Environmental Research 135 (2014) 81–87

Contents lists available at ScienceDirect

Environmental Research

journal homepage: www.elsevier.com/locate/envres

The 2011 heat wave in Greater Houston: Effects of land use on temperature

Weihe Zhou [a], Shuang Ji [a], Tsun-Hsuan Chen [b], Yi Hou [c], Kai Zhang [b,*]

[a] Division of Biostatistics, University of Texas School of Public Health, Houston, TX, USA
[b] Division of Epidemiology, Human Genetics and Environmental Sciences, University of Texas School of Public Health, Houston, TX, USA
[c] CDM Smith, 8140 Walnut Hill Ln, Dallas, TX, USA

ARTICLE INFO

Article history:
Received 31 May 2014
Received in revised form
30 July 2014
Accepted 11 August 2014

Keywords:
Heat wave
Land use regression
Quantile regression
Temperature

ABSTRACT

Effects of land use on temperatures during severe heat waves have been rarely studied. This paper examines land use-temperature associations during the 2011 heat wave in Greater Houston. We obtained high resolution of satellite-derived land use data from the US National Land Cover Database, and temperature observations at 138 weather stations from Weather Underground, Inc (WU) during the August of 2011, which was the hottest month in Houston since 1889. Land use regression and quantile regression methods were applied to the monthly averages of daily maximum/mean/minimum temperatures and 114 land use-related predictors. Although selected variables vary with temperature metric, distance to the coastline consistently appears among all models. Other variables are generally related to high developed intensity, open water or wetlands. In addition, our quantile regression analysis shows that distance to the coastline and high developed intensity areas have larger impacts on daily average temperatures at higher quantiles, and open water area has greater impacts on daily minimum temperatures at lower quantiles. By utilizing both land use regression and quantile regression on a recent heat wave in one of the largest US metropolitan areas, this paper provides a new perspective on the impacts of land use on temperatures. Our models can provide estimates of heat exposures for epidemiological studies, and our findings can be combined with demographic variables, air conditioning and relevant diseases information to identify 'hot spots' of population vulnerability for public health interventions to reduce heat-related health effects during heat waves.

1. Introduction

Heat waves have been reported to be associated with elevated risk of mortality and morbidity (Kovats and Hajat, 2008). Some people are more vulnerable to heat than others, including those with existing chronic disease, athletes, the elderly, children, the mentally ill, poor, among others (O'Neill et al., 2003, 2005). Because temperature is not homogeneous in space, higher exposures at certain areas related to types of land use such as vegetation, urbanicity and built environment (Harlan et al., 2006) may magnify the health risks of heat to vulnerable populations. Improved understanding of the associations between land use and temperature on hot days, particularly during heat waves, is therefore critical in identifying 'hot spots' and thus in targeting public health interventions to prevent heat-related illness and death during heat waves.

Effects of land use on air temperatures measurements during heat waves have been rarely studied. Although some studies have examined the impacts of land use on land surface temperature estimates derived from satellite images (Buscail et al., 2012; Kestens et al., 2011), few studies have explored the associations between air temperature observations and selected land use types and/or physical geographic information, e.g., imperviousness, distance to water bodies and distance to a city center (Oswald et al., 2012; Zhang et al., 2011); imperviousness, elevation and vegetation (Kloog et al., 2012); longitude, latitude, elevation, distances to coastline, and land use types of agriculture, forest, water bodies and build environment (Yunus et al., 2012). Additionally, limited studies have investigated how land use affects ground temperature measurements during severe heat waves when land use might have larger impacts on temperature than on the average temperature days, although some studies have examined the impacts of land use on land surface temperature during heat waves (Buscail et al., 2012). Moreover, previous studies have primarily used linear regression (Oswald et al., 2012), kriging (Zhang et al., 2011) and mixed models (Kloog et al., 2012), and have not addressed a research question: whether certain land use types have larger impacts on temperatures at 'hot spots' compared to non-hot spots.

* Corresponding author. Fax: +1 713 500 9264.
E-mail address: kai.zhang@uth.tmc.edu (K. Zhang).

http://dx.doi.org/10.1016/j.envres.2014.08.025
0013-9351/

This paper aims to examine effects of a variety of land use types on temperatures during the 2011 heat wave in Greater Houston using a land use regression approach, and then to conduct a quantile regression analysis to assess whether the impacts of selected land use on temperatures varied with a spectrum of quantiles. The 2011 heat wave in Houston led to the hottest August since records began in 1889 (National Weather Service, 2011) and the most consecutive days with issued heat wave alerts (August 2 to 30) by National Weather Service (Wood, L, personal communication, January 30, 2013). This heat wave lasted much longer than typical severe heat waves usually lasting for one week. It provides a unique research opportunity to explore land use-temperature associations. This analysis first takes advantage of the state-of-the-art technique developed in air pollution exposure assessment: land use regression, a linear regression built on a variety of variables related to land use, traffic and other variables derived from Geographic Information System (GIS) (Health Effects Institute, 2010; Zhang et al., 2014a). This analysis also applies a quantile regression method, which has been rarely used in heat-related studies.

2. Methods

2.1. Study area

The study area is Houston-Sugar Land-Baytown Metropolitan Statistical Area (MSA) in Texas that was delineated in December, 2009, a 10-county area including Austin, Brazoria, Chambers, Fort Bend, Galveston, Harris, Liberty, Montgomery, San Jacinto and Waller counties in Texas (US Census Bureau, 2013a). It had a total area of 25, 890 km^2 with a population of 5.95 million in 2010, and it was ranked the 6th largest US metropolitan statistical area (US Census Bureau, 2013b). According to the 2010 Census, the total Houston Metro population included a racial-ethnicity makeup of 62.7% white (39.7% non-Hispanic), 18% black or African American, 7.2% Asian, among others (US Census Bureau, 2013b).

2.2. Data source

Temperature observations were obtained from Weather Underground Inc. (WU). Specially, we downloaded daily average, minimum and maximum temperature measurements at all available weather stations from August 1 to August 31 in 2011 (Weather Underground, 2013). We then excluded observations with outliers or invalid values or a lack of internal consistency following the procedure suggested by Durre et al. (2010). There were a total of 138 weather stations in Greater Houston in the study period (Fig. 1).

Land cover and imperviousness data at the 30 m × 30 m resolution were obtained from the National Land Cover Database (U.S. Geological Survey, 2013a). These data were used to calculate the percentage of land cover types and imperviousness around Weather Underground stations within buffer distances varying from 50 m to 750 m using ArcGIS 10.0 (ESRI, Redlands, CA). The nearest distance to the coastline and the nearest distance to water bodies were derived by the shortest straight line distance between weather stations and the coastline of the Gulf of Mexico and the Galveston Bay or between weather stations and lake/pond or reservoirs using ArcGIS 10.0. Coastline and water body data were obtained from the National Hydrography Dataset (U.S. Geological Survey, 2013b). Table 1 illustrates all independent variables used in this study.

2.3. Statistical analysis

We utilized land use regression and quantile regression methods to estimate effects of land use and geographic variables (Table 1) on temperatures during the 2011 heat wave in Greater Houston. We examined three temperature measures as response variables: the monthly averages of daily minimum, mean and maximum at each weather station in August, 2011.

Land use regression models have been receiving considerable attention recently due to its ability to predict long-term exposures to air pollutants, especially in the studies examining the health effects caused by traffic-related air pollution (Health Effects Institute, 2010; Zhang, 2014a). They are essentially built on multiple linear regressions using many variables generally derived using GIS, e.g., land use, traffic intensity, nearby emission sources, and distance to major roads. This study uses GIS-derived variables listed in Table 1.

Quantile regression is a statistical technique for estimating the associations between quantiles of a response variable and independent variables (Koenker, 2005). Quantile regression has two major advantages compared to linear regression. First, quantile regression can provide a more complete picture of the estimated associations by charactering the effects of the covariates on a spectrum of quantiles of the response variable, while linear regression only quantifies the associations between the means of a response variable and independent variables (Koenker, 2005). This feature is particularly important when the covariates' effects vary with quantiles. Second, quantile regression is more robust than linear regression (Koenker, 2005). Linear regression assumes that variables are independent and identically distributed and residuals follow a normal distribution. However, quantile regression can estimate the associations without distributional assumptions.

We conducted the data analysis by following a procedure of screening, buffer size selection, variable selection, land use regression and quantile regression. First, given 114 geographic variables (Table 1), we excluded covariates with less variation when their 85th percentile values were the same with their 15th percentile values. Second, for the variables with varying buffer radii, we determined the 'best' buffer size where a variable had the highest Pearson correlation coefficient with a temperature variable of interest. Third, because some variables were dominated by zero values, we converted those with more than 50% zero values into binary variables. Forth, we built land use regression models by including all previous identified variables and then conducted the stepwise selection procedure to reduce the number of variables by setting p value < 0.15. Then we reran land use regression models based on those selected variables to keep variables with p value < 0.05. For those variables with marginal significance, we also used Bayesian information criterion (BIC) to compare the full model to the reduced model without the covariate of interest, and determined the model with smaller BIC values. Fifth, we ran quantile regression models based on the variables picked from previous stepwise selection procedure, and then made statistical inference on whether estimated effects of land use on temperature varied across different quantiles, following the procedures proposed by Peng and Huang (2008). We repeated the procedure for each of three response variables (monthly averages of daily average, maximum and minimum temperatures) separately. Statistical analysis was conducted using SAS (version 9.2, SAS Institute Inc., Cary, NC, USA) and R statistical software (R Development Core Team; http://R-project.org).

3. Results

3.1. Descriptive statistics

Monthly averages of daily maximum, mean and minimum at 138 weather stations in August, 2011, were 40.63, 32.38 and 25.23 °C, respectively. Daily maximum temperatures across stations have the largest variation (33.33–45.93 °C), followed by daily minimum temperatures (21.67–29.63 °C) and daily average temperatures (30.00–36.30 °C). The lowest temperatures occurred