An Analysis of an Incomplete Marked Point Pattern of Heat-Related 911 Calls

Matthew J. HEATON, Stephan R. SAIN, Andrew J. MONAGHAN, Olga V. WILHELMI, and Mary H. HAYDEN

We analyze an incomplete marked point pattern of heat-related 911 calls between the years 2006–2010 in Houston, TX, to primarily investigate conditions that are associated with increased vulnerability to heat-related morbidity and, secondarily, build a statistical model that can be used as a public health tool to predict the volume of 911 calls given a time frame and heat exposure. We model the calls as arising from a nonhomogenous Cox process with unknown intensity measure. By using the kernel convolution construction of a Gaussian process, the intensity surface is modeled using a low-dimensional representation and properly adheres to circular domain constraints. We account for the incomplete observations by marginalizing the joint intensity measure over the domain of the missing marks and also demonstrate model-based imputation. We find that spatial regions of high risk for heat-related 911 calls are temporally dynamic with the highest risk occurring in urban areas during the day. We also find that elderly populations have an increased probability of calling 911 with heat-related issues than younger populations. Finally, the age of individuals and hour of the day with the highest intensity of heat-related 911 calls varies by race/ethnicity. Supplementary materials are included with this article.

KEY WORDS: Differential vulnerability; Imputation; Kernel convolution; Log-Gaussian Cox process; Missing data.

1. INTRODUCTION

1.1 Research Motivation

During the summer of 2011, Houston experienced one of the warmest, and driest, summers on record. The summer of 2011 in Houston saw the hottest day in city history (August 27), 27 daily record highs tied or broken, 33 daily high minimums, the highest temperature ever recorded for June, the earliest 38°C (100°F) day, and the most days of greater than 41°C (105°F) temperatures in a year, to name only a few (see http://www.srh.noaa.gov/hgx/?n=Summer2011Records). Extreme heat scenarios such as these pose a threat to public health. For example, various studies have correlated extreme heat to increases in mortality (Peng et al. 2011; Heaton et al. 2014), emergency room visits (Rusticucci and Bettolli 2002), and other diseases (Li et al. 2012).

Given the potential negative consequences of high heat, this study uses 911 call data in Houston to understand demographics, locations, and temperatures that make individuals vulnerable to a heat-related illness. Ultimately, by understanding these conditions, demographics and locations, appropriate intervention strategies (e.g., education programs, heat advisories, construction of cooling centers, etc.) can be implemented to prepare for such situations and, we hope, reduce the occurrence of potential negative outcomes.

Matthew J. Heaton is Assistant Professor, Department of Statistics, Brigham Young University 204 TMCB, Provo, UT 84602 (Email: mheaton@stat.byu.edu). Stephan R. Sain (Email: ssain@ucar.edu) is Scientist, Andrew J. Monaghan (Email: monaghan@ucar.edu) is Scientist, Olga V. Wilhelmi (Email: olgaw@ucar.edu) is Scientist, and Mary Hayden (Email: mhayden@ucar.edu) is Scientist, National Center for Atmospheric Research, Box 3000, Boulder, CO 80307-3000. This work was supported by NASA contract award NNX10AK79G. Procedures for analysis and visualization of the 911 data were reviewed and approved by the National Center for Atmospheric Research IRB. The National Center for Atmospheric Research is managed by the University Corporation for Atmospheric Research under the sponsorship of the National Science Foundation. The authors wish to thank Vishnu Nepal and Deborah Banerjee from the Houston Department of Health and Human Services for providing the 911 and impervious surface data, Tamara Greasy and Christopher Uejio for helpful discussions and three anonymous reviewers for their time and effort in improving the article.
1.3 Methodological Overview and Contributions

Our primary objective in analyzing the described 911 call data is to investigate the trends in 911 call data and identify weather conditions, spatial locations, times of day, etc. that are associated with heat-related 911 calls. We emphasize that we are primarily interested in investigating conditions (e.g., spatial locations, temperatures, etc.) that are associated with an increased likelihood of calling 911 for a heat-related health issue. That is, this analysis focuses on the relationship between, for example, temperature and heat-related 911 calls rather than call volume. In this respect, this analysis is similar to vulnerability analyses of Dolney and Sheridan (2006), Golden et al. (2008), Bassil et al. (2011), and Uejo et al. (2011). A deeper understanding of trends in the 911 call data will, for example, allow officials to locate spatial regions and temperatures associated with an increased risk of a 911 call.

A secondary goal of this analysis is to develop a public health data-based tool to predict the number, spatial locations, and time of incidence of heat-related 911 calls. By building a prediction tool, we can predict 911 call volume given a weather forecast and appropriately give advisories or enact other intervention strategies. Broadly, both of these goals are considered to be under the umbrella of public health surveillance (Brookmeyer and Stroup 2003; Lawson and Kleinman 2005) for heat-related morbidity (Kovats and Hajat 2008; Hajat and Kosatky 2010; Wilhelmi, de Sherbinin, and Hayden 2012).

To accomplish these goals, our strategy is to develop a statistical model for 911 calls. One common approach to modeling such data is to aggregate (count) the number of calls within an areal unit and use a Poisson spatial autoregressive model, say, for the spatial structure (see, e.g., Besag 1974; Waller et al. 1997). However, given the intricacies of the 911 call data, such an approach would result in some information loss. For example, aggregating the 911 calls in this manner, while still likely to produce interesting inferences, would forfeit the information in the 911 call data at sub-areal unit spatial scales and, effectively, discretize a continuous spatial domain. Notably, the degree of information loss could be controlled based on the resolution of the areal units but our preference is to treat the spatial domain as continuous.

Our strategy is to statistically model the 911 calls described in Section 1.2 as a marked point pattern arising from a non-homogenous Poisson process (NHPP; Diggle 2003, Baddeley 2010). Let \( z \in \mathcal{Z} \) denote a vector containing the spatio-temporal and marked location of a 911 call. The intensity surface of a NHPP, \( \Lambda(z) \), is a nonnegative function and determines the number and marked location of the 911 calls. That is, the number of 911 calls over the domain \( \mathcal{Z} \) follows a Poisson distribution with mean \( \int_{\mathcal{Z}} \Lambda(z)dz \) and the corresponding locations (or marks) are determined by the normalized intensity \( \Lambda(z)/(\int_{\mathcal{Z}} \Lambda(z)dz) \). The inherent advantage of our point process approach is that we appropriately treat the spatial domain as continuous over the city of Houston and, hence, avoid the issues related to aggregation mentioned above.

Under our NHPP approach, the intensity surface \( \Lambda(z) \) is the primary parameter of interest that we wish to learn from the data. That is, by estimating \( \Lambda(z) \) we will be able to (i) perform inferences for locations and times that are associated with increased public health vulnerability and (ii) make predictions of heat-related 911 calls for a given temperature scenario. Specifying a valid model for \( \Lambda(z) \) over the domain \( \mathcal{Z} \) is challenging for a few reasons. For example, (i) the day of the year and hour of the day are circular variables and (ii) the spatial features (e.g., urban canyons, rivers, etc.) of the city of Houston are intricate requiring a high dimensional spatial model. To deal with both of these issues, we induce a log-Gaussian process for \( \Lambda(z) \) using circular kernels in a discrete kernel convolution framework (Higdon 2002).

A particular challenge for analyzing the 911 calls is the presence of incomplete observations (i.e., missing marks). A complete case analysis would reduce the number of observations by 32% which would severely affect our estimate of the total number of heat-related 911 calls (thereby, biasing predictions). Rather than discarding the information in the incomplete cases, we include the incomplete cases by marginalizing \( \Lambda(z) \) over the domain of the missing marks and, subsequently perform model-based imputation for the missing marks via compositional sampling.

This article is outlined as follows. Section 2 defines a hierarchical NHPP model for the 911 call data and details our statistical modeling strategy. Section 3 discusses the methods for estimating \( \Lambda(z) \) including imputation of missing marks. Section 4 discusses results and Section 5 concludes. Supplementary material for this article is available online.
where \( w(\cdot) \) is a known nonnegative weighting measure that captures features of the population (e.g., population density, age distribution, etc.), \( \lambda(\cdot) \) is an unknown nonnegative density with respect to \( w(\cdot) \) such that \( \int \lambda(z)w(z)dz = 1 \) and models deviations of the intensity surface \( \Lambda(z) \) from the known features of the population \( w(z) \). Finally, \( \delta \) is an unknown constant controlling the expected number of 911 calls.

From (1), we emphasize the following. First, by construction, the expected number of 911 calls over the entire domain \( Z \) is given by \( \int \Lambda(z)dz = \delta \int \lambda(z)w(z)dz = \delta \) where we use the integral notation loosely to denote integrating over continuous and discrete domains. In this way, \( \delta \) controls the total number of 911 calls over \( Z \). Second, \( \lambda(z) \) captures deviations of the intensity surface from the known population characteristics given by \( w(z) \). Notice if \( \lambda(z) \equiv \delta \) for all \( z \in Z \), then \( \Lambda(z) = \delta w(z) \) such that the intensity surface follows the characteristics of the population (e.g., the highest volume of calls will occur in areas of high population density). In this way, \( \lambda(z) \) represents the “relative-risk” of \( z \) with respect to the population characteristics.

For the point pattern of heat-related 911 calls described above, let \( z_i = (s_i, t_i, x_i) \) where the vector \( s_i \) denotes the spatial location of the \( i \)-th 911 call for \( i = 1, \ldots, 2054 \), \( t_i = (y, d, h) \) denotes the temporal location (i.e., the year, \( y \), day of the year, \( d \) where leap days are excluded, and hour of the day, \( h \)) of the \( i \)-th 911 call, and \( x_i = (a, g, e, C) \); denotes the attribute vector of the \( i \)-th 911 call with attributes age (\( a \)), gender (\( g \)), race/ethnicity (\( e \)), and temperature given in degrees Celsius (\( C \)). We treat \( \{s_i\} \) as a point pattern over the spatial domain \( S \), \( \{t_i\} \) as a point pattern over the temporal domain \( T \) and \( \{x_i\} \) as a point pattern over the attribute space \( \mathcal{X} \). The spatial domain \( S \) is taken to be the Houston city limits displayed in Figure 1: \( T = \mathcal{Y} \times D \times H \) where \( \mathcal{Y} = \{2006, \ldots, 2010\} \) are the years, \( D = \{1, \ldots, 365\} \) are the days of the year, and \( H \in \{0, \ldots, 23\} \) are the hours of the day; and, \( \mathcal{X} = \mathcal{A} \times \mathcal{G} \times \mathcal{E} \times \mathcal{C} \) where \( \mathcal{A} = \{0, \ldots, 110\} \) denotes the set of plausible ages, \( \mathcal{G} = \{F, M\} \) is the set of genders, \( \mathcal{E} \in \{ \text{Caucasian}, \text{African American}, \text{Asian} \} \) is the set of ethnicities, and \( \mathcal{C} = \{-\infty, \ldots, 46\} \) is the set of temperatures. The upper limit for the age set \( \mathcal{A} \) and both upper and lower limits for the set \( \mathcal{C} \) were chosen as reasonable end points. For example, the bounds chosen for \( \mathcal{C} \) reflect record temperatures where the lowest temperature ever recorded in Houston is \(-15^\circ \text{C}\) on January 23, 1930 and the highest temperature ever recorded was \(43^\circ \text{C}\) on August 27, 2011.

We define the weighting measure \( w(z) \) to reflect a “null hypothesis” of no change in 911 call volume from the population characteristics of Houston (Best, Ickstadt, and Wolpert 2000). Details on how we specified \( w(z) \) are given in the supplementary material but we highlight the primary features here. First, we assign uniform weights to all years (\( y \)), days of the year (\( d \)), and hours of the day (\( h \)) to reflect a “null hypothesis” that 911 calls for heat-related illness can occur at any time. Because at the time of the 2010 U.S. Census, Houston was 50.2% male and 49.8% female, we assume equal weights for male and female callers. We chose the spatial weights to vary by ethnicity where the weight of location \( s \) is the number of people (in thousands) of race/ethnicity \( e \) located in the census block group in which \( s \) resides (population numbers were obtained from the U.S. 2010 Census). We chose weights for temperatures to reflect the seasonal temperature changes in Houston. Obtaining data from the National Climatic Data Center (NCDC), the weight of temperature \( C \) on day \( d \) is set be the height of a Gaussian density function with mean equal to the 30-year average, average temperature on day \( d \) and standard deviation equal to 1/2 the 30-year average diurnal temperature range. Finally, the weights for age and race/ethnicity were set according to the estimated proportions from the 2010 U.S. Census.

For the deviation surface \( \lambda(z) \), we let

\[
\lambda(z) = \lambda(s, y, d, h, a, g, e, C) = \lambda_{HS}(s, h \mid e)\lambda_{ADC}(a, d, C \mid e)\lambda_{G}(g \mid e)\lambda_{E}(e),
\]

where each of \( \lambda_{HS}(s, h \mid e) \), \( \lambda_{ADC}(a, d, C \mid e) \), \( \lambda_{G}(g \mid e) \), and \( \lambda_{E}(e) \) describe deviations in \( \Lambda(z) \) from the respective population weights described in \( w(z) \). For example, \( \lambda_{HS}(s, h \mid e) > 1 \) indicates that the intensity of 911 calls at spatial location \( s \) and hour \( h \) for race/ethnicity \( e \) is higher than population characteristics dictate (i.e., more 911 calls occur at \( s, h \) than expected given the population demographics). The form for \( \lambda(z) \) in (2) was formulated primarily from the exploratory analysis of the 911 call data given in the supplementary materials. While we emphasize the following key features of \( \lambda(z) \) in (2) here, we refer to the supplementary materials for detailed justification. First, because year-to-year trends in the volume of 911 calls are not of primary interest for this study, we assumed \( \lambda(z) \) is not a function of \( y \) (i.e., \( \Lambda(z) \) is uniform with respect to year). Second, Equation (2) allows the intensity surface of every other mark to vary by race/ethnicity. This is reflected by the conditioning on \( e \) in (2) and, intuitively, means that race/ethnicity drives the spatial location and mark of the 911 call. Third, conditional on race/ethnicity, \( (h, s), (a, d, C) \), and \( g \) are modeled separately. In other words, we assume that there are no interactions between \( (h, s), (a, d, C), \) and \( g \) conditional on \( e \) in the intensity function.
2.2 Process Model

We desire a computationally tractable, yet flexible modeling strategy for each of $\lambda_{HS}$ and $\lambda_{ADC}$. Here, we chose to rely on the discrete process convolution (DPC) approach of Higdon (2002) to construct a log-Gaussian Cox process model (Brix and Diggle 2001; Isham 2010) for the 911 calls. The DPC approach of Higdon (2002) assumes the basis function expansions,

\[
\log(\lambda_{HS}(h, s \mid e)) = \sum_{b=1}^{B_{HS}} K_{HS,b}(h, s \mid e) \eta_{b|e} (3)
\]

\[
\log(\lambda_{ADC}(a, d, C \mid e)) = \sum_{b=1}^{B_{ADC}} K_{ADC,b}(a, d, C \mid e) \eta_{b|e}, (4)
\]

where \(K_{HS,b}(h, s \mid e)\) and \(K_{ADC,b}(a, d, C \mid e)\) are sets of kernel basis functions, \(B_{HS}\) and \(B_{ADC}\) are the number of basis functions, and \(\eta_{b|e}\) are ethnicity-specific coefficients modeled as Gaussian random variates. The kernel basis functions \(K_{HS,b}(h, s \mid e)\) and \(K_{ADC,b}(a, d, C \mid e)\) serve to smooth the white-noise processes \(\eta_{b|e}\) over the domains $\mathcal{H} \times S$ and $A \times D \times C$, respectively.

Using the DPC approach, we can easily constrain $\lambda_{HS}(h, s \mid e)$ and $\lambda_{ADC}(a, d, C \mid e)$ to circular domains through proper choice of kernel functions. To do so, we assume the kernel function $K_{HS,b}(h, s \mid e)$ to be the product of a Gaussian and von Mises kernel. Specifically, we assume,

\[
K_{HS,b}(h, s \mid e) \propto \exp \left\{ \frac{-||s - s^*_b||}{2\sigma^2_{se}} \right\} \exp \left\{ \frac{1}{2\sigma^2_{He}} \cos \left( \frac{2\pi}{24} |h - h^*_b| \right) \right\}, (5)
\]

where \(s^*_b, h^*_b\) are knot locations for the basis functions and \(\sigma^2_{se}\) and \(\sigma^2_{He}\) are smoothing parameters estimated from the data. Note that (5) appropriately constrains the process with respect to the circular domain of $h$ (hour of the day) through the use of the von Mises kernel (the second exponent term on the right-hand-side of (5)). That is, (5) appropriately correlates the intensity surface at hour $h = 0$ with hour $h = 23$. The knot locations \(s^*_b, h^*_b\) were taken to be an equidistant grid over the spatial and hour of the day product space $S \times \mathcal{H}$. We defined the kernel basis functions \(K_{ADC,b}\) in a similar fashion as Equation (5) where a Gaussian kernel was used for the age and temperature domains and a von Mises kernel was used for the day of the year domain. We refer to the supplementary materials for the technical details.

The specification in (3) and (4) allows the surfaces $\lambda_{HS}$ and $\lambda_{ADC}$ to be completely determined by a finite set of coefficients. This feature allows tractable computations in likelihood evaluations. We chose to use $B_{HS} = B_{ADC} = 1500$ which resulted in a sufficiently flexible model that is computationally feasible to work with.

2.3 Hyperprior Layer

Define $\eta$ and $\gamma$ to be the \((B_{HS} \times 4) \times 1\) and \((B_{ADC} \times 4) \times 1\) vectors of all $\eta$ and $\gamma$ coefficients from (3) and (4). The hyperpriors for $\eta$ and $\gamma$ under the DPC construction assumes $\eta \sim N(0, \sigma^2_{He}(R_{\mathcal{E}} \otimes I_{B_{HS}}))$ and $\gamma \sim N(0, \sigma^2_{He}(R_{\mathcal{E}} \otimes I_{B_{ADC}}))$ where $I_n$ is the rank $n$ identity matrix and $R$ is a correlation matrix. Notice that, under this hyperprior, the coefficients are allowed to be correlated so as to borrow information across race/ethnicity.

For \(\{\lambda_{g}(g \mid e) : g \in \mathcal{G}, e \in \mathcal{E}\}\) we assume $\lambda_{g}(M \mid e) \sim U(0, 2)$ independently for all $e$ such that, a priori, $E(\lambda_{g}(M \mid e)) = 1$ representing our null hypothesis that the intensity measure follow population characteristics. To induce a hyperprior for $\lambda_{g}(e)$ with the constraint that $\sum \lambda_{g}(e) w_{g}(e) = 1$ where $w_{g}(e)$ denotes the population weight for race/ethnicity $e$, we set

\[
\tilde{\lambda}_{g}(e) = \frac{1}{w_{g}(e)} \sum \lambda_{g}(e) \quad (6)
\]

where $\tilde{\lambda}_{g}(e)$ is $	ext{Gam}(w_{g}(e), 1)$ independently for each $e$ and $\text{Gam}(a, \beta)$ denotes the gamma distribution with shape $a$ and rate parameter $\beta$. By construction, $\tilde{\lambda}_{g}(e)/(\sum \tilde{\lambda}_{g}(e))$ follows a Dirichlet distribution with $E(\tilde{\lambda}_{g}(e)/(\sum \tilde{\lambda}_{g}(e))) = w_{g}(e)$; hence, (6) appropriately corresponds to a “null hypothesis” hyperprior that the ethnicities of 911 callers should correspond to the population of Houston.

Due to the equivalence of probability measures (Zhang 2004), we chose to use independent discrete hyperprior distributions for the bandwidth parameters in (5). As the hyperprior distribution for $\sigma^{-2}_{He}$ for example, we let $\sigma^{-2}_{He} \sim DU[0.5, \Delta, \ldots, 10\Delta]$ where $DU[\cdot]$ denotes the discrete uniform distribution and $\Delta = h_{b+1} - h_{b}$ is the common distance between knot locations. Jeffreys hyperprior distributions were used for the correlation matrices $R_{\mathcal{E}}$ and $R_{\mathcal{H}}$ (see Zhang, Boscardin, and Belin 2006) but other choices include hierarchical and shrinkage hyperpriors such as those suggested by Daniels and Kass (1999, 2001). Informally comparing hyperpriors, we found that the choice had little impact on the posterior distribution for $\lambda$. We also used a Jeffreys hyperprior $\delta, \sigma^{-2}_{He},$ and $\sigma^{-2}_{He}$.  

3. PARAMETER ESTIMATION AND IMPUTATION

For each call, partition $z_i = (z^O_i, z^M_i)$ where $z^O_i$ corresponds to the observed part of $z_i$ and $z^M_i$ corresponds to the missing part of $z_i$ and let $\theta$ denote the vector of parameters for the model described in Section 2. The complete case likelihood for $\theta$ is given by,

\[
L([z_i] \mid \theta) = \exp \left\{ -\int_{S \times \mathcal{E}} \sum_{e \in \mathcal{E}} \Lambda(s, t, x \mid \theta) ds \right\} \prod_{i=1}^{2054} \Lambda(z_i \mid \theta) 
\]

\[
= \exp \left\{ -\delta \right\} \delta^{2054} \prod_{i=1}^{2054} \lambda(z_i \mid \theta) w(z_i), (7)
\]

where 2054 are the total number of 911 calls within Houston city limits over the 5-year period 2006-2010 (see Diggle 2003; Möller and Waagepeterson 2004). We emphasize that calculating (7) requires $\{z_i\}$ to be fully observed whereas, in this analysis, 32% of $\{z_i\}$ are only partially observed. Discarding the incomplete observations would forfeit valuable information as well as drastically reduce the estimation of the total number of 911 calls over the study period.

A likelihood for $\theta$ using only $\{z^O_i\}$, is obtained by marginalizing (7) with respect to $\{z^M_i\}$. Let $Z = Z^O \times Z^M$ such that
\[ z^O_i \in Z^O_i \text{ and } z^M_i \in Z^M_i. \] The marginal intensity for \( z^O_i \) is

\[ \Lambda_i(z^O_i \mid \theta) = \delta \sum_{z^M_i \in Z^M_i} \lambda(z^O_i, z^M_i \mid \theta) w(z^O_i, z^M_i), \] (8)

where we can use the summation notation because the missing marks for this application are always defined over a discrete space (spatial locations are always known). Inserting \( \Lambda_i(z^O_i) \) into the likelihood function in (7) yields a likelihood function using only observed data and is given by

\[ L(\{z^O_i\} \mid \theta) = \exp \{-\delta\} \prod_{i=1}^{2054} \lambda_i(z^O_i \mid \theta) \]

\[ = \exp \{-\delta\} \delta^{2054} \sum_{i=1}^{2054} \sum_{z^M_i \in Z^M_i} \lambda(z^O_i, z^M_i \mid \theta) w(z^O_i, z^M_i). \] (9)

Despite having marginalized over the missing values \( z^M_i \), we can still impute \( \{z^M_i\} \) by sampling \( \{z^M_i\} \) from the conditional density \( \Lambda_{\text{Miss}}(z^M_i \mid z^O_i, \theta) \propto \lambda(z_i \mid \theta) w(z_i) \).

The above method for accounting for incomplete observations assumes that the data are missing at random (MAR; Little and Rubin 2002). Alternatively, if the data are not missing at random (NMAR), the missing data mechanism would need to be included in the likelihood in (9). Building such a model for the missing data mechanism is beyond the scope of this article and requires additional data (or strong prior assumptions) which are not available to us.

One challenge in calculating the likelihood in (9) is enforcing the density constraints for each component in (2). For example, the model constrains \( \int_S \sum_{h \in H} \lambda_{HS}(s, h \mid e) \, w(s \mid e)(24^{-1})ds = 1 \) for all \( e \). Enforcing such density constraints for \( \lambda_{ADF}, \lambda_D, \) and \( \lambda \) is a straightforward normalization because these are defined over discrete spaces. For \( \lambda_{HS} \), however, this normalization cannot be done exactly because \( S \) is a continuous domain. Hence, to ensure \( \lambda_{HS}(s, s \mid e) \) is a density we rely on the discrete Riemann approximation of the normalizing constant used by Liang, Carlin, and Gelfand (2009).

To obtain estimates of the parameters, we ran a Markov chain Monte Carlo (MCMC) algorithm to obtain 100,000 draws from the posterior distribution after a burn-in of 20,000 iterations. Note that the likelihood in (9) combined with the form for \( \lambda(z) \) in (2), factors into independent pieces conditional on race/ethnicity. This convenient forms allows for race/ethnicity specific parameters to be updated in parallel. Using parallel processing, one iteration using MATLAB took 0.35 sec on a 32 core Intel Xeon system with 512 GB of memory environment. Hence, one chain took approximately 12 hr of computation time to obtain the desired number of posterior draws. The Gelman and Rubin (1992) diagnostic on \(-2 \times \log\)-likelihood for 5 independent chains was used to assess convergence and indicated that the 20,000 iterations of burn-in was sufficient. Mixing of the post-burn draws was evaluated using trace plots. Accuracy of point estimates of posterior means and 0.025 and 0.975 quantiles were monitored using the mcmcse package in R(Flegal and Hughes 2012). The largest Monte Carlo standard error was associated with the posterior mean of \( \delta \) at a value 0.12.

4. RESULTS

4.1 Model Comparison

The fit of the full model described in Section 2 was compared to a null model, separable model, and a separable-in-ethnicity model using DIC. Each model is described in more mathematical detail in the supplementary materials but, briefly, the null model assumes \( \lambda(z) = 1 \) for all \( z \) so that the weighting measure \( w(z) \) adequately explains 911 calling patterns, the separable model assumes that \( \lambda(z) \) is separable (i.e., modeled additively on the log-scale) in all marks and, finally, the separable-in-ethnicity model is of the same form as (2) but modified so that each deviation surface is common across race/ethnicity. Each of these models represent increasing complexity such that subsequent comparison determines the degree of complexity necessary to explain the 911 call data.

The DIC values for the null, additive, separable-in-ethnicity, and full model were 122,200; 39,361; 37,971; and 37,235; respectively, suggesting that the null model is clearly inadequate and that the full model is preferred. The preference for the full model is substantiated by the fact that the full model had 4147 effective parameters compared to 85 for the separable and 1039 for the separable-in-ethnicity model. Hence, the increase in the number of parameters is justified by an increase in model fit. While the full model is preferred in terms of DIC, the decrease from the separable-in-ethnicity to the full model is small suggesting that including an additive race/ethnicity effect for the log-deviation surface may be an appropriate simplifying assumption.

4.2 Model Validation

We validated the full model fit by holding out the 2010 calls, fitting the model in Section 2 to data collected in 2006-2009 and subsequently predicting the locations of the 2010 calls according to a temperature scenario. Generating predictions under our point process model requires a temperature field for each day on which predictions are to be made. That is, locations, ages, genders, etc. of 911 callers are generated conditional on a temperature scenario. To generate a predicted point pattern of heat-related 911 calls, let \( R \) denote the joint region for \( (d, C) \) that corresponds to a predicted temperature scenario. Bayesian posterior simulation of a predicted point pattern occurs by (i) drawing the total number of 911 calls from a Poisson distribution with mean \( \Lambda_R = \sum_{i \in R} \Lambda(d, C) \) where \( \Lambda(d, C) \) is the intensity for \( (d, C) \) after marginalizing over the other marks and (ii) drawing the spatial locations, time periods and marks according to the conditional intensity \( \Lambda(z \mid (d, C) \in R) \) for each draw of \( \theta \) from the posterior distribution. Because \( s \) is continuous over the spatial domain \( S \), we draw \( s \) by discretizing \( S \) into 1301 discrete locations given by the centroids of the 2010 census block groups within city limits (see Figure 1).

While high-resolution, hourly temperature predictions across Houston may be available on a short-time scale, high-resolution predictions for an entire summer are not typically available for predictive purposes. Indeed, the HRLDAS model cannot be used for predictive purposes because it is constrained by observational data. However, it is reasonable to assume that regional, daily maximums and minimums could be estimated...
from historical data and used for predictive purposes. For this reason, we used the daily minimum and maximum temperatures from the IAH airport weather station for June, July and August 2010 from the NCDC website (http://www.ncdc.noaa.gov/) as the 2010 temperature scenario.

Figure 2(a) displays a histogram of the posterior predictive draws of the total number of heat-related 911 calls during summer 2010 where the vertical bar indicates the observed number of calls (229). Clearly, our model is useful in predicting the total number of calls that will occur under summer 2010 temperatures because the observed value of 229 falls in the center of our predictive distribution.

To validate our model at finer spatial scales, let \( N_b \) and \( \tilde{N}_b \) represent the observed and predicted number of calls in block group \( b \) during the summer of 2010, respectively. Figure 2(b) displays the average predictive probability \( p_k(n) = \frac{1}{n} \sum_{b:N_b=k} \Pr(\tilde{N}_b = n \mid \text{data}) \) where \( n_b \) is the number of block groups such that \( N_b = k \). Strong predictive accuracy is indicated by large probabilities along the increasing (left-to-right) diagonal. Figure 2(b) suggests that the model predicts well when \( N_b = 0 \) because \( p_0(0) \) is high when \( N_b = 0 \). Also, encouragingly, the model frequently places positive posterior predictive mass at the observed \( N_b \) where \( N_b > 0 \) (probabilities along increasing diagonal). For example, the block groups where \( N_b = 1 \) had an average predictive probability \( p_1(1) \) of approximately 0.45. Qualitatively, however, the model seems to under-predict higher counts (e.g., counts greater than 1) but \( N_b > 1 \) only occurred in 2% of the hold-out sample.

Results from Figure 2 indicate that our model accurately predicts the total number of calls and frequently places substantial posterior predictive mass at the observed \( N_b \). We can substantiate the statistical learning in our model further by comparing \( \Pr(\tilde{N}_b = N_b \mid \text{data}) \) to the same probabilities using a “null model” that assumes heat-related 911 calls are completely explained by population demographics (i.e., a model where \( \lambda(z) \equiv 1 \) for all \( z \)). On average the posterior predictive probability \( \Pr(\tilde{N}_b = N_b \mid \text{data}) \) under the full model was 6.9 times larger than under the null model suggesting successful statistical learning and improved prediction.

To validate predictive accuracy using interval estimators rather than point estimators, let \( \mathcal{P}_b \) represent a 95% central posterior predictive set for block group \( b \) where \( \Pr(\tilde{N}_b \in \mathcal{P}_b \mid \text{data}) \approx 0.95 \). For predicting heat-related 911 calls during summer of 2010, 94.2% of \( \mathcal{P}_b \) contained the observed \( N_b \) suggesting appropriate coverage of the predicted values.

4.3 Estimated Intensity Surface

Using all the data from 2006–2010, Figure 3(a) displays the posterior mean of \( \log(\lambda_{A,\sigma}(a, d, C \mid \epsilon)) \) averaged over age and race/ethnicity. Positive (negative) values in Figure 3(a) correspond to increased (decreased) relative risk with respect to the weighting measure. Not surprisingly, high temperatures during the summer exhibit the strongest increased risk from the baseline measure \( w(\cdot) \). Interestingly, note that the regions of greatest increased risk are above the average daily temperature. Investigating this further, let \( \Pr(N > 0 \mid C, d) \) denote the probability of at least one heat-related 911 call for a given temperature (\( C \)) and day of the year (\( d \)) and \( \overline{C}_d \) denote the average temperature on day \( d \). Figure 3(b) displays the posterior mean of the number of degrees above average (say, \( \epsilon \)) such that \( \Pr(N > 0 \mid C = \overline{C}_d + \epsilon, d) \) is 25% larger than \( \Pr(N > 0 \mid C = \overline{C}_d, d) \). From Figure 3(b), notice that on January 1, temperatures need to be about 6–9 degrees above average to increase \( \Pr(N > 0 \mid C = \overline{C}_d, d) \) by at least 25%. On July 1, however, a 1-degree increase in the temperature can lead to a 25% increase in \( \Pr(N > 0 \mid C, d) \). Hence, in terms of public
health, slightly above average summer temperatures can lead to an increased risk of a heat-related 911 call.

Panels (a) and (b) in Figure 3 are in terms of air temperatures. However, humidity is also an important factor when considering the effect of weather on public health (Smoyer-Tomic and Rainham 2001; Rainham and Smoyer-Tomic 2003). Perhaps for this reason, the National Weather Service (NWS) uses a heat index (a combination of temperature and relative humidity) threshold of 42°C to issue heat advisories. To validate this threshold, we also fit the model in Section 2 using the NWS heat index in place of temperature from the HRLDAS model output. Figure 3(c) displays the posterior predictive probability of at least one heat-related 911 call as a function of day of the year and heat index (note that Figure 3(c) is restricted to summer months because the NWS heat index is only well defined during this time period). Notably, our results indicate that the current threshold of 42°C (108°F) is too high. That is, Figure 3(c) shows that the highest probability of at least one 911 call occurs toward the end of July at a heat index value of 39.5°C (103°F). This result has prompted discussion among Houston health officials regarding appropriate weather conditions at which to issue a heat advisory.

Figure 4 displays the posterior mean of $\lambda_{ADC}(a, d, C | e)$ for $h = 9$ (between 8:31–9:29 AM) and $h = 16$ (between 3:30–4:29 PM) and all ethnicities where $\lambda_{ADC}(s, h | e) > 1$ is interpreted as an increased risk. Hour 9 occurs a few hours after the daily warming cycle begins (minimum temperatures occur around 6:00–7:00 am), whereas hour 16 is about an hour after the warmest time of the day throughout the city. From Figure 4, the risks for $h = 16$ are, on average, higher than for $h = 9$ across all ethnicities suggesting that the hottest hours of the day pose a higher risk for heat-related 911 calls. Notice, also, that the regions of elevated risk during the $h = 9$th hr of the day seem to be concentrated in more residential areas than during the $h = 16$th hr (compare with Figure 1). This result seems to suggest that, during the work hours, urban areas present a higher risk for 911 calls. The exception to this pattern was with the Asian population that seems to have the opposite temporal trend than the other ethnicities (i.e., regions of elevated risk are more concentrated in urban areas in the morning rather than the afternoon).

Figure 5 displays the posterior mean of $\lambda_{ADC}(a, d, C | e)$ for all ethnicities averaged across day of the year where a value greater than 1 indicates elevated risk. Marginally, the elevated risk of older populations for heat-related 911 emergencies is apparent across all ethnicities and higher temperatures. That is, $\lambda_{ADC}(a, d, C | e)$ for $a > 80$ and $C > 35$ is large for all ethnicities. For the Hispanic population, there is statistically significant evidence that the average value of $\lambda_{ADC}(a, d, C | e)$ for ages $a \in [25, 40]$ was greater than 1 (in the sense that a 95% credible interval for the average value of $\lambda_{ADC}(a, d, C | e)$ in this region was greater than 1). This same increased risk was not significant for the Asian population. We hypothesize that this increase in risk is likely due to an underlying unmeasured sociological covariate (e.g., occupation or income level) but further investigation is required.

The posterior mean (95% credible interval) for the (normalized) gender intensity of male callers (i.e., the percent of male callers) for Caucasian, African American, Hispanic, and Asian populations was 0.58 (0.54, 0.62), 0.63 (0.60, 0.66), 0.74 (0.69, 0.78), 0.60 (0.56, 0.66), respectively. Each of these proportions differed from the population percentages of 50%, 47%, 51%, and 49% male Caucasians, African Americans, Hispanics, and Asians, respectively. Notably, the Hispanic population seems to exhibit different 911 calling patterns for heat-related calls than the other ethnicities by having a higher concentration of Male callers. Finally, the posterior mean of the (normalized) race/ethnicity intensity (i.e., distribution of calls by race/ethnicity) is 0.31, 0.44, 0.23, and 0.02 for Caucasian, African American, Hispanic, and Asian race/ethnicity, respectively, suggesting that African Americans are the most common heat-related 911 callers. These race/ethnic proportions are statistically different than those reflected in the population of Houston wherein Caucasians, African Americans, Hispanics.

![Figure 3](image-url)  
Figure 3. (a) Posterior mean of $\log(\lambda_{ADC}(a, d, C | e))$ averaged over age and race/ethnicity. Positive (negative) values correspond to increased (decreased) relative risk with respect to the weighting measure. The solid line corresponds to the average daily temperature according to records from the National Climatic Data Center (NCDC). The dashed line indicates the zero level contour. (b) Posterior mean (solid line) and 95% credible interval limits (dashed lines) of the number of degrees above average ($\varepsilon$) such that $Pr(N > 0 | C = C_{d}, d)$ is 25% larger than $Pr(N > 0 | C = T_{d}, d)$. (c) Posterior mean of the probability of at least one heat-related 911 call as a function of the NWS heat index. The dashed line corresponds to the current threshold of 42°C (108°F) which, when exceeded, a heat advisory is issued.
Figure 4. Posterior mean of the marginal intensity $\lambda_{HS}(s, h | e)$ for all ethnicities and 2 hr of the day. The rows correspond to Caucasian, African American, Hispanic, and Asian populations, respectively. The first column corresponds to the intensity at the $h = 9$th hr of the day (between 8:31–9:29 AM) and the second column corresponds to the intensity at the $h = 16$th hr of the day (between 3:30–4:29 PM).
and Asians make up 49.3%, 19.6%, 37%, and 7.0% of the population.

4.4 Comparison to Complete Case Analysis

For comparison, we also estimated model parameters using only the 1385 complete 911 calls. The greatest difference between the complete case and the full analysis was the difference in the posterior mean of $\delta$. In the complete case analysis, the posterior mean (95% credible interval) for $\delta$ was 1385.2 (1314.7, 1454.0) compared to 2054.7 (1960.5, 2138.2) in the full analysis. This result is expected because, in the complete case analysis, 669 calls were discarded leading to a gross underestimation of the total number of 911 calls over the domain $Z$.

The posterior mean of $\Lambda(z)$ and $\lambda(z)$ in the complete case analysis changed little from those displayed in Figures 3, 4, and 5. However, the width of credible intervals increased by using only complete cases. For example, the 95% credible interval for the percent of male African American callers was (0.60, 0.66) using all the data compared to (0.58, 0.67) using only complete cases.

4.5 Imputing Missing Marks

We demonstrate our model’s capability of imputing missing values for four hypothetical callers (imputation for the actual observations contains personal health information which cannot be shared publicly). We assume that these incomplete calls occur during the $h = 16$th hr of the day on August 1, 2010. The locations of these calls will be denoted by $s_1^M, \ldots, s_4^M$ and were taken to be the centroid of the 2010 census tract where the intensity $\lambda_{HS}(h = 16, s \mid e)$ achieved its maximum for each of the four ethnicities. Assume, also, that the temperature was known to be 38°C at the time of the call. For this example we assume that the marks for gender, age, and race/ethnicity are missing. Imputing the missing values for this scenario amounts to sampling the missing components from the normalized conditional intensity,

$$
\Lambda(a, g, e \mid s, d, h, C) \propto \Lambda(a, g, e, s = s_i^M, d = 227, h = 16, C = 38) \quad (10)
$$

for $i = 1, \ldots, 4$ where the normalization constant is calculated as $\sum_{a, g, e} \Lambda(a, g, e, s_i^M, d = 227, h = 16, C = 100)$. Sampling from (10) for each iteration of the MCMC algorithm results in an estimate of the posterior distribution for the missing marks.

Table 1 displays the posterior mean of the distribution for the race/ethnicity of each caller (i.e., the posterior mean of $\sum_{a, g, e} \Lambda(a, g, e \mid s, d, h, C)$). We emphasize that because the calls were chosen at the location of maximum intensity for each race/ethnicity, calls 1–4 should, roughly, coincide with each of the four different ethnicities. This fact is reflected in Table 1 as the maximum probability for callers 1–3 corresponds to the hypothesized race/ethnicity. For caller 4, note that the posterior probability of an Asian caller is 300% larger than the marginal probability of an Asian caller (0.02). This gives credence that information from the spatial location, hour of the day, day of
Table 1. Posterior mean of the distribution of race/ethnicity for the imputation example in Section 4.5

| Caller | Location ($s^M$) | Caucasian | African American | Hispanic | Asian |
|-------|------------------|-----------|------------------|----------|-------|
| 1     | ($-95.3632, 29.7351$) | 0.46      | 0.28             | 0.26     | 0.00  |
| 2     | ($-95.3527, 29.6957$) | 0.17      | 0.68             | 0.14     | 0.00  |
| 3     | ($-95.3212, 29.7482$) | 0.15      | 0.27             | 0.57     | 0.00  |
| 4     | ($-95.5208, 29.7088$) | 0.32      | 0.30             | 0.29     | 0.08  |

the year and temperature is being correctly used to impute the missing values.

4.6 Using the Statistical Model as a Public Health Tool

To demonstrate how our statistical model can be used as a public health tool for emergency preparedness, in this section we compare the predicted volume of 911 calls (along with all marks) under the temperature scenario during the summer of 2011 described in the Introduction to the observed number of calls during the same time period in 2010. Similar to the validation exercise in Section 4.2, we use the maximum and minimum daily temperatures at IAH as the summer 2011 heat scenario to generate predictions.

For the high temperatures experienced in 2011, the mean of the posterior predictive distribution for the number of heat-related 911 calls was 250 with a 95% predictive interval of (220, 273). In comparison, there were 229 heat-related 911 calls in 2010 such that, with probability 0.93, a summer similar to 2011 will have more heat-related 911 calls than observed in 2010. Stated differently, we predict the number of heat-related 911 calls under the weather scenario in 2011 to be approximately 10% more than the observed scenario in 2010.

As this research focuses on assessing public vulnerability to heat, it is important to understand the time of day, locations, and attributes of those who are predicted to call 911 for a heat-related health emergency. Figure 6 summarizes the simulated realizations of the 911 call point pattern under the summer 2011 temperature scenario. Figure 6(a) displays the average number of calls for each day of the summer and hour of the day. The simulated calls tended to cluster near the 15th hr (2:30-3:30 PM) which is approximately the hour where temperature reaches a maximum. An interesting feature in Figure 6(a) is the presence of a few days that did not realize a high volume of heat-related 911 calls (note the light vertical bands in Figure 6(a)). For example, June 22, on average, saw very few realizations of heat-related 911 calls. This is because the temperature profile for June 22, 2011, was relatively cool with a minimum of 20°C (68°F) and a maximum of only 29°C (84°F). Hence, the data-based tool is appropriately predicting a fewer number of 911 calls for June 22 due to lower temperatures.

In Figure 6(b) we classify block groups as having high, moderate or low volume based on the predicted number of heat-related 911 calls within the block group after controlling for population. That is, Figure 6(b) classifies a block group based by its rank of the total number of simulated 911 calls within the block group divided by the total population. One strategy implemented by Houston to counteract the negative effects of high heat is to construct cooling centers that are available for public use (see http://www.houstonoem.org/go/doc/4027/1871258/City-of-Houston-Cooling-Center-Locations). Using this map, public officials may be able to pinpoint high risk areas that may benefit from a cooling center.

Table 2 displays the average number of calls for various age categories and all ethnicities. By considering the predicted ages and ethnicities of callers, public health officials will be able to
more effectively design intervention strategies for each group. We predict that the age of Hispanic callers will be younger than those of Caucasian callers suggesting, perhaps, a different form of intervention for these two ethnicities. Comparing the numbers to the observed number of calls in 2010, we predict the largest increase in calls by African American and Hispanics in the (30, 40] age group.

### 5. DISCUSSION

In this article, we developed a tool for performing inference on the number and marked location of 911 calls for specified time periods and weather scenarios using an incomplete marked point pattern of 911 calls within the city of Houston. We relied on kernel convolution models to model the intensity surface of the observed point pattern and constrain the marks over the appropriate circular and cylindrical domains. To incorporate information from incomplete records, we marginalized the full intensity over the domain of the missing marks yielding a marginal intensity for the observed marks. By including incomplete cases, our estimate of the total number of 911 calls was more accurate and the width of credible intervals decreased.

The effect of heat stress on the public is manifest in many different forms such as emergency room visits, 911 calls, and, in some cases, mortalities. We emphasize that this research analyzes the incidence of a heat-related 911 call which is but one manifestation of a heat-related public health problem and the cumulative impact of heat on public health is more broad. Furthermore, this analysis focused on 911 calls where “heat,” “dehydration,” “hyperthermia,” “heat exhaustion,” and “heat stroke” were listed as a diagnosis. We admit that this is a restrictive definition but extending our definition of a “heat-related” 911 call, subsequent analyses of the cumulative impact of heat could simultaneously analyze the many different types of public health data. This is part of the ongoing work associated with the SIMMER project.

Many extensions of our model are possible which we leave for future work. For example, several studies have found lagged temperatures to be associated with increases in negative public health outcomes (see, e.g., Hajat et al. 2005; Anderson and Bell 2009; Heaton and Peng 2012). However, including lagged effects is a difficult problem for this application because we don’t have information on the whereabouts of the 911 caller prior to the call. Hence, any lagged temperature estimate would be an indirect proxy of the patient’s heat exposure. Likewise, considering lagged effects of a day or more would necessitate including characteristics of the patient’s home, such as the presence of air conditioning, which would certainly influence the timing and spatial location of a call.

As another extension, humidity and various heat metrics (e.g., the heat index, apparent temperature, etc.) could also be included because these also may affect the occurrence of a 911 call. Recent studies by Kovats and Hajat (2008), Gosling et al. (2009), and Heaton et al. (2014) have formally compared heat indices to raw temperatures and concluded that raw temperatures can be more explanatory of heat-related illnesses than heat indices. Rather than including only a single measure of heat, multiple heat indices could be included but this greatly increases the complexity of the model because these alternate heat metrics would need to be modeled nonseparately from the temperature intensity surface.

As mentioned in Section 4.2, model predictions rely on accurate portrayals of future temperatures. For the prediction examples described here, the temperature scenarios were simplified daily minimum and maximums at IAH and assumed to be constant across the city. While high-resolution, accurate temperature forecasts may be available for near future predictions (1–5 days), constructing accurate temperature scenarios over a long temporal range is difficult. Furthermore, any predicted temperature scenario is also associated with a degree of uncertainty. Similar to predictions of future climate (Tebaldi and Knutti 2007), incorporating the uncertainty in temperature scenarios into model predictions could be done by averaging predictions across various temperature scenarios weighted by the likelihood of each scenario.

---

Table 2. Mean number of calls by age group and race/ethnicity across simulated realizations of 911 calls using the summer 2011 temperature scenario compared to the observed number of calls in each category in 2010. Ninety-five percent predictive intervals are given in parentheses. Note that 229 calls occurred in 2010 but only 192 are displayed due to incomplete information.

| Age Group | Caucasian | African American | Hispanic | Asian | Sum |
|-----------|-----------|------------------|----------|-------|-----|
| (0,10]    | 0         | 1 (0.4)          | 0        | 0     | 1   |
| (10,20]   | 2         | 3 (0.6)          | 7 (3.13) | 6     | 7 (2.13) |
| (20,30]   | 5         | 6 (2.11)         | 12       | 1     | 14 (8.22) |
| (30,40]   | 7         | 10 (5.17)        | 16       | 6     | 12 (6.20) |
| (40,50]   | 9         | 13 (7.20)        | 15       | 9     | 6 (2.12) |
| (50,60]   | 14        | 14 (8.22)        | 13       | 4     | 3 (0.7) |
| (60,70]   | 10        | 11 (5.18)        | 11       | 3     | 2 (0.6) |
| (70,80]   | 4         | 8 (3.16)         | 6        | 3     | 2 (0.6) |
| (80,90]   | 6         | 7 (2.12)         | 3        | 1     | 1 (0.3) |
| (90,110]  | 1         | 3 (0.7)          | 1        | 0     | 1 (0.3) |
| Sum       | 58        | 77 (60.93)       | 87       | 44    | 52 (33.68) |

| Race/ethnicity | 2010 | 2011 | 2010 | 2011 | 2010 | 2011 | 2010 | 2011 | 2010 | 2011 |
|---------------|------|------|------|------|------|------|------|------|------|------|
| Caucasian     | 0    |      | 0    |      | 1    |      | 0    |      | 0    |      |
| African       | 1    | (0.4)| 3    | (0.6)| 2    | (0.6)| 0    | (0.1)| 1    | (2.12)|
| American      | 10   |      | 17   |      | 11   |      | 17   |      | 28   |      |
| Hispanic      | 0    |      | 0    |      | 0    |      | 0    |      | 18   |      |
| Asian         | 0    |      | 0    |      | 0    |      | 0    |      | 18   |      |
| Sum           | 58   |      | 77   |      | 87   |      | 114  |      | 44   |      | 52   | 33.68 | 192  | 2.237 |
We note that our treatment of the incomplete cases implicitly assumes that the marks are missing at random. That is, we assume that the missing data mechanism only depends on the observed marks. This is likely an unrealistic assumption. For example, we imagine that race/ethnicity was often not recorded for callers whose race/ethnicity was not obvious, was of a race/ethnicity not familiar to the paramedics, or was not recorded by request of the patient. This suggests that the marks are not missing at random. To account for this, additional data would need to be gathered to develop an appropriate missing data mechanism and either (i) include the missing data mechanism in the likelihood or (ii) impute missing marks at each iteration of the MCMC chain using the missing data mechanism.

In conclusion, the value of this work is realized when considered in the context of climate change and urbanization. That is, projections of increases in the number of warm days and nights (Oleson et al. 2013) and urbanization (Alig, Kline, and Lichtenstein 2004) will affect the volume of heat-related calls. For this reason, public health tools such as the one developed herein will be important to help mitigate the negative consequences of such changes. Using this tool to assess impacts of climate change and urbanization presents interesting statistical challenges—particularly the challenge of extrapolation. For example, under climate change temperatures greater than 38°C (100°F) could become more common. Whereas, in this dataset, such high temperatures were rarely observed rendering our estimated intensity surface to be low for this temperature region. Hence, to use this tool for predictive purposes under a climate change scenario would require additional constraints on the intensity surface to account for potential shifts in temperatures.

**SUPPLEMENTARY MATERIALS**

**Technical Details:** Technical details regarding the exploratory analysis of the 911 call data to justify modeling assumptions, the kernel convolution construction, and alternative models for model comparison.

[Received August 2013. Revised September 2014.]

**REFERENCES**

Alig, R. J., Kline, J. D., and Lichtenstein, M. (2004), “Urbanization on the us Landscape: Looking Ahead in the 21st Century,” *Landscape and Urban Planning,* 69, 219–324. [134]

Anderson, B. G., and Bell, M. L. (2009), “Weather-Related Mortality: How Heat, Cold, and Heat Waves Affect mortality in the United States,” *Epidemiology,* 20, 205–213. [133]

Baddeley, A. (2010), “Multivariate and Marked Point Processes,” in *Handbook of Spatial Statistics,* eds. A. E. Gelfand, P. J. Diggle, M. Fuentes, and P. Guttorp, Boca Raton, FL: Chapman & Hall/CRC Press, pp. 371–402. [124]

Bassil, K. L., Cole, D. C., Moineddin, R., Lou, W., Craig, A. M., Schwartz, B., and Rea, E. (2011), “The Relationship Between Temperature and Ambulance Response Calls for Heat-Related Illnesses in Toronto, Ontario, 2005,” *Journal of Epidemiology and Community Health,* 65, 829–831. [124]

Besag, J. (1974), “Spatial Interaction and the Statistical Analysis of Lattice Systems” (with discussion), *Journal of the Royal Statistical Society,* Series B, 36, 192–236. [124]

Best, N. G., Ickstadt, K., and Wolpert, R. L. (2000), “Spatial Poisson Regression for Health and Exposure Data Measured at Disparate Resolutions,” *Journal of the American Statistical Association,* 95, 1076–1088. [125]

Brix, A., and Diggle, P. J. (2001), “Spatio-Temporal Prediction for Log-Gaussian Cox Processes,” *Journal of the Royal Statistical Society,* Series B, 63, 823–841. [126]

Brookmeyer, R., and Stroup, D. (eds.), (2003), *Monitoring the Health of Populations: Statistical Methods for Public Health Surveillance,* New York: Oxford University Press. [124]

Chen, F., Manning, K., LeMone, S., Alfieri, J., Roberts, R., Tewari, M., Niyogi, D., and Horst, T. (2007), “Description and Evaluation of the Characteristics of the NCAR High-Resolution Land Data Assimilation System.” *Journal of Applied Meteorology,* 46, 694–713. [123]

Cosgrove, B., Lohmann, D., Mitchell, K. E., Houer, P. S., Wood, E. F., Schaeke, J. C., Robock, A., Marshall, C., Sheffield, J., Duan, O., Luo, L., Higgins, R. W., Fischer, R. T., Tarpley, J. D., and Meng, J. (2003), “Real-Time and Retrospective Forcing in the North American Land Data Assimilation System (nldas) Project,” *Journal of Geophysical Research,* 108, 8842. [124]

Daniels, M. J., and Kass, R. E. (1999), “Nonconjugate Bayesian Estimation of Covariance Matrices and Its Use in Hierarchical Models,” *Journal of the American Statistical Association,* 94, 1254–1263. [126]

Daniels, M. J., and Kass, R. E. (2001), “Shrinkage Estimators for Covariance Matrices,” *Biometrics,* 57, 1773–1184. [126]

Diggle, P. J. (2003), *Statistical Analysis of Spatial Point Patterns,* New York: Arnold. [124,126]

Dolney, T. J., and Sheridan, S. C. (2006), “The Relationship Between Extreme Heat and Ambulance Response Calls for the City of Toronto, Ontario, Canada,” *Environmental Research,* 101, 94–103. [124]

Flegal, J. M., and Hughes, J. (2012), mcmcse: Monte Carlo Standard Errors for MCMC, Riverside, CA and Minneapolis, MN, R package version 1.0-1. [127]

Gelman, A., and Rubin, D. (1992), “Inference From Iterative Simulation Using Multiple Sequences,” *Statistical Science,* 7, 457–511. [127]

Golden, J. S., Hartz, D. B., Brazael, A., Luber, G., and Phelan, P. (2008), “A Biometeorology Study of Climate and Heat-Related Morbidity in Phoenix From 2001 to 2006,” *International Journal of Biometeorology,* 52, 471–480. [124]

Gosling, S. N., Lowe, J. A., McGregor, G. R., Pelling, M., and Malamud, B. D. (2009), “Associations Between Elevated Atmospheric Temperature and Human Mortality: A Critical Review of the Literature,” *Climatic Change,* 92, 299–341. [133]

Hajat, S., Armstrong, B. G., Gouveia, N., and Wilkinson, P. (2005), “Mortality Displacement of Heat-Related Deaths: A Comparison of Delhi, Sao Paulo, and London,” *Epidemiology,* 16, 613–620. [133]

Hajat, S., and Kosatky, T. (2010), “Heat-Related Mortality: A Review and Exploration of Heterogeneity,” *Journal of Epidemiology and Community Health,* 64, 753–760. [124]

Heaton, M. J., and Peng, R. D. (2012), “Flexible Distributed Lag Models Using Random Functions With Application to Estimating Mortality Displacement From Heat-Related Deaths,” *Journal of Agricultural, Biological and Environmental Statistics,* 17, 313–331. [133]

Heaton, M. J., Sain, S. R., Greasby, T. A., Uejio, C. K., Hayden, M. H., Monaghan, A., Boehnert, J., Sampson, K., Banerjee, D., Nepal, V., and Wilhelmi, O. V. (2014), “Identifying Vulnerability to Heat-Related Mortality Using a Spatially Varying Coefficient Model,” *Spatial and Spatio-Temporal Epidemiology,* 8, 23–33. [123,133]

Higdon, D. (2002), “Space and Space-Time Modeling using Process Convolutions, in Quantitative Methods for Current Environmental Issues,” eds. C. Anderson, V. Barnett, P. C. Chatwin, and A. H. El-Shaarawi, London, UK: Springer-Verlag, pp. 37–56. [124,126]

Hu, L., Brunsell, N. A., Monaghan, A. J., Barlage, M., and Wilhelmi, O. V. (2014), “How Can We Use Modis Land Surface Temperature to Validate Long-Term Urban Model Simulations?” *Journal of Geophysical Research,* 119, 3185–3201. [124]

Isham, V. (2010), “Spatial Point Process Models,” in *Handbook of Spatial Statistics,* eds. A. E. Gelfand, P. J. Diggle, M. Fuentes, and P. Guttorp, Boca Raton, FL: Chapman & Hall/CRC Press, pp. 283–298. [126]

Kovats, R. S., and Hajat, S. (2008), “Heat Stress and Public Health: A Critical Review,” *Annual Review of Public Health,* 29, 41–55. [124,133]

Kusaka, H., Kondo, H., Kikegawa, Y., and Kimura, F. (2001), “A Simple Single Layer Urban Canopy Model for Atmospheric Models: Comparison with Multi-Layer and Slab Models,” *Boundary-Layer Meteorology,* 101, 329–358. [124]

Lawson, A. B., and Kleinman, K. (2005), *Spatial and Syndromic Surveillance for Public Health,* New York: Wiley. [124]

Li, B., Sain, S., Mears, L. O., Anderson, H. A., Kovats, S., Ebi, K. L., Bekkedahl, M., Kanarek, M. S., and Patz, J. A. (2012), “The Impact of Extreme Heat on Morbidity in Milwaukee, Wisconsin,” *Climatic Change,* 110, 959–976. [123]

Liang, S., Carlin, B. P., and Gelfand, A. E. (2009), “Analysis of Minnesota Colon and Rectum Cancer Point Patterns With Spatial and Nonspatial Covariate Information,” Annals of Applied Statistics, 3, 943–962. [127]

Little, R. J. A., and Rubin, D. B. (2002), *Statistical Analysis With Missing Data,* Hoboken, New Jersey: Wiley. [127]
Møller, J., and Waagepeterson, R. P. (2004), *Statistical Inference and Simulation for Spatial Point Processes*, Boca Raton, FL: Chapman and Hall/CRC Press.

Monaghan, A. J., Hu, L., Brunsell, N. A., Barlage, M. P., and Wilhelmi, O. V. (2014), “Evaluating the Impact of Urban Morphology Configurations on the Accuracy of Urban Canopy Model Temperature Simulations With Modis,” *Journal of Geophysical and Atmospheric Research: Atmospheres*, 119, 6376–6392. [124]

Oleson, K., Monaghan, A., Wilhelmi, O., Barlage, M., Brunsell, N., Feddema, J., Hu, L., and Steinhoff, D. (2013), “Interactions Between Urbanization, Heat Stress, and Climate Change,” *Climatic Change*, doi:10.1007/s10584-013-0936-8. [134]

Parrish, D. D., Allen, D. T., Bates, T. S., Estes, M., Feltsenfeld, F. C., Feingold, G., Ferrare, R., Hardesty, R. M., Meagher, J. F., Nielsen-Gammon, J. W., Pierce, R. B., Ryerson, T. B., Seinfeld, J. H., and Williams, E. J. (2009), “Overview of the Second Texas Air Quality Study (TexAQS II) and the Gulf of Mexico Atmospheric Composition and Climate Study,” *Journal of Geophysical Research*, 114, D00F13. [124]

Peng, R. D., Bobb, J. F., Tebaldi, C., McDaniel, L., Bell, M. L., and Dominici, F. (2011), “Toward a Quantitative Estimate of Future Heat Wave Mortality Under Global Climate Change,” *Environmental Health Perspectives*, 119, 701–706. [123]

Rainham, D., and Smoyer-Tomic, K. E. (2003), “The Role of Air Pollution in the Relationship Between a Heat Stress Index and Human Mortality in Toronto,” *Environmental Research*, 93, 9–19. [129]

Rusticiucci, M., and Bettolli, M. L. (2002), “Association Between Weather Conditions and the Number of Patients at the Emergency Room in an Argentine Hospital,” *International Journal of Biometeorology*, 46, 42–51. [123]

Smoyer-Tomic, K. E., and Rainham, D. G. C. (2001), “Beating the Heat: Development and Evaluation of a Canadian Hot Weather Health-Response Plan,” *Environmental Health Perspectives*, 109, 1241–1248. [129]

Tebaldi, C., and Knutti, R. (2007), “The Use of the Multi-Model Ensemble in Probabilistic Climate Projection,” *Philosopical Transactions of the Royal Society*, Series A, 365, 2053–2075. [133]

Uejio, C. K., Wilhelmi, O. V., Golden, J. S., Mills, D. M., Gulino, S. P., and Samenow, J. P. (2011), “Intra-Urban Societal Vulnerability To Extreme Heat: The Role of Heat Exposure and the Built Environment, Socioeconomics, and Neighborhood Stability,” *Health and Place*, 17, 498–507. [124]

Waller, L. A., Carlin, B. P., Xia, H., and Gelfand, A. E. (1997), “Hierarchical Spatio-Temporal Mapping of Disease Rates,” *Journal of the American Statistical Association*, 92, 607–617. [124]

Wilhelmi, O., de Sherbinin, A., and Hayden, M. (2012), “Exposure to Heat Stress in Urban Environments: Current Status and Future Prospects in a Changing Climate,” in *Ecologies and Politics of Health*, eds. B. King, and K. Crews, Routledge Press, pp. 219–238. [124]

Zhang, H. (2004), “Inconsistent Estimation and Asymptotically Equal Interpolations in Model-Based Geostatistics,” *Journal of the American Statistical Association*, 99, 250–261. [126]

Zhang, X., Boscardin, W. J., and Belin, T. R. (2006), “Sampling Correlation Matrices in Bayesian Models With Correlated Latent Variables,” *Journal of Computational and Graphical Statistics*, 15, 880–896. [126]