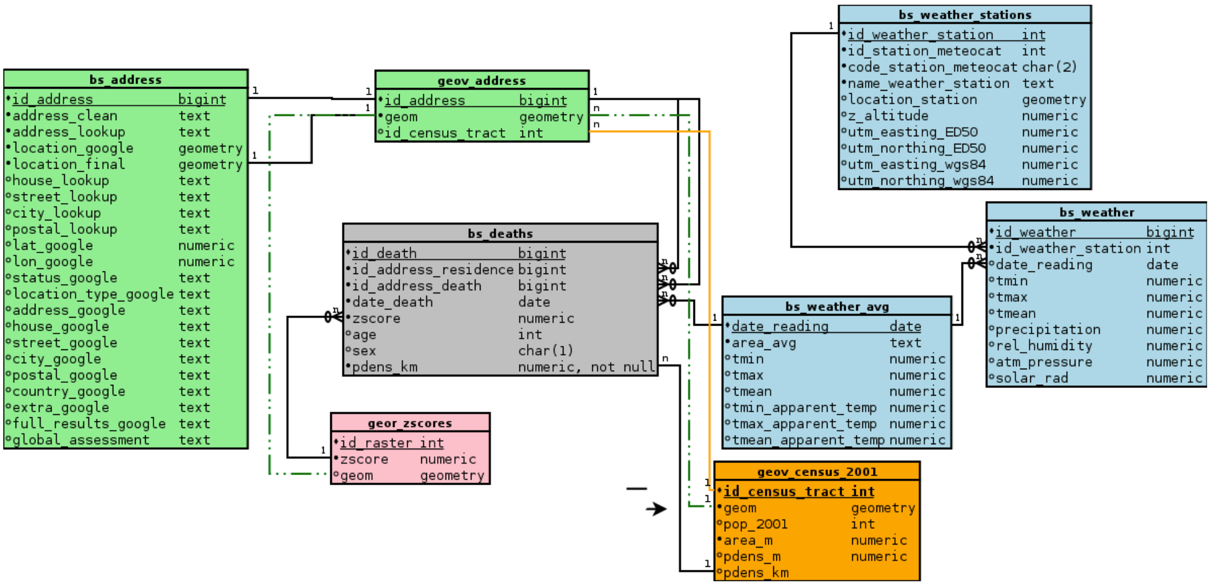


Illustration 19: Schematic of PostGIS 2.0 database used for project analysis.

Standard SQL joins and updating methods were used to link tables and extract corresponding values. In the case of tables containing spatial vector or raster geometries, the PostGIS functions used are described below.

Deaths from the *geov_address* vector table were linked to census tracts in the *geov_census_2001* vector table using the PostGIS *ST_Intersects* function, as follows:

```
UPDATE analysis.geov_address_32631
SET codi_cs_2001 = cs.codi_cs_int
FROM staging.geov_census_2001_32631 AS cs
WHERE ST_Intersects(analysis.geov_address_32631.the_geom, cs.geom);
```

where *geov_address_32631* is a spatial table of vector values for each address, and *geov_census_2001_32631* is a spatial vector table (Illustration 19) describing the census tracts in Barcelona (in schemas named "analysis" and "staging", respectively). The number suffix "*32631*" identifies the tables as containing values in the EPSG:32631 spatial reference system.

Values from the *geor_zscores* raster table at each address were extracted in a JOIN statement linking the vector address position to a point on the raster layer and creating an on-the-fly table with the results, as follows:

```
JOIN (
	SELECT v.id_address
		, v.codi_cs_2001
		, ST_Value(r.rast,1,v.the_geom) AS val
		, ST_X(v.the_geom) as x
		, ST_Y(v.the_geom) as y
	FROM analysis.geov_address_32631 AS v
		JOIN analysis.geor_zscores AS r
		ON ST_Intersects(v.the_geom, r.rast)
) AS zvals
```

where *geov_address_32631* is a vector table containing the locations of each address, indexed by the *id_address* field, and *geor_zscores* is a table containing the geographical information for the raster of the averaged Landsat 7 z-scores (Illustration 19). The *ST_Intersects* function identifies the intersection of the two geometries (the address point on the vector layer and the corresponding cell on the raster layer), the *ST_Value* function returns the value at the address position from the raster and the *ST_X* and *ST_Y* functions return the x and y coordinates of the address. This temporary table, "*zvals*", was then used to update values in the *bs_deaths* table. The final data for analysis was later extracted from the *bs_deaths* table into a comma separated values (CSV) file for analysis using R.

6.5 Regression

All regression analysis was done using the R statistical package (version 2.13.1) after exporting the processed study data from the PostGIS database to a CSV file. The following R packages were used: *sp*, *spdep*, *rgdal*, *RANN*, *dlnm*, *splines* and *AICcmodavg*.

6.5.1 Model

The ***glm***() function was used to create a general linear model with a logit link for the regression analysis before adjusting for spatial autocorrelation, using the following R code structure:

```
fit <- glm(formula=<formula>, family=binomial(link="logit"))
```

- **Formula**: the regression formula of the type "hrr0 ~ hat_tmin90_l2 +age +sex +pdens_km".
- **Family**: as the dependent variable was binary, a binomial link function for the glm was used (family=*binomial*(link=*logit*) which is the default for the binomial family).

6.5.2 Moran eigenvector GLM filtering

The ME() function, part of the *spdep* package, was used to generate the filtered set of eigenvectors. The code for the filtering process in R was the following:

```
library(spdep)
lag <- ME(formula=<formula>, family=binomial(link="logit"), alpha=0.05,
```

```
listw=lw_dd, data=d23)
```

where d23 was the subset of all mortality and weather data between the years 2000 and 2003 (inclusive) and between April 1 and September 30 (inclusive). The formula was, in the reference case,

```
hrr0 ~ hat_tmin90_l2 +age +sex +pdens_km
```

The *ME* function uses *listw* objects, which are weight lists, and all networks generated were converted to weight lists using the *nb2listw*() function with "W" row-standardized weighting (*Tiefelsdorf et al 1999*), as described in the *Network Weighting* section.

The lag model was then reintroduced to the regression formula to produce the final results partially adjusted for spatial autocorrelation, with this code:

```
fit_spatial <- glm(formula=hrr0 ~ hat_tmi90_l2 +age +sex +pdens_km +fitted(lag),
family=binomial(link="logit"))
```

6.6 Modeling Time Lag

It has been proposed for many years that heat-associated mortality shows non-linear effects over time, affecting mortality days after heat impacts, both through delayed mortality (*Gosling et al. 2008; Bayentin et al. 2010; Kinney et al. 2008; Metzger et al. 2010*) and the "harvesting" effect (*Gosling et al. 2008; Metzger et al. 2010*) of bringing forward imminent deaths, thus reducing subsequent mortality for an undetermined period.

The R module DLNM (*Distributed Lag Non-linear Models*) was used to profile the effects of heat across a range of temperatures and lag times, using the study data reorganized into a time series.

6.6.1 How it works

The DLNM module performs a fairly complex analysis, but its basic functionality is to calculate the effects of a factor using non-linear models across both a range of factor values and a range of time lags. DLNM was developed for risk factors that produce time-lagged effects, such as temperature and air pollution, and the documentation provides several examples about how to apply the package to these types of datasets.

Starting with a vector of exposures x_j, with j = 1...n, the basic formula (excluding other predictors) for a series of outcomes *Y*t with time t = 1...n is:

$$g(\mu_t) = \alpha + J \sum_{j=1} s_j(x_{tj};\beta_j)$$

where $\mu \equiv E(Y)$ and g is a link function. The function s_j smooths the variable x_j and the parameter vector β. A set of *basis functions* can be defined for the relationship between x and $g(\mu)$, of which s is one, and this set of functions is called a *basis*. The matrix of variables x is transformed by the set of basis functions to create a new matrix of *basis variables*. These transformations can be more or less complex, depending on the functions chosen. At this point, the matrix of basis variables just describes the effect of the variable x.

Another matrix *q* of basis variables can be defined, based on the delayed effects of the variable *xt-l* where *t* is the time of exposure and *l* is the lag time. By applying a set of basis functions describing lag effects to the original array *x*, a new matrix of time-lagged basis variables is created for lags.

Combining the two sets of basis functions creates a set of cross basis functions, which can be used to describe the behavior of the variable in three dimensions: the variable scale, the effect scale (as a risk ratio) and the lag scale.

6.6.2 Methods

The dataset was reorganized as a time series of days from 2000-04-01 to 2003-09-30 - the warm season - during the study period, with daily counts of deaths and and daily temperature measures. Apparent temperatures were only available for 2000-2003.

First, a crossbasis was established for both the variable and lag arrays (the *basis* function is internal, and called by the *crossbasis* function):

```
btmean = crossbasis(dts$tmean, vartype="bs", vardf=5, vardegree=2, lagdf=5,
maxlag=5)
```

in this case for mean temperature, using a B-spline, five degrees of freedom in the variable and lag axes, and with a maximum lag of five days. The relationship between *x* and both the direct effect and the lag effect are non-linear, and the lag was defined as a natural spline (the default). Then, a general linear model with a quasi-poisson link is fitted to the data using only the temperature variable as an explanatory factor, and a natural spline of the date variable to adjust for regular time effects, as follows:

```
mtmean5 = glm(formula=deaths~btmean5 + ns(date_death, 7*14),
family=quasipoisson(), data=dts)
```

The quasi-poisson is a type of poisson function which allows for greater variability in the data than a normal poisson function. Then the model is combined with the crossbasis to produce a prediction object:

```
ptmean5 <- crosspred(btmean5, mtmean5, by=1)
```

which could then be plotted.

6.7 Network Selection and Weighting

6.7.1 Network Selection

The ME function takes a series of arguments and returns a lag object that can then be entered into the formula of a GLM regression as a fitted model. The following parameters to the ME function call are the ones that are relevant to this analysis:

- **Formula**: this is the same formula as the regression model for the lag model, in the form <dependent variable> ~ <independent variable1> + <independent variable 2>, etc.
- **Alpha**: the threshold for the eigenvector selection stopping rule. The *ME*() process will include all eigenvectors up to and including the first eigenvector that exceeds this threshold.
- **Listw**: a weighted list, created from a neighborhood network (*nb*) object, using the *nb2listw*()

function.

The two most important parameters are the *alpha* setting and the *listw* object. Higher alpha settings generate lag models that fit the data better, at the expense of greater computational resources. The individual vectors identified tend to have lower p-values than the threshold when included in the final regression model (typically, a threshold of 0.05 will produce a set of vectors all below a p-value of 0.001), and only the final vector will exceed this value. When using a higher threshold, there is also the danger of overfitting the model, in which case a backwards stepwise regression can be used to remove some of the vectors. This was not done in this study, as the alpha threshold chosen was high enough that all (or nearly all) eigenvectors produced were significant at the 0.001 level (the highest p-values produces were in the range of 0.0002, and the majority were significant to more than a dozen decimal places).

A more complex process is involved in the generation of a *listw* object. First, an *nb* object must be generated using one of several different methods (see *Network Types*, below). This object is then converted to a *listw* object, with the important parameters being whether the network should be made symmetrical (required in the case of Gabriel and Relative Neighbor networks, which normally do not allow more than one node in the same position) and whether zero-weight nodes should be allowed - this is necessary when the distance to the closest node is greater than the maximum distance set for a DNN, which creates a discontinuous network. The minimum distance needed to create a continuous network with the study dataset was 979.62 meters. The tests done allowed zero-weight nodes and set distance thresholds at 1000, 750 and 500 meters.

6.7.1.1 Network Types

Dray, Legendre and Peres-Neto (*2005*) tested different spatial weighting matrices for their fit to a dataset using the *ME* eigenvector filtering techniques. They tested Delaunay triangulation (TRI), Gabriel networks (GAB), Relative Neighbor networks (RNN), minimum spanning trees (MST), and distance-based networks (DNN), and compared their results using the corrected Akaike Information Criteria (AICc) measure. Their best results were produced using DNN with specific types of weighting functions and RNN with binary weighting.

In this study, the networks that could be produced for analysis were GAB, RNN and DNN, although several of the networks generated required that zero-weight nodes be allowed. DNN networks had, in addition to the set of point coordinates, two distance parameters corresponding to the range of distance bounds within which points were considered neighbors. The lower bound was set to zero, so the upper bound functioned as a neighbor membership threshold. Because a distance below 980 meters created zero-weight nodes, they were allowed in the generation of the *listw* object for all distances.

The *ME* function uses *listw* objects, which are weight lists, and all networks generated were converted to weight lists with “W”, row-standardized weighting (*Tiefelsdorf et al 1999*), as described in the following section, *Network Weighting*.

6.7.1.2 *Network Weighting*

The weighting used when processing a network for inclusion in the eigenvector filtering process is very important (*Dray, Legendre and Peres-Neto, 2006; Teifelsdorf, Griffith and Boots, 1999*). The default weighting scheme used was the "W" row-standardized weighting scheme, although "B", "C", "U" and "S" weighting schemes were also tested. The coding schemes are derived from a *n x m* matrix combining the *n* nodes of the network to be studied, called G.

The "B" scheme is the basic binary association between two nodes (0 or 1, neighbors or not). If a node is within the radius of the selected distance from the test node, it is considered a neighbor and given a weight of one, otherwise it has weight zero.

The "C" scheme is globally standardized over the connectivity of all the elements in G. The "W" scheme is row standardized, where each row in the matrix is the set of connections to a single node. "U" weights are equal to the "C" divided by the number of neighbors and the "S" scheme was devised by Tiefelsdorf et al. (*1999*) to stabilize variance.

Tiefelsdorf et al. (*1999*) summarize the characteristics of the two principal schemes in this way: "*Results show that the C-coding scheme emphasizes spatial objects with relatively large numbers of connections, such as those in the interior of a study region. In contrast, the W-coding scheme assigns higher leverage to spatial objects with few connections, such as those on the periphery of a study region.*" (*page 165*). The S-coding scheme was devised to strike a balance between the properties of the W and U schemes.

All weighting options were tested on the same data set and compared based on AICc value. The B, C and U schemes produced exactly the same results, but otherwise the B weighting scheme showed the best fit to the data, with better AICc scores than the S and W schemes. It also has the advantage of being computationally simpler than some of the other schemes, but is more difficult to interpret. *Appendix 3: Moran's I Values* shows the effects of different networks and weighting schemes on Moran's I values. The different weighting schemes compared (B, W and S) show similar results.

For the final model, W-weighting was chosen due to its familiarity and ease of interpretation. Unless stated otherwise, all regression models in this study used W-style row-standardized weighting.

6.7.1.3 *Network Selected*

The network type chosen for the analysis was a DNN (with W-weighting), for several reasons. First, the fixed-parameter GAB and RNN networks are computationally intensive, and the only way to simplify the required calculations was to sample the data rather than use the entire dataset. This was done for a 36% sample (2000 cases). The results are shown in the *Other Network Types* section, below. Adjusting the distance parameter of a DNN produces models that can be filtered using the *ME* function in a reasonable amount of time. Second, the interpretation of a DNN is fairly intuitive, and distance thresholds can be compared to the size of HRR areas. Third, using a DNN makes it possible to explore the effect of distance scaling on the performance of a regression model.

6.7.2 The effect of distance thresholds on the model

Table 19 shows the effect of reducing the maximum distance in a DNN model on spatially adjusted results for the full data set (Table 20 shows a slightly broader range of distances using a sample of cases). Distance thresholds are shown in the table, along with the average time taken for an analysis and the number of eigenvectors identified.

It was not possible to do extensive testing, but the tendency in these limited results was that the lower the distance threshold, the larger the number of significant eigenvectors in the matrix and the more demanding the filtering process. Based on the testing of small samples, there was a threshold between 500-250 meters where the calculation of results became effectively impossible, even for a limited sample, and no results could be obtained for the 250 meter threshold.

One exception to the tendency toward increasing time and eigenvectors produced was the 750 meter threshold, which required slightly less time and produced slightly less eigenvectors than the 1000 meter threshold.

Table 19: Results of distance threshold testing with best model (tmin90_l2)

Threshold	Coefficient	OR	95% CI	P-value	AICc	RelAIC	Hours	eV
meters								***eigenvectors***
1000	0.138	**1.148**	**1.003-1.313**	**0.045**	6618.893	2.67E-68	3.80	17
750	0.141	**1.151**	**1.005-1.319**	**0.043**	6516.797	3.95E-46	3.32	16
500	0.141	**1.152**	**1.002-1.324**	**0.047**	6307.707	1.00	6.69	26

All models tested at alpha=0.05 level, using W-weighting. OR = Odds Ratio, eV = eigenvectors.

The progression toward shorter distance thresholds produced models with better (lower) AICc scores, but the trend toward increasing significance in the p-values was not maintained as the fit improved, and in fact the reduction of the distance threshold resulted in a slight decrease in the significance of the results. This may indicate that models at the 750 meter level were an anomaly, or that this distance had some important characteristics that influenced the data analysis. In any case, the results remained significant across this range of distance thresholds (1000-500 meters), indicating that this association was robust.

6.7.3 Other Network Types

To compare the DNN networks to GAB and RNN networks, a 36% sample (2000 points) was selected randomly from the dataset and analyzed in the tmin90_l2 model, with spatial autocorrelation adjustment. The results are shown in Table 20.

Table 20: Results of a 36% (2000 point) sample, by network type and distance

Network	Distance	Coefficient	OR	95% CI	P-value	AICc	Hours	eV
	meters							***eigenvectors***
GAB	--	0.367	1.443	1.084-1.924	0.012	1882.318	5.73	91
RNN	--	0.123	1.131	0.848-1.507	0.403	1835.285	6.86	94
DNN	1000	0.101	1.106	0.888-1.378	0.367	2461.055	0.23	8
DNN	750	0.132	1.141	0.912-1.427	0.248	2405.831	0.34	13
DNN	500	0.060	1.062	0.844-1.334	0.609	2359.156	0.89	22
DNN	50	6.95E+13	Inf	Error	0	65990.130	0.55	10

All models tested at alpha=0.05 level, using W-weighting and a single random sample of 3000 cases from the study dataset. OR = Odds Ratio, eV = eigenvectors.

The same pattern as seen with the full dataset emerged with the distance-based models – shorter distances required more computing time to adjust for spatial autocorrelation and produced better fitting models with more eigenvectors. P-values were more unstable, and dipped at the 750 meter level, then rose again. The shorter distances provided a better fit for the data.

The very best fit, based on AICc values, came from the RNN. The GAB network also provided a better fit than the DNN samples tested. These results, however, were at the cost of greater computing resources. Two important aspects of these types of networks to consider are the computational requirements for using them, as well the interpretation of the results.

Another noteworthy result was for the 50m threshold. This model calculated quickly, had a terrible AICc value, and nonsense results. While it was not a useful result from the study perspective, it did show that smaller distance thresholds did not automatically provide a better fit or require more time to calculate in a mathematical progression – a point was reached where the network essentially collapsed as there were not enough points with neighbors.

Appendix 3: Moran's I Values

Moran's I values were calculated for all study variables used in the regression models.

Table 21: Moran's I values for study variables, by Network and Weighting Type

Network	Distance threshold	Wt	Variable	MI	P-Value
Gabriel	na	B	Tmin	0.019	0.01135
Relative Neighbor	na	B	Tmin	0.016	0.04565
Distance	1000	B	Tmin	-0.001	0.79863
Distance	750	B	Tmin	-0.001	0.70428
Distance	500	B	Tmin	-0.001	0.70684
Distance	250	B	Tmin	0.004	0.10977
Distance	100	B	Tmin	0.017	0.02227
Gabriel	na	W	Tmin	0.019	0.02122
Relative Neighbor	na	W	Tmin	0.016	0.07233
Distance	1000	W	Tmin	0.000	0.58862
Distance	750	W	Tmin	-0.002	0.89187
Distance	500	W	Tmin	-0.004	0.94944
Distance	250	W	Tmin	0.002	0.32095
Distance	100	W	Tmin	0.008	0.20969
Gabriel	na	S	Tmin	0.019	0.01420
Relative Neighbor	na	S	Tmin	0.016	0.05217
Distance	1000	S	Tmin	-0.001	0.79696
Distance	750	S	Tmin	-0.001	0.82216
Distance	500	S	Tmin	-0.002	0.85473
Distance	250	S	Tmin	0.004	0.14605
Distance	100	S	Tmin	0.013	0.06907
Gabriel	na	B	Tmax	0.012	0.07876
Relative Neighbor	na	B	Tmax	0.013	0.07831
Distance	1000	B	Tmax	-0.002	0.94371
Distance	750	B	Tmax	-0.001	0.79048
Distance	500	B	Tmax	-0.003	0.93839
Distance	250	B	Tmax	-0.001	0.60661
Distance	100	B	Tmax	-0.001	0.53237
Gabriel	na	W	Tmax	0.003	0.38555
Relative Neighbor	na	W	Tmax	0.004	0.35300
Distance	1000	W	Tmax	-0.002	0.95315
Distance	750	W	Tmax	-0.002	0.90669

Network	Distance threshold	Wt	Variable	MI	P-Value
Distance	500	W	Tmax	-0.005	0.98392
Distance	250	W	Tmax	-0.005	0.84302
Distance	100	W	Tmax	-0.008	0.76452
Gabriel	na	S	Tmax	0.007	0.21652
Relative Neighbor	na	S	Tmax	0.008	0.20858
Distance	1000	S	Tmax	-0.002	0.95159
Distance	750	S	Tmax	-0.002	0.85315
Distance	500	S	Tmax	-0.004	0.96382
Distance	250	S	Tmax	-0.002	0.67120
Distance	100	S	Tmax	-0.003	0.63952
Gabriel	na	B	Tmean	0.015	0.03489
Relative Neighbor	na	B	Tmean	0.017	0.03364
Distance	1000	B	Tmean	-0.001	0.72275
Distance	750	B	Tmean	0.000	0.50323
Distance	500	B	Tmean	-0.002	0.75725
Distance	250	B	Tmean	0.004	0.14305
Distance	100	B	Tmean	0.004	0.33398
Gabriel	na	W	Tmean	0.011	0.11044
Relative Neighbor	na	W	Tmean	0.014	0.09835
Distance	1000	W	Tmean	0.000	0.44644
Distance	750	W	Tmean	-0.001	0.64685
Distance	500	W	Tmean	-0.003	0.91174
Distance	250	W	Tmean	0.002	0.34135
Distance	100	W	Tmean	0.000	0.48899
Gabriel	na	S	Tmean	0.014	0.05882
Relative Neighbor	na	S	Tmean	0.016	0.05642
Distance	1000	S	Tmean	-0.001	0.69942
Distance	750	S	Tmean	0.000	0.58454
Distance	500	S	Tmean	-0.002	0.83698
Distance	250	S	Tmean	0.003	0.16718
Distance	100	S	Tmean	0.002	0.39466
Gabriel	na	B	Age	0.072	0.00000
Relative Neighbor	na	B	Age	0.086	0.00000
Distance	1000	B	Age	0.020	0.00000
Distance	750	B	Age	0.022	0.00000
Distance	500	B	Age	0.027	0.00000
Distance	250	B	Age	0.025	0.00000
Distance	100	B	Age	0.022	0.00484

Network	Distance threshold	Wt	Variable	MI	P-Value
Gabriel	na	W	Age	0.040	0.00001
Relative Neighbor	na	W	Age	0.035	0.00067
Distance	1000	W	Age	0.022	0.00000
Distance	750	W	Age	0.023	0.00000
Distance	500	W	Age	0.031	0.00000
Distance	250	W	Age	0.030	0.00000
Distance	100	W	Age	0.021	0.02064
Gabriel	na	S	Age	0.054	0.00000
Relative Neighbor	na	S	Age	0.056	0.00000
Distance	1000	S	Age	0.020	0.00000
Distance	750	S	Age	0.021	0.00000
Distance	500	S	Age	0.028	0.00000
Distance	250	S	Age	0.027	0.00000
Distance	100	S	Age	0.022	0.00682
Gabriel	na	B	Sex	0.055	0.00000
Relative Neighbor	na	B	Sex	0.066	0.00000
Distance	1000	B	Sex	0.009	0.00000
Distance	750	B	Sex	0.009	0.00000
Distance	500	B	Sex	0.010	0.00000
Distance	250	B	Sex	0.009	0.00463
Distance	100	B	Sex	0.015	0.03711
Gabriel	na	W	Sex	0.029	0.00095
Relative Neighbor	na	W	Sex	0.024	0.01480
Distance	1000	W	Sex	0.010	0.00000
Distance	750	W	Sex	0.010	0.00000
Distance	500	W	Sex	0.012	0.00000
Distance	250	W	Sex	0.016	0.00023
Distance	100	W	Sex	0.009	0.19461
Gabriel	na	S	Sex	0.039	0.00000
Relative Neighbor	na	S	Sex	0.040	0.00003
Distance	1000	S	Sex	0.009	0.00000
Distance	750	S	Sex	0.009	0.00000
Distance	500	S	Sex	0.010	0.00000
Distance	250	S	Sex	0.011	0.00228
Distance	100	S	Sex	0.012	0.09174
Gabriel	na	B	PDens-Km	0.669	0.00000
Relative Neighbor	na	B	PDens-Km	0.743	0.00000
Distance	1000	B	PDens-Km	0.171	0.00000

Network	Distance threshold	Wt	Variable	MI	P-Value
Distance	750	B	PDens-Km	0.211	0.00000
Distance	500	B	PDens-Km	0.280	0.00000
Distance	250	B	PDens-Km	0.415	0.00000
Distance	100	B	PDens-Km	0.613	0.00000
Gabriel	na	W	PDens-Km	0.665	0.00000
Relative Neighbor	na	W	PDens-Km	0.710	0.00000
Distance	1000	W	PDens-Km	0.208	0.000
Distance	750	W	PDens-Km	0.250	0.000
Distance	500	W	PDens-Km	0.330	0.000
Distance	250	W	PDens-Km	0.466	0.000
Distance	100	W	PDens-Km	0.590	0.000
Gabriel	na	S	PDens-Km	0.665	0.000
Relative Neighbor	na	S	PDens-Km	0.725	0.000
Distance	1000	S	PDens-Km	0.180	0.000
Distance	750	S	PDens-Km	0.221	0.000
Distance	500	S	PDens-Km	0.293	0.000
Distance	250	S	PDens-Km	0.427	0.000
Distance	100	S	PDens-Km	0.601	0.000
Gabriel	na	B	HRR0	0.552	0.000
Relative Neighbor	na	B	HRR0	0.642	0.000
Distance	1000	B	HRR0	0.113	0.000
Distance	750	B	HRR0	0.134	0.000
Distance	500	B	HRR0	0.166	0.000
Distance	250	B	HRR0	0.268	0.000
Distance	100	B	HRR0	0.474	0.000
Gabriel	na	W	HRR0	0.520	0.000
Relative Neighbor	na	W	HRR0	0.581	0.000
Distance	1000	W	HRR0	0.117	0.000
Distance	750	W	HRR0	0.143	0.000
Distance	500	W	HRR0	0.183	0.000
Distance	250	W	HRR0	0.284	0.000
Distance	100	W	HRR0	0.474	0.000
Gabriel	na	S	HRR0	0.528	0.000
Relative Neighbor	na	S	HRR0	0.603	0.000
Distance	1000	S	HRR0	0.115	0.000
Distance	750	S	HRR0	0.138	0.000
Distance	500	S	HRR0	0.173	0.000
Distance	250	S	HRR0	0.275	0.000
Distance	100	S	HRR0	0.474	0.000

Appendix 4: Variable Correlations

Table 22: Correlations between study variables

Variable		Regression	Coefficient	P-value
Dependent	***Independent***	***type***		
Age		linear		
	sex_m0_f1		6.840	0.00000
	pdens_km		0.000	0.00017
	hat_tmean90		1.155	0.15407
	hat_tmin90		0.412	0.51413
	hat_tmax90		-0.375	0.52884
Sex		logistic		
	age		0.054	0.00000
	pdens_km		0.000	0.00032
	hat_tmean90		-0.109	0.45900
	hat_tmin90		0.054	0.04188
	hat_tmax90		0.022	0.36573
Population density		linear		
	age		-99.269	0.00017
	sex_m0_f1		-2286.478	0.00032
	hat_tmean90		-399.066	0.80267
	hat_tmin90		314.740	0.80045
	hat_tmax90		-364.094	0.75654
Mean temperature		linear		
	age		0.000	0.15407
	sex_m0_f1		-0.004	0.45900
	pdens_km		0.000	0.80267
	hat_tmin90		0.520	0.00000
	hat_tmax90		0.459	0.00000
Minimum temperature		linear		
	age		0.000	0.51413
	sex_m0_f1		0.014	0.04188
	pdens_km		0.000	0.80045
	hat_tmean90		0.855	0.00000
	hat_tmax90		-0.031	0.01303
Maximum temperature		linear		
	age		0.000	0.52884
	sex_m0_f1		0.007	0.36573
	pdens_km		0.000	0.75654
	hat_tmean90		0.849	0.00000
	hat_tmin90		-0.35	0.01303

Appendix 5: Spatially-Adjusted Regression Results

Table 23: Regression results after adjustment for spatial autocorrelation

Temperature		Perc.	Lag	Coef.	OR	95% CI	P-value	AICc	RelAIC	EV
type	***measure***									
Apparent	Mean	90^{th}	0	0.038	1.039	0.905-1.191	0.587	6520.603	0.149	16
Normal	Mean	90^{th}	0	0.038	1.039	0.905-1.191	0.587	6520.603	0.149	16
Apparent	Minimum	90^{th}	0	0.078	1.081	0.939-1.244	0.277	6519.718	0.232	16
Normal	Minimum	90^{th}	0	0.087	1.091	0.952-1.251	0.210	6519.329	0.282	16
Apparent	Maximum	90^{th}	0	0.023	1.023	0.899-1.164	0.728	6520.777	0.137	16
Normal	Maximum	90^{th}	0	0.000	1.000	0.873-1.144	0.995	6520.897	0.129	16
Apparent	Mean	90^{th}	1	0.100	1.105	0.963-1.267	0.154	6518.869	0.355	16
Normal	Mean	90^{th}	1	0.100	1.105	0.963-1.267	0.154	6518.869	0.355	16
Apparent	**Minimum**	$\mathbf{90^{th}}$	**1**	**0.139**	**1.149**	**1.001-1.32**	**0.048**	**6517.011**	**0.899**	**16**
Normal	Minimum	90^{th}	1	0.123	1.131	0.987-1.296	0.076	6517.750	0.621	16
Apparent	Maximum	90^{th}	1	0.062	1.064	0.935-1.21	0.349	6520.021	0.199	16
Normal	Maximum	90^{th}	1	0.051	1.052	0.92-1.204	0.457	6520.344	0.170	16
Apparent	Mean	90^{th}	2	0.123	1.131	0.987-1.295	0.076	6517.765	0.616	16
Normal	Mean	90^{th}	2	0.123	1.131	0.987-1.295	0.076	6517.765	0.616	16
Apparent	**Minimum**	$\mathbf{90^{th}}$	**2**	**0.140**	**1.150**	**1.003-1.319**	**0.046**	**6516.910**	**0.945**	**16**
Normal	**Minimum**	$\mathbf{90^{th}}$	**2**	**0.141**	**1.151**	**1.005-1.319**	**0.043**	**6516.797**	**1.000**	**16**
Apparent	Maximum	90^{th}	2	0.079	1.082	0.951-1.232	0.231	6519.463	0.264	16
Normal	Maximum	90^{th}	2	0.063	1.066	0.929-1.222	0.364	6520.075	0.194	16
Apparent	Mean	90^{th}	3	0.093	1.097	0.957-1.257	0.184	6519.131	0.311	16
Normal	Mean	90^{th}	3	0.093	1.097	0.957-1.257	0.184	6519.131	0.311	16
Apparent	Minimum	90^{th}	3	0.134	1.143	0.996-1.312	0.057	6517.295	0.780	16
Normal	Minimum	90^{th}	3	0.128	1.137	0.992-1.302	0.065	6517.488	0.708	16
Apparent	Maximum	90^{th}	3	0.056	1.057	0.928-1.204	0.404	6520.201	0.182	16
Normal	Maximum	90^{th}	3	0.045	1.046	0.913-1.199	0.516	6520.476	0.159	16
Apparent	Mean	90^{th}	5	0.122	1.130	0.917-1.207	0.085	6517.927	0.568	16
Normal	Mean	90^{th}	5	0.122	1.130	0.917-1.207	0.085	6517.927	0.568	16
Apparent	Minimum	90^{th}	5	0.124	1.132	0.983-1.299	0.077	6517.783	0.611	16
Normal	Minimum	90^{th}	5	0.050	1.051	0.986-1.299	0.459	6520.350	0.169	16
Apparent	Maximum	90^{th}	5	-0.002	0.998	0.921-1.198	0.978	6520.897	0.129	16
Normal	Maximum	90^{th}	5	0.051	1.052	0.869-1.146	0.469	6520.373	0.167	16