

Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

Gregory J. Schwartz

Date

The Association of Ozone and Particulate Matter Exposure to Asthma Related
Hospital Visits in Mississippi

By

Gregory J. Schwartz

Masters of Science

Biostatistics

Lance A. Waller, Ph.D.
Advisor

Robert H. Lyles, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the Graduate School

_____ Date

The Association of Ozone and Particulate Matter Exposure to Asthma Related
Hospital Visits in Mississippi

By

Gregory J. Schwartz
B.A., University of Colorado, 1997

Advisor: Lance A. Waller, Ph.D.

An abstract of
A thesis submitted to the Faculty of the Graduate School of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science
in Biostatistics
2008

Abstract

The Association of Ozone and Particulate Matter Exposure to Asthma Related Hospital Visits in Mississippi

By Gregory J. Schwartz

In 2006, 22.9 million Americans were estimated to have asthma. This thesis looks at the relationship between asthma related hospital visits and exposure to ambient concentrations of ozone and particulate matter $< 2.5 \mu\text{m}$ ($\text{PM}_{2.5}$). Exposure is measured remotely via satellite allowing the study to cover entire state of Mississippi.

A regular network of 1,318 grid cells was overlain the state of Mississippi. Mean daily Ozone and $\text{PM}_{2.5}$ exposure, number of asthma related hospital visits, and demographic characteristics were determined for each cell from January 1st, 2003 to December 31st, 2005. Dual-exposure models were built using Generalized Estimating Equations (GEE) to determine the association between hospital visits and exposures controlling for demographic characteristics. The study stratified by urban or rural designation considering the entire population as well as the black portion of the population. These models were extended with hierarchical Bayesian models to account for conditional autoregressive (CAR) spatial and non-spatial exchangeable random effects at the grid cell level.

Significant dual-exposure models were found for both urban and rural regions when considering the total population. A 2-day lag for ozone and 5-day lag for $\text{PM}_{2.5}$ were used in the urban only model with relative risks of 1.003 (95% CI = (1.001, 1.006)) and 1.004 (95% CI = (1.000, 1.008)) respectively. The rural only model used a 2-day lag for ozone and 4-day lag for $\text{PM}_{2.5}$ with respective relative risks of 1.002 (95% CI = (1.001, 1.005)) and 1.002 (95% CI = (1.000, 1.004)). A significant dual-exposure model was found for the black only urban area using a 4-day lag for ozone and a 2-day lag for $\text{PM}_{2.5}$ with relative risks of 1.004 (95% CI = (1.001, 1.006)) and 1.003 (95% CI = (1.001, 1.006)) respectively. The Bayesian analysis found exchangeable random effects in the urban region and spatial random effects in the rural region improved the model fit allowing for risk estimates to be made at the grid cell level.

The Association of Ozone and Particulate Matter Exposure to Asthma Related
Hospital Visits in Mississippi

By

Gregory J. Schwartz
B.A., University of Colorado, 1997

Advisor: Lance A. Waller, Ph.D.

A thesis submitted to the Faculty of the Graduate School of Emory University
in partial fulfillment of the requirements for the degree of
Master of Science
in Biostatistics
2008

Acknowledgments

I wish to express my sincerest gratitude to my advisor, Dr. Lance Waller, for his guidance and support in the preparation of this thesis. I felt privileged to have been able to work with Dr. Waller who not only took interest in my education, but devoted many hours of his time providing invaluable direction and advice. I would also like to thank the Applied Science Division of the National Aeronautics and Space Administration for providing the satellite data and financial support for the project that made this thesis possible.

Table of Contents

Introduction.....	1
<i>Background</i>	1
<i>Remote sensing of ground based pollution concentrations</i>	4
<i>Models for repeated data</i>	4
Methods.....	6
<i>Data Source</i>	6
<i>Missing Data</i>	7
<i>Analytical Methods</i>	7
<i>Descriptive Statistics</i>	8
<i>Longitudinal Models using Generalized Estimating Equations</i>	9
<i>Hierarchical Bayesian Models</i>	11
Results.....	16
<i>Descriptive Statistics</i>	16
<i>GEE Models</i>	18
<i>Single Exposure Models</i>	18
<i>Dual-Exposure Models</i>	19
<i>Hierarchical Bayesian Models</i>	20
Discussion.....	23
<i>GEE Model Discussion</i>	24
<i>Hierarchical Bayesian Model Discussion</i>	27
<i>Strengths and Limitations</i>	28
<i>Conclusion</i>	32
Figures 1 - 8	33
Tables 1 - 10.....	42
Appendix.....	55

Introduction

Background

Asthma is a serious lung disease characterized by chronic airway inflammation. Individuals with the disease experience episodes where airway restriction results in difficulty in breathing. The severity of asthma attacks can range from mild to severe with hospitalization and sometimes death resulting from the most severe cases. In 2006, an estimated 22.9 million (77.7 / 1,000 population) Americans had asthma with significant differences existing in gender, sex and age specific prevalence. With 12.8 million (85.7 / 1,000 population) female and 10.0 million (69.7 / 1,000 population) male asthmatics, prevalence in females was 23% higher than in males. Prevalence in the black population was 23.9% greater than that in the white population with 3.7 million (94.2 / 1,000 population) of the black population being asthmatic and 18.1 million (76.1 / 1,000 population) of the whites population being asthmatic. Children less than 18 years old had a 27.3% higher prevalence than the adult population. 6.8 million (92.8 / 1,000 population) children had asthma compared to 16.1 million (70.9 / 1,000 population) adults (American Lung Association 2007).

It is likely that air pollutants in the environment act to increase both the prevalence and morbidity of asthma and exacerbate asthma related symptoms. The Clean Air Act of 1970 required that the United States Environmental Protection Agency set national standards for the allowable ground level ambient concentrations of six common air pollutants referred to as the criteria pollutants. The pollutants include particulate matter, ground-level ozone, carbon monoxide, sulfur oxides, nitrogen oxides, and lead.

Each of these pollutants is thought to cause harm to human health or the environment (Environmental Protection Agency 2008).

A large body of clinical and epidemiological studies supports an association between ambient air pollution and increased asthmatic events or decreases in other measures of lung functions. The focus of this work is on the effect of increased ozone and particulate matter (PM) concentrations on asthma related hospital visitation. Ozone and particulate matter affect the lungs by causing irritation and inflammation which may result in decreased lung function and the exacerbation of asthmatic symptoms (Balmes 1993; Delfino et al. 2002; Koren 1995; Lewis et al. 2005; van Eeden et al. 2001).

Studies have shown that ozone and PM provide the strongest evidence of an association with exacerbation of asthmatic symptoms (Schildcrout et al. 2006). Lewis et al. (2005) looked at single and dual-exposure models to determine the association between PM less than 2.5 μm ($\text{PM}_{2.5}$) and ozone exposures and lung function as measured by forced expiratory volume. The study included a population of asthmatic children and stratified based on the presence of upper respiratory infection and use of maintenance corticosteroids. Exposure lags of 1 and 2 days along with an averaged lagged exposure from 3 to 5 days were considered. $\text{PM}_{2.5}$ with the 3 to 5 day average lag and ozone with a 2 day lag were found to be associated with decreased lung function in single exposure models. Dual-exposure models did not consider different combinations of lag times. In general, models with $\text{PM}_{2.5}$ and ozone tended to show significant effects in the longer lag times considered. Delfino et al. (2002) compared the effect of PM_{10} on asthma symptoms of asthmatic children by looking at 1-hour and 8-hour peak measurements compared to the 24-hour mean concentrations using a 0-day lag time while

stratifying by the use of anti-inflammatory medication. The 24-hour mean is the value used for regulatory purposes as well as most epidemiological studies. It was found that models using 1-hour maximum concentration had the strongest association with adverse asthma events. The study also found the strongest 24-hour mean lag time association at a 0-day lag and a 3-day moving average. A study of asthma-related hospital admissions on a nonelderly population (younger than 65 years old) was done by Sheppard et al. (1999). Here $PM_{2.5}$ and ozone were found to be associated with increased rate of asthma related hospital visits. Lag times used for $PM_{2.5}$ and ozone were 1-day and 3-day respectively. $PM_{2.5}$ and O_3 were not found to be jointly significant together in the same model, but $PM_{2.5}$ was found to be jointly significant with carbon monoxide (CO). Sun et al. (2006) performed an analysis of asthma-related emergency department visits that considered ozone and $PM < 10 \mu m$ (PM_{10}) independently. An association between children younger than 16 years of age, PM_{10} and emergency department visits was found. A relationship was not found between children and ozone and emergency department visits, or between adults (as defined as older than 16 years) and asthma or ozone.

Most studies, including the studies just reviewed, looked at the effect of ambient pollution concentrations on individuals who are already asthmatic. For example, emergency department visits likely result from high levels of pollution that have severely exacerbated a preexisting asthma condition. Few prospective studies have been performed that look at the association of ambient air pollution and incidence of new asthma cases. Gilmour et al. (2006) reviewed five prospective studies that together lend only modest support that increased ambient levels of pollution including $PM_{2.5}$ and ozone may result in increased incidence of asthma in an otherwise healthy and non-asthmatic

population. In general, any prospective study of a chronic disease onset is difficult because of the difficulty associated with clear identification of disease onset as well as the identification of appropriate exposures and controls.

Remote sensing of ground based pollution concentrations

This thesis explores the association between asthma-related hospital visits and ambient concentrations of ozone and PM_{2.5}. Previous research typically relied on ground-based monitoring systems to estimate exposure with relatively large scale study regions that were generally at the city or metropolitan level. An approach to modeling the association of exposure to ambient air pollution and asthma by utilizing remotely sensed ozone and PM_{2.5} concentrations from satellites to estimate exposure, is developed for a small geographical scale study that in this paper covers the entire state of Mississippi. This study design differs markedly from many previous asthma studies by considering a regular grid of cells overlain over Mississippi as the “subjects”, or basic units of observation, whereas previous works consider the individual. All data about the individual is aggregated in this grid-based design and the outcome is the number of individual asthma-related hospitalizations on any one day with respect to a particular grid cell. This study is also unique in its exhaustive search for lag time combinations of ozone and PM_{2.5} in dual exposure models so that it could be determined if their effects are differentiable.

Models for repeated data

The study area was decomposed into the regular grid network as seen in Figure 1. Because the outcome is count data, a Poisson process is assumed to describe the relationship between grid cell, pollution exposure and outcome. The properties of the Poisson model force a major assumption that the mean and variance of modeling residuals are equivalent. A traditional quasi-likelihood approach to longitudinal modeling using generalized estimating equations (GEE) is first implemented under this assumption, even though overdispersion in variance likely exists. The GEE approach offers a means to model overdispersion by determining an overdispersion factor greater than unity that can be used to inflate the standard error of parameter estimates.

It also is desirable however to utilize the spatial structure of the study design to model overdispersion. A second set of hierarchical Bayesian models is therefore proposed that directly considers spatial residuals. This is a more robust approach to modeling overdispersion as spatial and non-spatial random effects are parameterized to directly account for extra variability in an easy and straightforward (albeit computationally demanding) way. Random effects can also be included in various levels of the hierarchy as required by the study design adding modeling flexibility. Inclusion of random effects in this way supports a cluster- or subject-specific interpretation of model parameters, a shift from the population averaged or marginal interpretation obtained from GEE fixed effect estimation. The generalized linear mixed model (GLMM) approach based on pseudo-likelihood estimation (Breslow and Clayton 1993) also supports mixed effects modeling. Because the GLMM class of models is not well suited for the specification of spatial random effects this class of models was not considered in this work.

Methods

Data Source

A regular grid of cells each approximately 8.3 km by 10 km was superimposed over the state of Mississippi (Figure 1). Each cell contained spatially aggregated data measuring the number of asthma related hospital visits, concentration of ambient ozone (O_3 ppb), 2.5 micron ambient particulate mater ($PM_{2.5}$ $\mu g/m^3$), and demographic characteristics. Each cell had values recorded daily over the three year study period from January 1st, 2003 to December 31st, 2005 resulting in 1,096 observations per cell. The 1,318 grid cells that fell completely within the state boundary were initially considered. Hospital visitation records were georeferenced at the residential level and counts were then aggregated to the grid cell level for each day of the study period.

Raw O_3 and $PM_{2.5}$ concentrations were collected from a satellite producing measurements in a grid based format. The satellite grid was registered to the study grid and overlain. Daily pollution values for each grid cell were derived from a weighted average that considered the proportion of satellite grid cells each study grid cell intersected. Multiple pollution concentration values were recorded throughout the day, so average daily values were also calculated. The result is an O_3 and $PM_{2.5}$ value aggregated for each day and for each grid cell over the study period. Demographic

characteristics were assigned to the study grid based on United States 2000 census data with projections applied for the beginning of each of the three study years¹.

Missing Data

In order to minimize potential bias in defining the strongest temporal lags, only counties in Mississippi known to have complete data with respect to asthma related hospital visitations were included in this study. Of the 82 counties within Mississippi, 9 counties were identified to have incomplete data. Grid cells inside, or having more than 50% of their areas within the identified county boundaries were excluded from the study. In this way, the number of grid cells used in this study was reduced from 1,318 to 1,045 with a total of 1,145,320 data points representing a 20% reduction in data.

The number of missing values generated also depended on the choice of the maximum lag time for O₃ and PM_{2.5} and was small relative to the number of observations for each grid cell. For example, if a model used a 2-day and 5-day lag time for O₃ and PM_{2.5} exposures, respectively, then each grid cell would have 5 missing values at the beginning of the first study year. Missing data also exists for PM_{2.5} on October 10th, 2004 for all grid cells due to satellite measurement errors on that day. The actual day when these data are recorded as missing is dependent on the PM_{2.5} lag time used in the model.

Analytical Methods

¹ This is a general description of the data aggregation process. Data used in this thesis was received in a form aggregated to the grid cell level with no identifying information at the individual level.

Descriptive Statistics

Descriptive statistics were calculated at the grid cell level to summarize asthma-related hospital visitation, O₃ and PM_{2.5} levels, and demographic characteristics. Basic demographic characteristics included the percentage of black and female populations, with respect to grid cell population, to account for race and gender respectively. Average per capita income (PCI) and the average median household income were also considered. Because census projections were applied at the beginning of each study year, summaries were calculated for each calendar year of the study. Two regions were considered for modeling that divided the study grid into spatially distinct rural and urban sub-regions. The union of these two regions composed the entire study region and defined a third modeling region. Demographic characteristics were also calculated for each of these three regions.

Seasonal fluctuations of Ground level concentrations of O₃ and PM_{2.5} were examined to determine a suitable mechanism for controlling for seasonality. Daily values of both pollutant levels and hospital visits were plotted and their variations smoothed using loess curves (Cleveland et al. 1992) to better visualize trends throughout the 3 year study period. To control for the seasonal variability in asthma-related hospital visitation, a month variable was introduced for every observation.

In order to investigate potential associations between time-lagged exposure to pollutants and asthma-related events, a comprehensive evaluation of lag times was conducted for each pollutant ranging from 1 to 14 days. Covariance matrices were also generated to assess the degree of collinearity between lagged pollution measurements of O₃ and PM_{2.5} both jointly and independently.

Longitudinal Models using Generalized Estimating Equations

Poisson regression modeling of the effects of O₃ and PM_{2.5} on the daily counts of asthma-related hospital visits was done using generalized estimating equations (GEE) (Liang and Zeger 1986). Spatial correlation was not accounted for in the GEE models due to difficulty in specifying this type of correlation within a working correlation matrix. To standardize the number of hospital visits by the size of the at-risk population within each grid cell, Poisson regression was performed on the rate of visits (total visits / total population) where the total population size is used as an offset in the model (McCullagh and Nelder 1989). Within this study each grid cell is considered to be a cluster containing exposure and outcome measures repeated daily throughout the study period. An autoregressive (AR1) working correlation matrix was used to account for the temporal correlation between observations within each grid cell, assuming higher correlations in measurements closest in time. The model takes the form given in Equation 1.

Equation 1

$$y_{g,t} \sim \text{Poisson}(\mu_{g,t})$$

$$\log(\mu_{g,t}) = \alpha + \mathbf{X}_{g,t}^T \boldsymbol{\beta} + \log(n_{g,t})$$

$g \in (1, 2, \dots, G)$, where G is the number of Grid Cells

$t \in (1, 2, \dots, T)$, where T are the dates of each sample

Here $y_{g,t}$ and $\mu_{g,t}$ are the observed and expected number of hospital visits in grid cell g at time t respectively, and $n_{g,t}$ is the population at risk in grid cell g on date t . \mathbf{X} includes the exposure and demographic variables with values recorded for each grid on each date.

Two general model types were considered in this analysis. The first type was based on the proportion of all aggregate visits relative to the entire population in each grid cell. The second type represented race-specific models based on the proportion of aggregate visits from the black population relative to the total black population in each grid cell. These two types of models differ only in the aggregate visits outcome and the total population offset term from Equation 1. The models were applied to each of the three geographic regions (urban, rural, and combined urban and rural) previously defined, resulting in a total of 6 different models.

Analysis of Lag Times

Since the data involve grid-based remotely sensed exposure values and aggregated case counts, it was not immediately clear whether previous exposure lags for either O_3 or $PM_{2.5}$ would apply to our data. As a result, a comprehensive regression analysis was performed to investigate which lag times yielded the strongest associations between grid-based outcomes and each exposure, O_3 and $PM_{2.5}$. Single-exposure models were built using all covariates, without consideration of more parsimonious models, to generate a set of positive and significant parameter estimates for both O_3 and $PM_{2.5}$ independent of each other. These two sets provided a reduced search space for possible O_3 and $PM_{2.5}$ combinations. In this way a comprehensive set of dual-exposure models

was generated. Finally, only dual-exposure models that resulted in positive and significant parameter for O_3 and $PM_{2.5}$ having lag times within 5 days of each other were considered candidates for a final model. This process was repeated for each of the 6 model types considered resulting in a single dual exposure model for each model type. The final selection was based on lagged exposure combinations having the greatest positive and significant parameter estimates along with the lowest associated standard errors. A final model selection was then performed for each model that removed from the initial set of control variables any covariates that were found to be insignificant and not part of a confounding relationship.

Hierarchical Bayesian Models

Hierarchical Bayesian models were used to extend the results of the GEE models and use random effects to model extra variance and more complex correlation structures. Parameterizing additional variation within the model this way extends the usual Poisson assumption of equality of variance and mean to allow residual extra-Poisson variation or (in our case) spatially structured correlation between neighboring grid cells. WinBUGS (Spiegelhalter et al. 2003) was used to fit all Bayesian models considered in this study.

The spatial-temporal structure of the data presents many levels where random effects may be studied. In the temporal dimension random effects can be used to test for correlation at the year, month and possibly the day level. Random effects can be inserted at the grid level to consider both spatial and non-spatial correlation. Equation 2 is the general mixed effects model used in this study.

Equation 2

$$\log(y_{g,y,m,d}) = \alpha + \mathbf{X}_{g,y,m,d}^T \boldsymbol{\beta} + u_g + \gamma_g + \log(n_{g,y,m,d})$$

$g \in (1, 2, \dots, G)$, where G is the number of grid cells
 $y \in (1, 2, \dots, Y)$, where Y is the number of years
 $m \in (1, 2, \dots, M)$, where M is the number of months
 $d \in (1, 2, \dots, D)$, where D is the number of days

Here the linear predictor is a linear combination of fixed effects, 2 random effects and the offset term. The fixed effects include the lagged ozone and PM exposure variables as well as other demographical covariates as determined by the GEE models. The random effects model a *convolution* structure as it includes both spatial and non-spatial random effects, u_g and γ_g respectively. Lower level random effects were not included in this model primarily due to computational constraints.

The Bayesian hierarchical model specifies the posterior densities of parameters as proportional to the product of the likelihood, prior and hyper-prior distributions. The Bayesian hierarchical form of the joint posterior of the fixed and random effect parameters is:

Equation 3

$$f(\beta_p, u_G, \gamma_G, \sigma_u, \sigma_\gamma | y) \propto f(y | \beta_p, u_G, \gamma_G) \cdot f(u_G | \sigma_u) \cdot f(\gamma_G | \sigma_\gamma) \cdot f(\beta_p) \cdot f(\sigma_u) \cdot f(\sigma_\gamma)$$

In Equation 3, the posterior distribution contains P fixed effects parameters as determined by the models obtained by GEE. In addition to the fixed effects, there are two sets of G random effects for each grid cell, one defining extra-Poisson variation and the other

inducing spatial correlation. Non-informative prior distributions are placed on each of the fixed effects where for each $\beta_p \in \beta_p$, $\beta \sim N(0, \sigma_\beta^2)$ and $\sigma_\beta^2 = 10,000$ for each fixed effect. Spatially uncorrelated random effects are given the following exchangeable prior distribution:

$$\begin{aligned} \gamma_g &\overset{ind}{\sim} N(0, \sigma_\gamma^2), \quad g \in (1, \dots, G) \\ \sigma_\gamma^{-2} &\sim \text{gamma}(\alpha, \beta) \end{aligned}$$

This prior and hyper-prior distribution generates a common normal distribution with mean 0 and shared prior variance σ_γ^2 for each nonspatial random effect. With this structure, each random effect, γ_g , is independent of the others and not dependent on location g (Waller and Gotway 2004). Spatial correlated random effects are modeled by the Conditional Auto Regressive (CAR) distribution (Besag et al. 1991):

$$u_g | u_{\{-g\}} \sim N \left(\bar{u}_{\{-g\}}, \frac{\sigma_u^2}{\sum_{\{-g\}} w_{g,g'}} \right)$$

Where $\{-g\}$ is the set of grid cell neighbors of g exclusive of g ,

$$\bar{u}_{\{-g\}} = \frac{1}{\sum_{\{-g\}} w_{g,g'}} \sum_{\{-g\}} u_g w_{g,g'},$$

$$\text{and } w_{g,g'} = \begin{cases} 1 & \text{if } g \text{ and } g' \text{ are adjacent } (g \neq g'), \\ 0 & \text{otherwise} \end{cases}$$

$$\sigma_u^{-2} \sim \text{gamma}(\alpha, \beta)$$

In this study, the neighborhood of a grid cell is defined by all grid cells that touch its boundary. A single grid cell can then have up to eight neighbors. Un-normalized,

binary weights ($w_{g,g'}$) indicating adjacency are assigned the value 1 for each neighboring grid cell. This results in the conditional distribution of each individual grid cell random effect being centered about the simple average of the random effects of its neighbors. The overall variance is then scaled by the sum of the weights (in this case the number of neighbors). In this specification the CAR model is improper in that the overall mean of the spatial random effect is not defined. The WinBUGS implementation of the CAR normal model handles this by constraining the random effects to sum to zero which requires that the intercept term α in Equation 2 be given an improper and unbounded uniform prior distribution ($\alpha \sim \text{uniform}(-\infty, \infty)$) (Besag and Kooperberg 1995). The same parameters for α and β in the gamma distribution for the exchangeable and CAR precision parameters are not required.

The relatively large dataset for this study made running the hierarchical models for each candidate GEE model computationally prohibitive. As a result, focus was limited to the full population models with the race-specific models not being considered. Limitations on computational resources also forced the urban and combined regional models to use subsets of the data that were used for the GEE models (Figure 2). Data subsets were obtained by selecting a contiguous set of grid cells that had similar parameter estimates when calculated with the same model as the full dataset using GEE. Each of the three models was initially based on Equation 2 and Equation 3 containing the convolution prior. The effect on each of these models while including either the spatially correlated or uncorrelated random effects independently was also considered.

Bayesian model diagnostics

Diagnostics were performed to determine if the MCMC simulations had converged on a stationary distribution, to determine model sensitivity to the specification of hyper-prior variance distributions on random effects and to determine the effect of spatially correlated and uncorrelated random effects on overall model fit. Model convergence was tested using the Gelman/Rubin R statistic (Gelman and Rubin 1992) to measure the within and between variance components of Markov chains. Values close to one are achieved when the two measures of variance approach each other, which occurs upon convergence. To support this statistic, two Markov chains were initially run for each simulation. The chains were extended as necessary for convergence. So that a meaningful R statistic could be calculated as the chains converged from different locations, initial values for the chains were calculated as either ± 1 or ± 2 standard errors from the parameter estimates as determined by the GEE models. Initially 2,000 iterations were performed for each Markov chain with incremental extensions of 1,000 iterations added as required for sufficient convergence (R close to one). Posterior densities were determined from the last 1,000 iterations of each Markov chain. The Deviance Information Criterion (DIC) (Spiegelhalter et al. 2002) was used to determine the best-fitting model with respect to convolution, CAR or exchangeable random effects for each of the three models considered. The sensitivity of the variance for each of the distributions of random effects was evaluated by using DIC to compare the model fit resulting from different hyper-prior specifications. This final analysis was carried out only on the model determined to have the best fit for each of the three geographic regions.

Results

Descriptive Statistics.

Descriptive statistics of the grid cells are shown in Table 1. Comparisons were made by geographical region designated as urban and rural, rural only or urban only. The urban & rural region is a superset of grid cells composed of the union of the non-intersecting urban and rural regions. There were 1,045 total grid cells of which 933 were considered rural and 112 urban. The descriptive statistics within each region are broken down further by the study years 2003 - 2005. In all regions it is seen that the total population and average population per grid cell is increasing. In the composite urban-rural region the percentage of black population grew from 37.86% to 38.44% and the percentage of females in the population adjusted only slightly from 50.69% to 50.58% during the years under study. This trend is followed closely by the rural region, a likely reflection of the higher proportion of rural cells in the composite urban-rural region. A slight decrease in the female population is observed in the urban region ranging from 51.31% to 51.25% over the 3 year period. The urban percentage of black population is slightly smaller than the rural and composite regions and it increases from 36.66% to 37.41% over the study period. Within the composite region there were 19,241 total hospital visits in 2003. This number decreases to 18,021 and then jumps to 22,487 in years 2004 and 2005 respectively. A similar trend exists in the rural and urban regions. The average number of hospital visits per grid cell for each study year is much larger in the urban areas as would be expected from increased population density. Here the average number of visits ranges from 67.76 per year to 90.57. This compares to 11.66 –

13.23 and 17.24 – 21.52 visits per year in the urban and combined regions respectively.

A noticeable disparity exists in the proportion of black individuals accounting for hospital visits. Within the composite urban and rural region, the percentage increases from 66.09% to 67.15% which coincides with a black population of approximately 40% throughout the study period. This disparity is more pronounced in the urban region where the percentage increases from 73.37% to 76.58% during the study period. A decreasing trend in the proportion of black hospital visits is seen within the rural region. Here the percentage decreases from 61.28% to 59.39%.

Figure 3 presents time series plots of the averaged daily values for grid cells throughout the duration of the study for O_3 and $PM_{2.5}$ concentrations as well as the averaged number of daily hospital visits. Hospital visits are shown for the whole study population over the combined region as well as the visitation from the black population over the urban and rural sub-regions. The O_3 and $PM_{2.5}$ time series are omitted for the separate urban and rural regions as their patterns are very similar to those seen for the entire region. For each plot, a smooth loess curve generalizes the day to day mean trend while the variance in measured values is visible in the background. General seasonal trends are apparent in each plot. Ozone concentrations have a major peak from March to May with a less defined peak around August and September. $PM_{2.5}$ concentrations tend to peak towards the end of the summer from August to October. A time series for total visits is shown for all regions along with the series for visits by black patients for all regions and urban regions. The seasonal cycle in the number of visits is similar throughout with peak visitations taking place around the month of October. This pattern is particularly evident in 2003 and 2005, but appears diminished in 2004 which is

consistent with the overall reduced number of hospital visits observed for this year.

Finally, a valley is observed about the month of July that appears consistent across study years.

A Pearson's correlation matrix for O₃ and PM_{2.5} lag times appears in Table 2. All correlation coefficients in the matrix were found to be significant at the 0.05 level. It is apparent that all lagged PM_{2.5} concentrations are positively correlated with those of O₃. Each matching lag time for the two pollutants shown has a correlation coefficient of about 0.38. As the time period between the lag times of the pollutants increases, the correlation coefficient decreases. Tables 3-4 show the Pearson's correlation matrixes for the lag times of O₃ and PM_{2.5} independent of each other. All correlation coefficients in both matrixes are significant at the 0.05 level. In both cases the strongest correlation (with the exception of equal lag times) is found between lag times one day apart ($r = 0.79$ for O₃ and $r = 0.74$ for PM_{2.5}). The rate of decrease in r as the time period between lags increases is greatest for PM_{2.5}.

GEE Models

Single Exposure Models.

The results of single exposure models by geographic region and by race are presented in Table 5. The models each controlled for per capita income and the percentages of each grid cell which are female or black. The seasonal variations in hospital visits were controlled for by way of a month of observation variable. The parameter estimates in this table are crude in the sense that a full model selection was not done for each model type. As a result, only the parameter estimates for O₃ and PM_{2.5} are

shown. Table 5 only lists the single largest positive and significant parameter estimate for both O₃ and PM_{2.5} for both the low and high range lag values for each model type. There can be at most 4 values for each model type. This selection comes from a set of 15 models for each pollutant looking at lag times of 0 – 14 days and provides a useful comparison for the dual-exposure model.

Dual-Exposure Models

Final dual-exposure model results are shown in Tables 6 – 9. Not all model types had a valid model because there were no combinations of O₃ and PM_{2.5} lag times that had both positive and significant parameter estimates. It is also apparent that sets of significant covariates varied between model types. All models shown use the same month variable to control for seasonal variations in hospital visitation. The parameters % black Q1 – Q4 and % female Q1 – Q4 are quartile indicator variables for the percentage of black and female populations respectively with Q1 being the reference level.

Table 6 shows results from the all racial models for the lower range of exposure lagged combinations. The composite urban-rural and rural regions both have O₃ lag 2 and PM_{2.5} lag 4 for lagged exposures. The parameter estimates for these exposures are identical and the models share the same set of covariates. However, they do differ in the non-exposure covariate parameter estimates. The urban only model in Table 6 also uses O₃ lag 2 and has a similar parameter estimate as the other two regional models. This model also differs in that gender is not a significant variable and there is a much stronger effect from race. Here grid cells in the highest quartile of the percentage of the black

population have a relative risk of 3.7 compared to 1.4 for the rural only region and 2.0 for the composite region.

The models based on the total population for the upper lagged combination range are shown in Table 7. A dual-exposure model with a significant O₃ and PM_{2.5} lag time combination was not found for the rural only type. The combined urban-rural model uses O₃ lag 12 and PM_{2.5} lag 10 parameter estimates along with the % black and % female covariates. Again, the urban only model drops the % female covariate and uses the O₃ lag 10 and PM_{2.5} lag 12 exposure combination.

The models limited to the black population are shown in Table 8 and Table 9. In the lower range only the urban model has a significant combination and other than controlling for month, there are no other significant covariates. The model uses O₃ lag 4 and PM_{2.5} lag 2 compared to O₃ lag 2 and PM_{2.5} lag 5 in the all race urban only model. In both cases the parameter estimates are similar. In the upper range only the composite urban-rural and urban only models were found to be valid models. Both models use the same covariate set. The composite model uses O₃ lag 10 and PM_{2.5} lag 13 compared to O₃ lag 12 and PM_{2.5} lag 12 used in the urban only model.

Hierarchical Bayesian Models

The parameter estimates from the hierarchical Bayesian analysis are given in Table 10. The urban area estimates are based on the exchangeable prior and the rural and combined urban-rural areas are based on the CAR prior. These estimates along with the cell specific random effects were then used to determine the overall risk as well as the expected number of hospital visits for each cell. These values are shown in Figure 6-8.

The number and percentage of significant random effects at the 0.05 level were observed to be 61% or 447 of 727 for the combined urban-rural region, 71% or 80 of 112 for the urban only region, and 61% or 299 of 494 for the rural only region.

The initial run of each of the 3 models was based on the convolution prior for random effects. These models revealed a general inability of the models to distinguish between the spatial and non-spatial random effects. This is evident in Figure 4 where the Markov chains for the CAR and exchangeable random effect variances show a strong inverse correlation. In the case of the rural and urban-rural models the first 750 to 1000 iterations appear stationary for the CAR variance distribution with a value of about 2.7. At the same time the exchangeable variance is close to 0 and shows signs of decreasing stability after 500 iterations. A rapid transition then occurs in how the variance is modeled. This is evident as the CAR variance parameter destabilizes and rapidly declines towards 0 as the exchangeable prior rapidly increases towards 1 and appears to be in convergence itself during the last iterations of the simulations. This process is also observed for the urban simulation, but with a less abrupt change in how the variances are modeled.

The results of the convolution prior models required that either the CAR or exchangeable random effects be included independently in a final model. The urban only model showed better fit using exchangeable compared to CAR random effects as measured by a reduction in the DIC of 21.5. Both models for the rural and combined urban-rural regions were best fit using the CAR random effect compared to the exchangeable random effect. The reduction in the DIC was 269.00 and 206.00 for the combined and rural only models respectively.

shows how the choice of a CAR or exchangeable prior for the random effects influences the expected number of hospital visitations. The plots show the expected values for January 5th 2005 for each of the 3 geographical regions with the yearly mean concentrations of O₃ and PM_{2.5} used to calculate the expectations. If the prior distributions have no effect on expectation, then a one-to-one relationship would exist between the two expectations and the plotted values would fall on the line with slope 1. There is however a general tendency for the exchangeable random effects to be larger than the CAR random effects suggesting that they do not follow a one-to-one relationship. This difference becomes most obvious when the expectations are large relative to the main mass of points. For example, the point in the upper right quadrant of the urban-rural plot has an expectation of 10.91 and 7.81 from the exchangeable and CAR prior respectively. This difference highlights the utility of the CAR prior in adjusting parameter estimates towards the mean value of its neighbors and the capture of local deviations.

The sensitivity of each model to the specification of the hyper-prior inverse-gamma distribution $\left(\text{Gamma}(\alpha, \beta) \text{ where the mean} = \frac{\alpha}{\beta} \right)$ of the random effect precision was tested. Two specifications were compared: $(\alpha, \beta) = (0.001, 0.001)$ and $(\alpha, \beta) = (0.5, 0.005)$. The specification did not have any discernable difference on model fit as measured by the DIC. The model results in Table 10 are based on the $\alpha = .001, \beta = .001$ gamma distribution. The posterior distributions and traces of the Markov chains for the random effects of each model are found in the appendix. The WinBUGS code for the urban and urban-rural regional models is also listed in the appendix.

Discussion

The GEE and Bayesian models developed in this paper differ markedly from previous works both in study design and exposure assessment. Previous investigations have focused on the asthmatic population presenting studies looking at how ambient air pollution exacerbates symptoms in a specific population. By focusing on a population of individual asthmatics as the unit of observation, studies have been able to support analyses based on specific individual characteristics. For example the analysis done by Delfino et al. (2002) stratifies by use of anti-inflammatory medication, and Tolbert et al. (2002) predicts the probability of an emergency room patient being an asthma case. These studies have also been performed at the city scale and have used ground based pollution monitoring stations to determine exposure levels for individual subjects.

The analyses in this study contrast with previous works in both the unit of observation and the method of exposure assessment. The study presented in this work encompasses the entire state of Mississippi, and the regular network of grid cells overlain the state are the units of observation as opposed to the individual. Because ground-based measurements of exposure are not possible in such a small scaled study, satellite measurements were used. As discussed below, the satellites provide an estimate of exposure on the ground and these estimated exposure values should be considered as an exposure proxy. The models presented here make no assumptions about who is going to the hospital and the entire population within each grid cell is considered to be at risk. Risk assessments derived from the modeling results are therefore interpreted with respect to the entire population residing within each grid cell.

GEE Model Discussion

The urban-rural combined models appear to be more similar to the rural only models than compared to the urban only models. This is particularly evident for the urban only GEE models looking at all races in Table 6 where the rural and urban-rural models use the same set of parameters and exposure lag times. The urban only model did not show a significant effect for gender and uses a 5-day lag for PM_{2.5} instead of the 4-day lag used for the other two regions. The urban and urban-rural similarity is most likely a result of the relatively small number of grid cells classified in the urban region (112) compared to those in the rural region (933). The urban-rural model is therefore influenced mostly by the rural region and can be interpreted as a weighted average of the separate urban and rural regions. In this regard it is more useful to compare the urban and rural models directly. Both models have similar parameter estimates for the O₃ and PM_{2.5} exposures and the month of observation with the urban model giving more weight to the effect of PM_{2.5}. The largest difference in parameter estimates is found for grid cells being in the 4th quartile for percent black. In the urban model grid cells in this quartile have 3.69 times the risk of grid cells in the lower quartile. This compares to the rural only model where there is only 1.40 times the risk as compared to the lower quartile. This difference may be a result of the demographic characteristics of the urban and rural regions (

Table 1). Considering the 2003 study year in, the percentage of the population that is black in each grid cell is similar for both geographical regions. It is also apparent that the black population accounts for a higher proportion of hospital visits. This is consistent with the 2006 national asthma prevalence being 23.9% greater in the black population as compared to the white population. The urban model is picking up on these demographic trends by associating more risk on grid cells that have a greater proportion of black in their local population. The fact that gender appears significant in the rural and urban-rural combined models is also consistent with the 23% greater prevalence of asthma in females as compared to males. The models did consider the proportion of the population younger than 18 years old, but this was not found to be a significant variable even though this segment of the population has a 27.3% higher prevalence of asthma than the adult population.

It is believed that the black population may be overrepresented in the aggregate count of hospital visits for each grid cell. This of course has the potential to introduce bias into the model which could affect the risk adjustments estimated for race in particular. In order to limit this possible source of bias, race-specific models were examined that considered the rate of black only visits within the black population. Only the urban region had significant dual-exposure model when limiting consideration to black hospital visits in the lower lag range. This may be a result of the racial differences previously discussed between the rural and urban regions. In comparing the urban-only model for the black population in Table 8 to the urban only model for the entire population in Table 6 only a small difference is apparent. The lag combination of O₃ lag 4 and PM_{2.5} lag 2 is used in the former model compared to O₃ lag 2 and PM_{2.5} lag 5 in the

later. The magnitude of the exposure estimates is however similar for both models. The percent black variable was not a significant covariate in the black-only models including those for the upper lag range (Table 9). This is a likely result of the outcome of the models being the ratio of black hospital visits to the black population and therefore the use of the real size of the black population is no longer a significant factor within the model.

Reviewing the single parameter estimates in Table 5 reveals that the parameter estimates tend to be larger for single-exposure models as compared to dual-exposure models. For example, the two lower range parameter estimates for the all races urban and rural model are 1.004 and 1.005 for O₃ lag 3 and PM_{2.5} lag 5 respectively. This compares to 1.003 and 1.002 in the dual-exposure model for O₃ lag 2 and PM_{2.5} lag 4 respectively. The dampening of the respective effects of exposures is likely a result of correlation between O₃ and PM_{2.5} lagged exposure times (Table 2). The positive correlation between ozone and PM_{2.5} implies that the two pollutants likely represent competing risks for asthma related hospital visits. Previous research has also reported significant correlations between multiple exposures which in some cases resulted in insignificant dual-exposure models (Sheppard et al. 1999; Tolbert et al. 2000). The independent lag time correlation matrixes for O₃ and PM_{2.5} also suggest that because of very high correlations observed particularly between 1 day lag times (Table 3- Table 4) that the lagged exposure parameters presented in Table 6 - Table 9 may be considered as representing a range of effects over time, rather than as definitive lags of highest impact.

Interestingly, the actual lag times, ranging from 2 to 5 days, used for the models in Table 6 and Table 8 are consistent with those reported in the literature reviewed

previously. The interpretation of these lag times is meaningful in that a high exposure of O_3 or $PM_{2.5}$ on a given day has the potential to bring on severe asthmatic symptoms resulting in a hospitalization a few days later. It is more difficult to interpret the lag times for the models using the upper lag times (Table 7 and Table 9) that range from 10 to 13 days. The process that these models describe is not clearly understood. It is interesting to note that in both the black only and full racial models that a significant model was not found for the rural only region. It is possible that the urban region is what is driving this somewhat nebulous process and making the urban-rural model to be significant.

Hierarchical Bayesian Model Discussion

The use of a hierarchical Bayesian model allows for correlation between grid cells to be modeled with random effects. By breaking with the assumption in GEE that the grid cells are independent, the Bayesian models can “borrow” information from all grid cells in calculating parameter estimates. The resulting parameter estimates are smoothed as they are made to shrink towards a mean value and the calculation is based on a much larger effective sample size than if the grid cells were considered independently (Agresti et al. 2000). The resulting conditional or subject specific interpretation of the model allows risk to vary between grid cells as a result of the cell specific random effects (Figure 6 - Figure 8).

A direct comparison between the population-averaged GEE models and the conditional Bayesian model is difficult. However, it is noted that the parameter estimates are similar between the two model classes (Table 6 and Table 10). Each Bayesian model uses a single random effect that is modeled with an exchangeable prior for the urban

model and with a CAR prior for the rural and urban-rural models. The variances for each of the random effect distributions are listed in Table 10. In all cases the variance parameter is greater than 0 which implies that residual correlation does exist between grid cells and is being accounted for by the random effects. It follows that the standard errors for the parameter estimates in the GEE model are deflated and an over-dispersion parameter would be useful to obtain more realistic confidence intervals for the models listed in Table 6 - Table 9. The Bayesian model is a particularly useful extension to the marginal GEE model in this study as it allows the identification of particular grid cells that are associated with a relatively high risk. The random effects also give a more accurate estimate of the expected number of hospital visits. The display of this information on the maps in Figure 6 - Figure 8 further aids in the visual identification of areas where extreme values tend to congregate.

Strengths and Limitations

Previous studies have focused on summer or warm temperature ozone studies as this is the time when ozone is most concentrated due to the photochemical processes that are responsible for its generation (Schildcrout et al. 2006; Weisel et al. 1995). This study includes daily ozone data over the entire 3-year study period and models show it is significant while controlling for month thus supporting inference of the effect of ozone during all seasons. To the knowledge of the author, this is the first asthma related pollution study to be done based on local estimates of exposure over a large regional scale on the order in size of the state of Mississippi. This was made possible by the use of remotely sensed pollution via satellite which supports the inclusion of populated areas

where data from ground based sensors is not available. The satellite data also gives flexibility with respect to the scale of the study design. The study presented in this thesis was of relatively small scale covering the entire state of Mississippi with a regular grid cell network. The grid cells were large relative to the pixels resolution of the satellite data and so the satellite data resolution was effectively reduced to match that of the study grid. Studies that require a smaller study area defined for example by zip code or metropolitan or rural statistical area boundaries can spatially aggregate the satellite data to fit their particular needs.

Inference in this study is based at the grid cell level, with parameter estimates affecting the relative risk of hospital visitation. It is understood that this relative risk is based on the assumption of homogeneous risk over the entire population. This leads to a possible problem with the ecological fallacy. Asthma-related hospital visits within each grid cell are assumed to be from individuals with preexisting asthmatic conditions. However, the majority of the population within each grid cell is not asthmatic. It follows that inference is based on a population without homogeneous risk. Because the results of this work are similar to those based on individual-level covariates from previous studies, it is believed that any bias resulting from assumptions of homogeneous risk is not particularly strong.

The use of the regular network of grid cells was a convenient mechanism by which to aggregate hospital visitation, ozone and $PM_{2.5}$ exposure and demographic covariates from various geographic units. In this way a basis exists for straightforward statistical models. More complex models are possible that explicitly consider misaligned spatial data in support of analysis of the various data used in this thesis based on its

original spatial scale of collection (Banerjee et al. 2004). In this way it is possible to avoid problems associated with bias induced by spatial aggregation as mentioned above.

The results of the GEE and Bayesian models may also be biased to some degree from missing data. As stated previously, about 20% of the data was missing within the rural region of the state. By not including the hospital visitation rates and demographic characteristics of these areas the parameter estimates for the rural and urban-rural models may have some bias associated with them. It is also believed that the aggregated number of daily hospital visits for each grid cell may be biased towards the black population. This segment of the population is generally believed to be underserved in the healthcare system and more likely included in the aggregate visit outcome in this study. Individuals with better access to healthcare are more likely to manage their disease and be less likely to suffer from spikes in pollution. These individuals, assumed to be represented more within the white segment of the population, are less likely to be recorded. Finally, the use of the CAR prior in the urban only model did not take into account that the region is composed of a number of island or disconnected sub-regions. Hodges et al. (2003) shows that properly accounting for the number of islands may have an important impact on all posterior summaries. It is not known how this would affect the model DIC with respect to the urban model using the exchangeable prior listed in Table 10.

A possible problem with the convergence of the Markov chains was observed for the percent black and female categorical variables for the models in Table 10. The chains did not appear to reach full convergence after 3,000 iterations. The parameter estimates for these variables shown in Table 10 may be biased and the 95% credibility intervals too wide. It was observed that upon a second run of both the urban and urban-rural models

that this problem with convergence had little to no observable effect on the exposure parameter estimates. The chains and posterior distributions for all fixed effect parameters for the simulations are given in the appendix.

Conclusion

Satellites measure pollution concentrations through a column of air from the top of the atmosphere to the surface of the earth where a measurement is desired. Further research is required to enhance the quality of this measurement. PM_{2.5} measurements calculated from satellite data have been shown to have up to a 96% correlation with ground level measurement devices. Satellite measurements of O₃ are less reliable (Weinhold 2008). Because the daily value for each grid cell is an average of many satellite measurements, it is assumed that much of the variability is smoothed out. It is not known how precise the exposure measurements are in this study, but it is assumed at the very least that the models are showing the effects of O₃ and PM_{2.5} on a relative scale. Future work with satellite based epidemiological studies, such as the one in this thesis, will become more promising as remotely sensed exposure estimates for the criteria air pollutants become more accurate.

Figures 1 - 8

Figure 1: Mississippi Asthma Study area

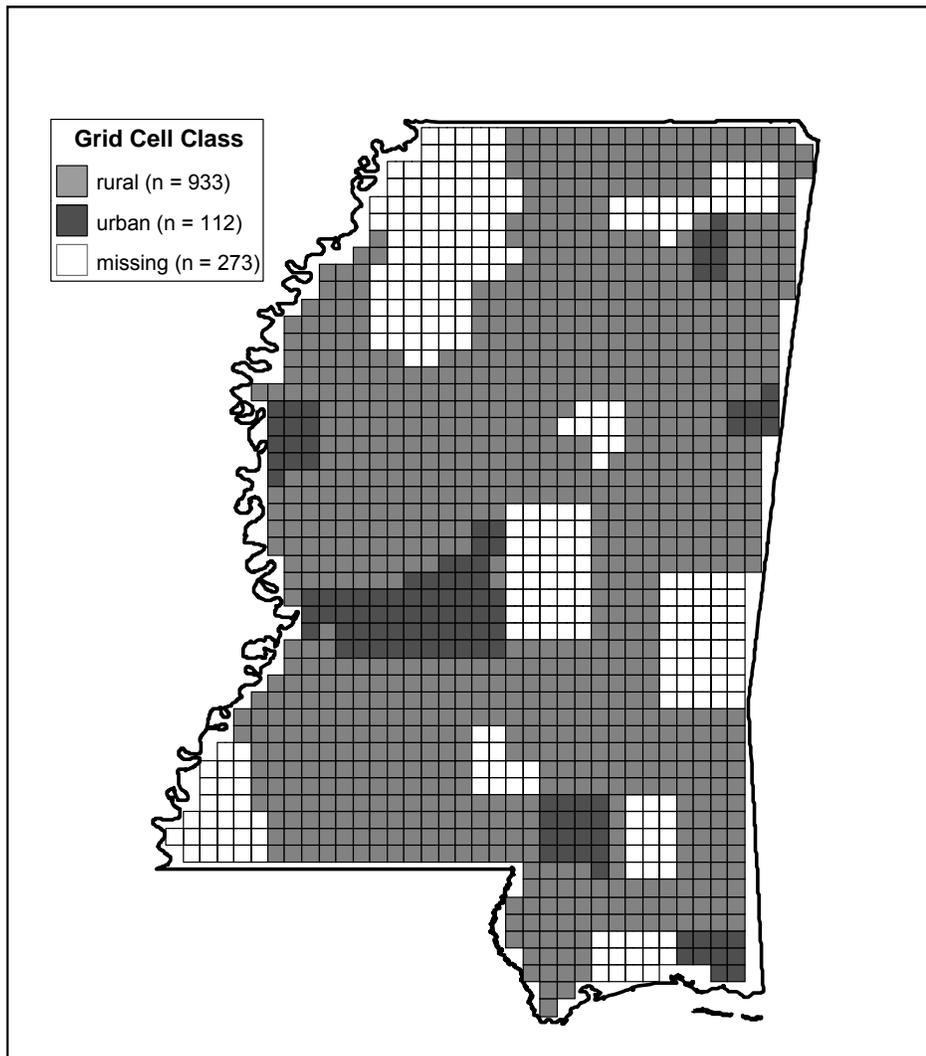


Figure 2: Data Subsets Used for Bayesian Analysis



Figure 3: Exposure and Hospital Visit Time Series

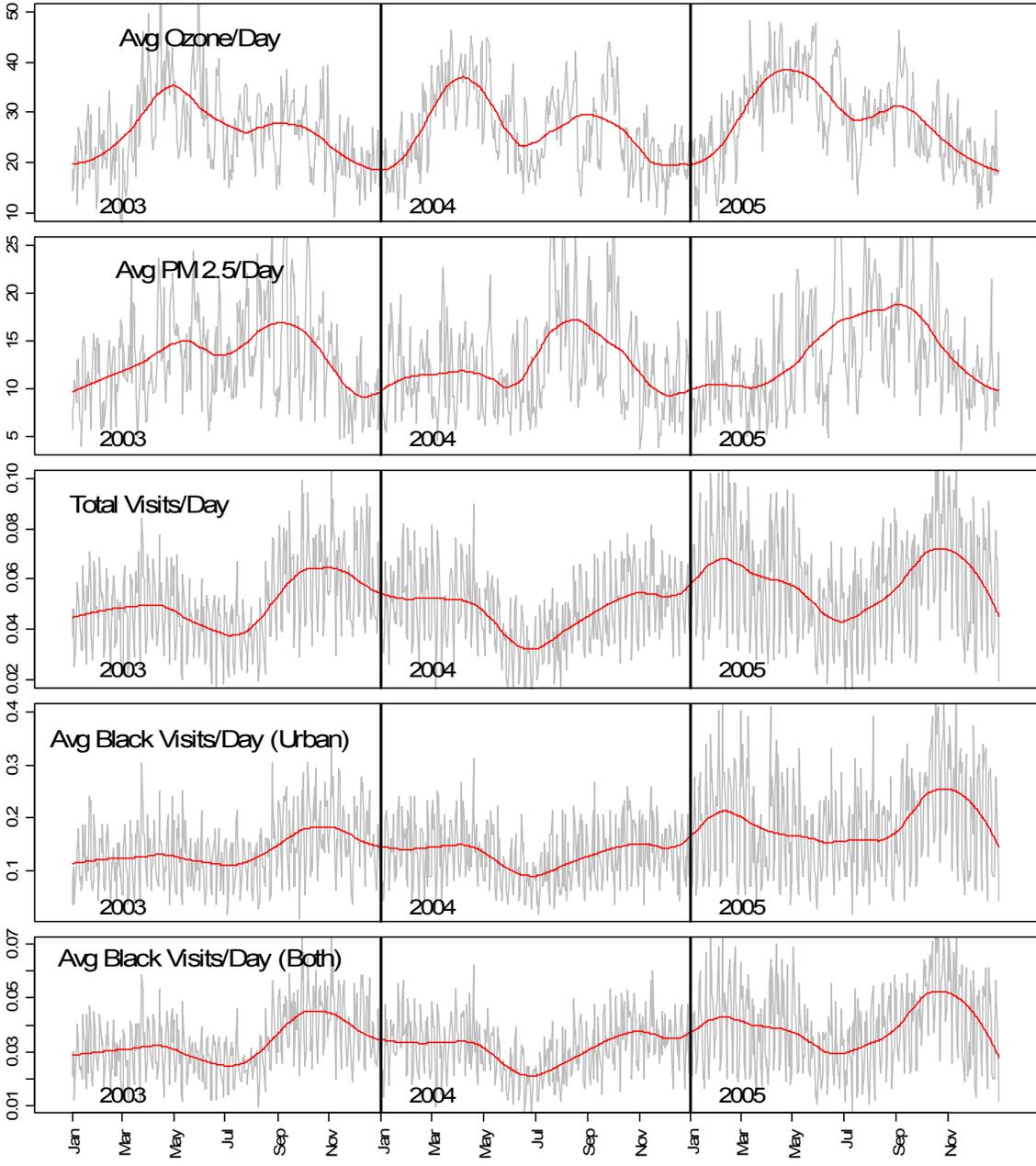


Figure 4: Inverse Correlation of the Convolution Prior Variances

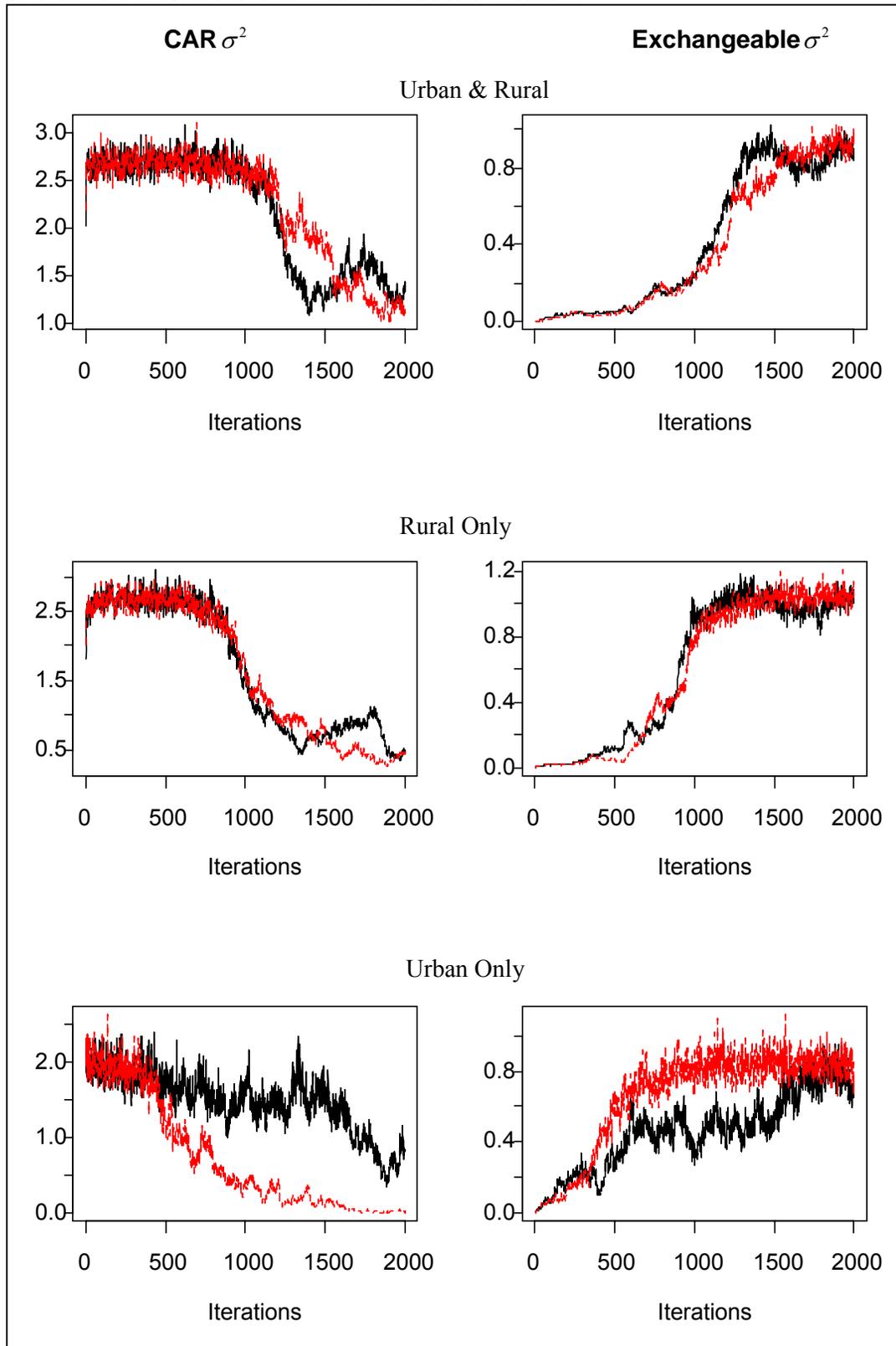


Figure 5: Expectation Plots

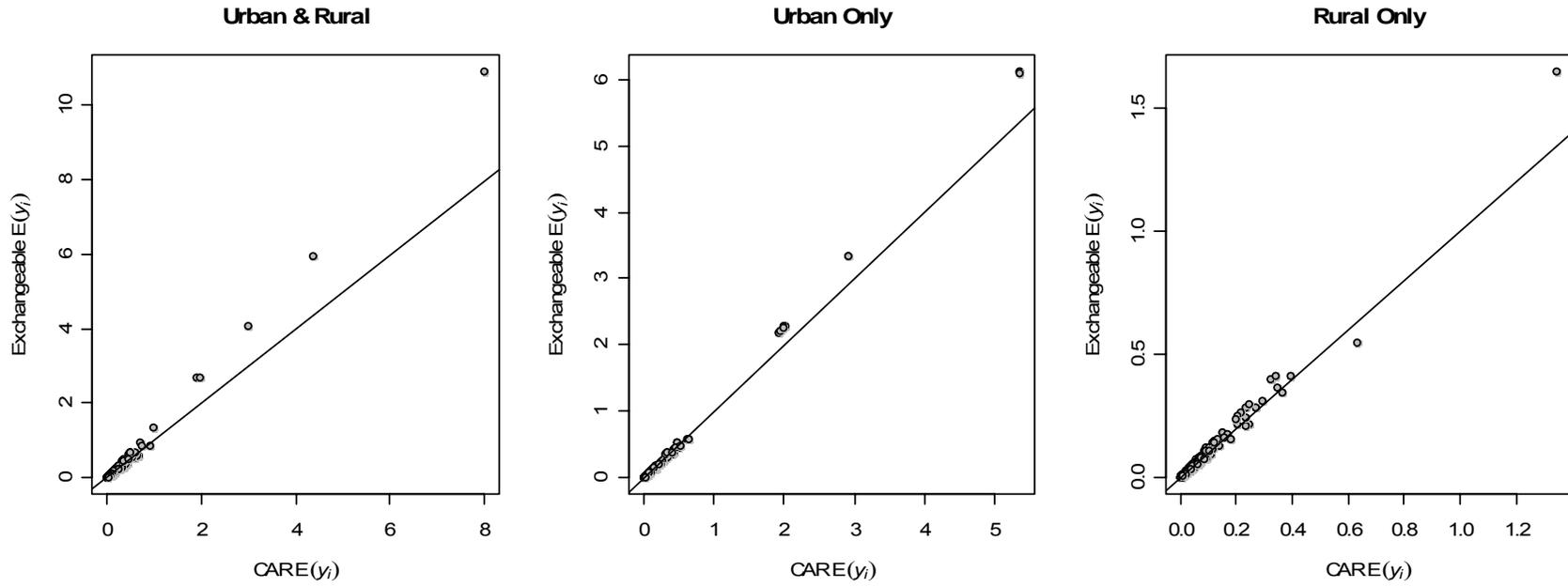


Figure 6: Spatial Distribution of Risk and Expectation - Urban and Rural

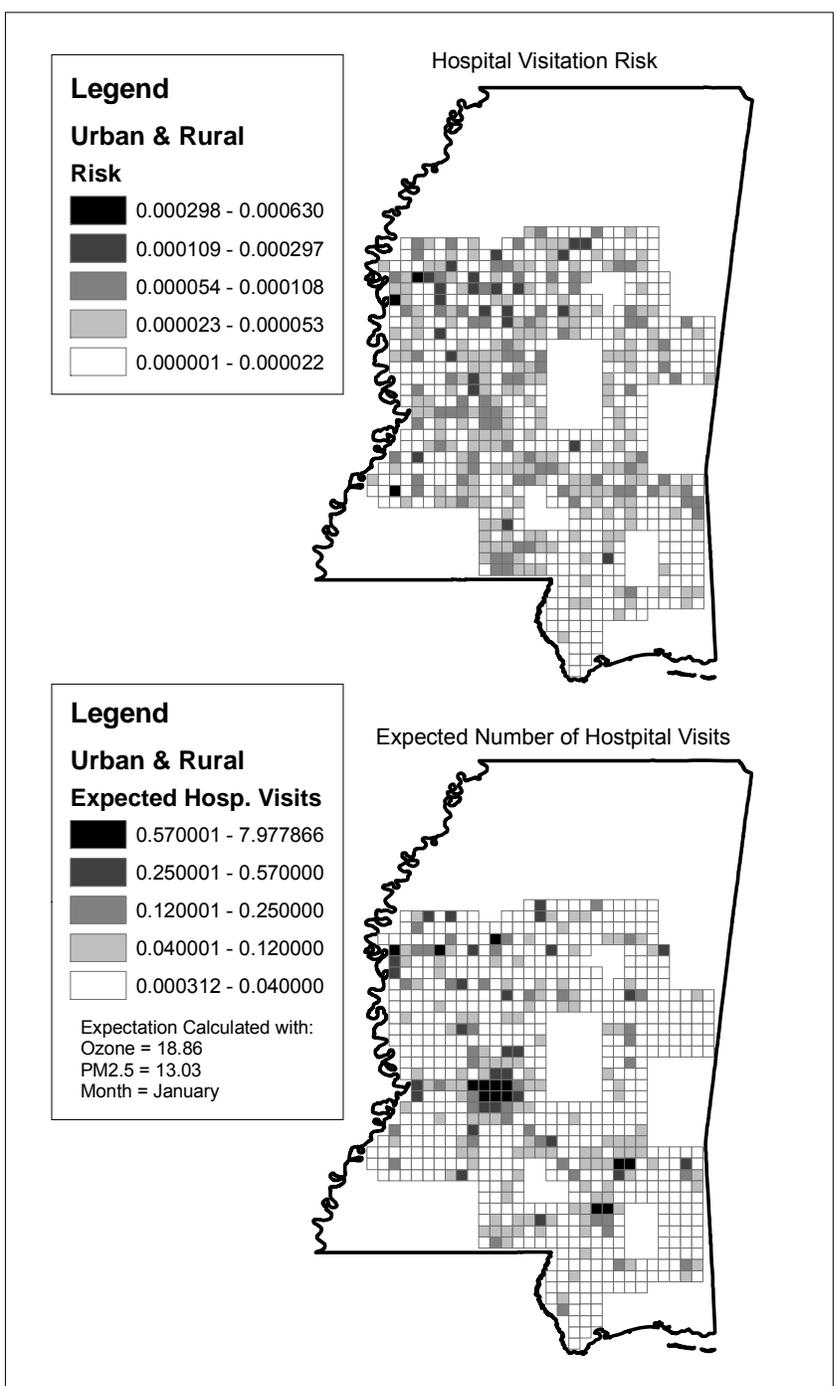


Figure 7: Spatial Distribution of Risk and Expectation - Urban

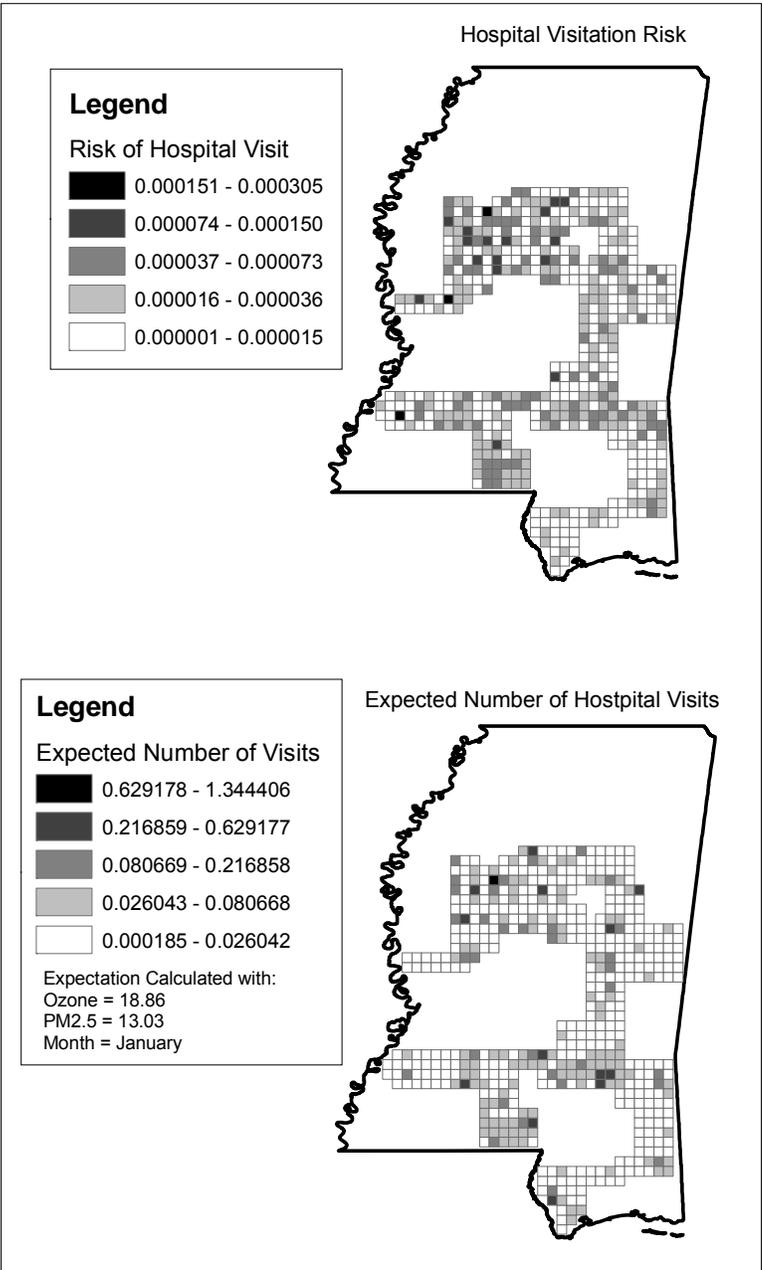
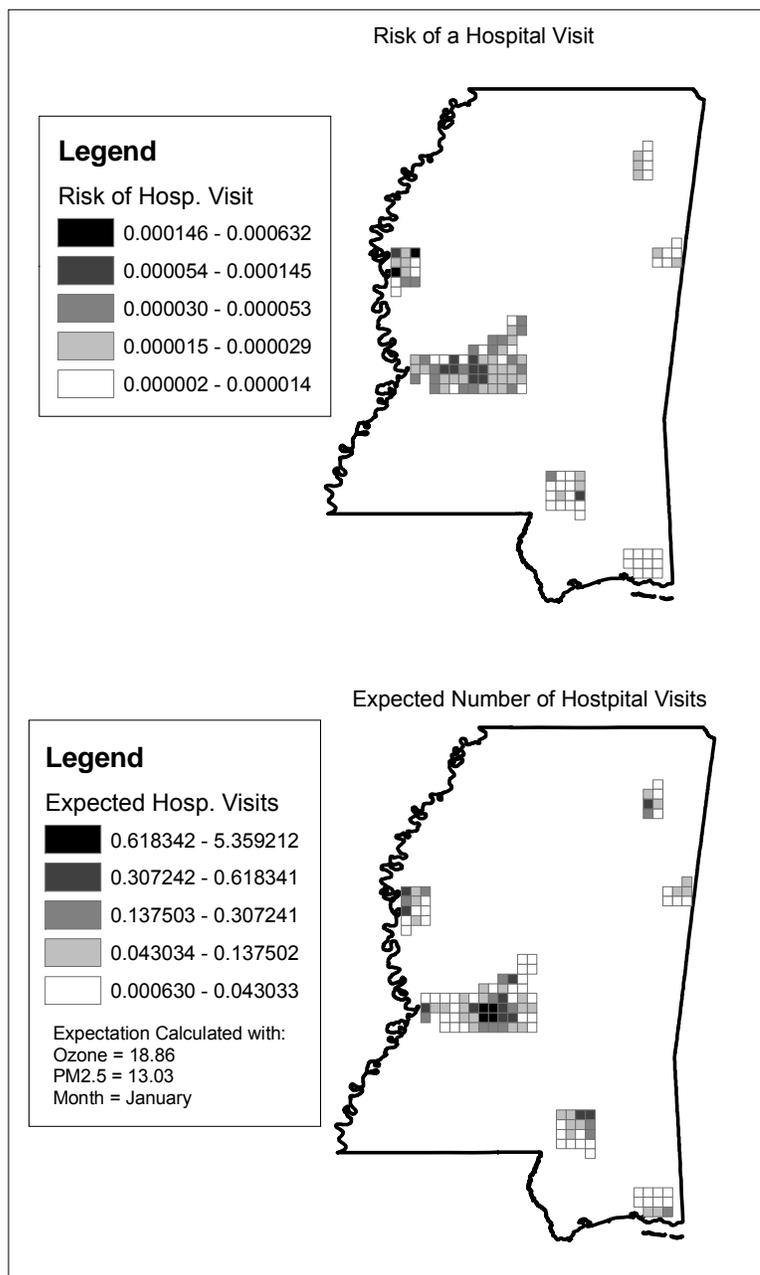


Figure 8: Spatial Distribution of Risk and Expectation - Rural

Tables 1 - 10

Table 1: Population Characteristics

	Composite Urban & Rural (n = 1045)		
Characteristic	2003 Mean (SD)	2004 Mean (SD)	2005 Mean (SD)
<i>Total Population</i>	1967626*	1973962*	1980706*
<i>Avg. Population (individuals)</i>	1882.9 (4215.94)	1888.96 (4229.56)	1895.41 (4244.48)
% Black	37.86% (0.24)	38.14% (0.24)	38.44% (0.25)
% Female	50.69% (0.04)	50.63% (0.04)	50.58% (0.04)
<i>Total Hospital Visits</i>	19241*	18021*	22487*
<i>Avg. Hospital Visits</i>	18.41	17.24	21.52
% Black	66.09%	66.38%	67.15%
<i>Per Capita Income</i>	16548.58 (3412.57)	16547.78 (3413.02)	16546.8 (3413.73)
<i>Median Household Income</i>	31499.13 (7447.75)	32251.13 (7633.82)	33021.98 (7827.05)
	Rural Only (n = 933)		
Characteristic	2003 Mean (SD)	2004 Mean (SD)	2005 Mean (SD)
<i>Total Population</i>	1257377*	1257973*	1258772*
<i>Avg. Population (individuals)</i>	1347.67 (1966.21)	1348.31 (1963.55)	1349.17 (1961.27)
% Black	38.01% (0.24)	38.28% (0.24)	38.57% (0.24)
% Female	50.61% (0.04)	50.56% (0.04)	50.50% (0.04)
<i>Total Hospital Visits</i>	11581*	10880*	12343*
<i>Avg. Hospital Visits</i>	12.41	11.66	13.23
% Black	61.28%	60.62%	59.39%
<i>Per Capita Income</i>	16126.86 (2682.33)	16125.26 (2681.33)	16123.58 (2680.56)
<i>Median Household Income</i>	30342.78 (5234.6)	31062.64 (5334.01)	31800.3 (5436.82)
	Urban Only (n = 112)		
Characteristic	2003 Mean (SD)	2004 Mean (SD)	2005 Mean (SD)
<i>Total Population</i>	710249*	715989*	721934*
<i>Avg. Population (individuals)</i>	6341.51 (10594.66)	6392.76 (10628.21)	6445.84 (10665.16)
% Black	36.66% (0.26)	37.04% (0.27)	37.41% (0.27)
% Female	51.31% (0.02)	51.28% (0.02)	51.25% (0.02)
<i>Total Hospital Visits</i>	7660*	7141*	10144*
<i>Avg. Hospital Visits</i>	68.39	63.76	90.57
% Black	73.37%	75.17%	76.58%
<i>Per Capita Income</i>	20061.71 (5930.63)	20067.58 (5932.41)	20072.37 (5935.3)
<i>Median Household Income</i>	41131.96 (13666.42)	42151.74 (14086.62)	43198.99 (14524.55)

* Total count for all cells for study year

Table 2: Exposure lagged correlation matrix

		PM 2.5														
		lag 0	lag1	lag 2	lag3	lag4	lag5	lag6	lag7	lag8	lag9	lag10	lag11	lag12	lag13	lag14
Ozone	lag 0	0.375	0.325	0.249	0.197	0.160	0.138	0.122	0.104	0.106	0.105	0.095	0.080	0.066	0.073	0.077
	lag1	0.330	0.375	0.325	0.248	0.197	0.159	0.138	0.121	0.104	0.107	0.106	0.096	0.080	0.066	0.073
	lag2	0.272	0.329	0.375	0.325	0.248	0.196	0.158	0.137	0.121	0.104	0.108	0.107	0.096	0.080	0.065
	lag3	0.238	0.271	0.330	0.375	0.324	0.247	0.195	0.157	0.137	0.121	0.106	0.108	0.107	0.095	0.080
	lag4	0.222	0.239	0.271	0.330	0.375	0.325	0.248	0.196	0.157	0.136	0.121	0.106	0.108	0.107	0.095
	lag5	0.226	0.223	0.239	0.272	0.330	0.376	0.325	0.248	0.196	0.157	0.136	0.121	0.106	0.109	0.107
	lag6	0.211	0.225	0.223	0.238	0.271	0.330	0.375	0.324	0.248	0.196	0.159	0.136	0.121	0.105	0.108
	lag7	0.176	0.211	0.226	0.223	0.238	0.271	0.329	0.375	0.324	0.248	0.198	0.159	0.136	0.120	0.105
	lag8	0.138	0.175	0.211	0.225	0.222	0.237	0.270	0.329	0.375	0.324	0.249	0.198	0.159	0.136	0.120
	lag9	0.121	0.138	0.176	0.211	0.225	0.222	0.237	0.270	0.329	0.375	0.325	0.249	0.198	0.159	0.136
	lag10	0.104	0.120	0.138	0.175	0.210	0.224	0.222	0.236	0.270	0.330	0.377	0.325	0.249	0.197	0.159
	lag11	0.091	0.103	0.121	0.137	0.174	0.209	0.224	0.220	0.236	0.271	0.333	0.378	0.326	0.249	0.197
	lag12	0.087	0.091	0.103	0.120	0.137	0.174	0.209	0.223	0.220	0.236	0.272	0.333	0.378	0.326	0.249
	lag13	0.100	0.086	0.091	0.102	0.120	0.136	0.173	0.207	0.223	0.221	0.238	0.273	0.333	0.378	0.325
	lag14	0.111	0.099	0.086	0.090	0.102	0.119	0.135	0.172	0.207	0.223	0.223	0.239	0.273	0.333	0.378

Table 3: PM 2.5 Lag Time Correlation Matrix

		PM 2.5														
PM 2.5	Variable	lag 0	lag 1	lag 2	lag 3	lag 4	lag 5	lag 6	lag 7	lag 8	lag 9	lag 10	lag 11	lag 12	lag 13	lag 14
	lag 0	1	0.74	0.5	0.39	0.32	0.28	0.25	0.19	0.17	0.17	0.15	0.14	0.14	0.15	0.17
	lag 1	0.74	1	0.74	0.5	0.39	0.32	0.28	0.24	0.19	0.17	0.17	0.15	0.14	0.14	0.15
	lag 2	0.5	0.74	1	0.74	0.5	0.39	0.32	0.28	0.24	0.19	0.17	0.17	0.15	0.14	0.14
	lag 3	0.39	0.5	0.74	1	0.74	0.5	0.39	0.32	0.28	0.24	0.2	0.17	0.17	0.15	0.14
	lag 4	0.32	0.39	0.5	0.74	1	0.74	0.5	0.39	0.32	0.28	0.25	0.2	0.17	0.17	0.15
	lag 5	0.28	0.32	0.39	0.5	0.74	1	0.74	0.5	0.39	0.32	0.28	0.25	0.2	0.17	0.17
	lag 6	0.25	0.28	0.32	0.39	0.5	0.74	1	0.74	0.5	0.39	0.32	0.28	0.25	0.2	0.17
	lag 7	0.19	0.24	0.28	0.32	0.39	0.5	0.74	1	0.74	0.5	0.39	0.32	0.28	0.25	0.19
	lag 8	0.17	0.19	0.24	0.28	0.32	0.39	0.5	0.74	1	0.74	0.5	0.39	0.32	0.28	0.25
	lag 9	0.17	0.17	0.19	0.24	0.28	0.32	0.39	0.5	0.74	1	0.74	0.5	0.39	0.32	0.28
	lag 10	0.15	0.17	0.17	0.2	0.25	0.28	0.32	0.39	0.5	0.74	1	0.74	0.5	0.39	0.32
	lag 11	0.14	0.15	0.17	0.17	0.2	0.25	0.28	0.32	0.39	0.5	0.74	1	0.74	0.5	0.39
	lag 12	0.14	0.14	0.15	0.17	0.17	0.2	0.25	0.28	0.32	0.39	0.5	0.74	1	0.74	0.5
	lag 13	0.15	0.14	0.14	0.15	0.17	0.17	0.2	0.25	0.28	0.32	0.39	0.5	0.74	1	0.74
	lag 14	0.17	0.15	0.14	0.14	0.15	0.17	0.17	0.19	0.25	0.28	0.32	0.39	0.5	0.74	1

Table 4: Ozone Lag Time Correlation Matrix

		Ozone														
Ozone	Variable	lag 0	lag 1	lag 2	lag 3	lag 4	lag 5	lag 6	lag 7	lag 8	lag 9	lag 10	lag 11	lag 12	lag 13	lag 14
	lag 0	1.00	0.79	0.62	0.54	0.50	0.48	0.46	0.46	0.44	0.43	0.41	0.40	0.38	0.37	0.36
	lag 1	0.79	1.00	0.79	0.61	0.54	0.50	0.48	0.46	0.46	0.44	0.43	0.41	0.40	0.38	0.37
	lag 2	0.62	0.79	1.00	0.79	0.62	0.54	0.50	0.48	0.46	0.46	0.44	0.43	0.41	0.40	0.38
	lag 3	0.54	0.61	0.79	1.00	0.79	0.62	0.54	0.50	0.48	0.46	0.46	0.44	0.43	0.41	0.39
	lag 4	0.50	0.54	0.62	0.79	1.00	0.79	0.62	0.54	0.50	0.48	0.46	0.46	0.44	0.43	0.41
	lag 5	0.48	0.50	0.54	0.62	0.79	1.00	0.79	0.62	0.54	0.50	0.48	0.46	0.46	0.44	0.43
	lag 6	0.46	0.48	0.50	0.54	0.62	0.79	1.00	0.79	0.62	0.54	0.50	0.48	0.46	0.46	0.44
	lag 7	0.46	0.46	0.48	0.50	0.54	0.62	0.79	1.00	0.79	0.62	0.54	0.50	0.48	0.46	0.46
	lag 8	0.44	0.46	0.46	0.48	0.50	0.54	0.62	0.79	1.00	0.79	0.62	0.54	0.50	0.48	0.46
	lag 9	0.43	0.44	0.46	0.46	0.48	0.50	0.54	0.62	0.79	1.00	0.79	0.62	0.54	0.50	0.48
	lag 10	0.41	0.43	0.44	0.46	0.46	0.48	0.50	0.54	0.62	0.79	1.00	0.79	0.62	0.54	0.50
	lag 11	0.40	0.41	0.43	0.44	0.46	0.46	0.48	0.50	0.54	0.62	0.79	1.00	0.79	0.62	0.54
	lag 12	0.38	0.40	0.41	0.43	0.44	0.46	0.46	0.48	0.50	0.54	0.62	0.79	1.00	0.79	0.62
	lag 13	0.37	0.38	0.40	0.41	0.43	0.44	0.46	0.46	0.48	0.50	0.54	0.62	0.79	1.00	0.79
	lag 14	0.36	0.37	0.38	0.39	0.41	0.43	0.44	0.46	0.46	0.48	0.50	0.54	0.62	0.79	1.00

Table 5: PM 2.5 and O3 single exposure model parameter estimates

All Races		Black Only	
<i>Urban and Rural</i>		<i>Urban and Rural</i>	
Parameter	Relative Risk (95% CI)	Parameter	Relative Risk (95% CI)
O ₃ lag 3	1.004 (1.001, 1.006)	O ₃ lag 4	1.004 (1.002, 1.006)
O ₃ lag10	1.005 (1.003, 1.007)	O ₃ lag 12	1.004 (1.002, 1.006)
PM _{2.5} 5	1.005 (1.001, 1.008)	PM _{2.5} lag5	1.005 (1.001, 1.009)
PM _{2.5} 12	1.008 (1.005, 1.01)	PM _{2.5} lag12	1.008 (1.004, 1.011)
<i>Rural only</i>		<i>Rural only</i>	
Parameter	Relative Risk (95% CI)	Parameter	Relative Risk (95% CI)
O ₃ lag 2	1.004 (1.002, 1.006)	O ₃ lag 2	1.004 (1.002, 1.006)
O ₃ lag 10	1.004 (1.002, 1.006)	O ₃ lag 10	1.004 (1.002, 1.007)
PM _{2.5} lag4	1.003 (1.001, 1.005)		
<i>Urban Only</i>		<i>Urban Only</i>	
Parameter	Relative Risk (95% CI)	Parameter	Relative Risk (95% CI)
O ₃ lag 3	1.003 (1.001, 1.005)	O ₃ lag 3	1.003 (1.001, 1.005)
O ₃ lag 10	1.004 (1.003, 1.006)	O ₃ lag 10	1.004 (1.002, 1.006)
PM _{2.5} lag4	1.003 (1.001, 1.004)	PM _{2.5} lag4	1.002 (1.001, 1.004)
PM _{2.5} lag12	1.003 (1.001, 1.005)	PM _{2.5} lag12	1.004 (1.001, 1.007)

Table 6: Parameter Estimates for All Races in Lower Lag Range

Urban & Rural Lower Range		Urban Only Lower Range		Rural Only Lower Range	
<i>Parameter</i>	<i>Relative Risk (95% CI)</i>	<i>Parameter</i>	<i>Relative Risk (95% CI)</i>	<i>Parameter</i>	<i>Relative Risk (95% CI)</i>
Intercept	0(0, 0)	Intercept	0(0, 0)	Intercept	0(0, 0)
ozone lag 2	1.003(1.001, 1.005)	ozone lag 2	1.003(1.001, 1.006)	ozone lag 2	1.003(1.001, 1.005)
pm 2.5 lag 4	1.002(1, 1.004)	pm 2.5 lag 5	1.004(1, 1.008)	pm 2.5 lag 4	1.002(1, 1.004)
% black Q4	1.969(1.529, 2.535)	% black Q4	3.688(2.688, 5.061)	% black Q4	1.402(1.161, 1.694)
% black Q3	1.301(1.048, 1.615)	% black Q3	1.331(0.949, 1.866)	% black Q3	1.242(1.045, 1.477)
% black Q2	1.109(0.9, 1.368)	% black Q2	1.462(0.987, 2.166)	% black Q2	1.141(0.96, 1.357)
% black Q1	Reference	% black Q1	Reference	% black Q1	Reference
% female Q4	1.508(1.15, 1.978)	Jan	0.926(0.875, 0.98)	% female Q4	1.954(1.603, 2.383)
% female Q3	1.084(0.825, 1.425)	Feb	0.926(0.862, 0.995)	% female Q3	1.336(1.106, 1.615)
% female Q2	1.321(0.98, 1.781)	Mar	0.88(0.817, 0.947)	% female Q2	1.313(1.105, 1.56)
% female Q1	Reference	Apr	0.837(0.785, 0.891)	% female Q1	Reference
Jan	0.95(0.908, 0.994)	May	0.731(0.683, 0.783)	Jan	0.97(0.91, 1.034)
Feb	0.988(0.939, 1.039)	Jun	0.715(0.676, 0.757)	Feb	1.017(0.952, 1.087)
Mar	0.904(0.855, 0.955)	Jul	0.664(0.621, 0.71)	Mar	0.907(0.84, 0.98)
Apr	0.909(0.863, 0.957)	Aug	0.766(0.715, 0.822)	Apr	0.932(0.866, 1.004)
May	0.773(0.735, 0.813)	Sep	0.865(0.811, 0.922)	May	0.782(0.73, 0.838)
Jun	0.686(0.649, 0.724)	Oct	1.071(1.016, 1.129)	Jun	0.65(0.602, 0.701)
Jul	0.657(0.624, 0.692)	Nov	0.991(0.948, 1.035)	Jul	0.638(0.59, 0.689)
Aug	0.819(0.774, 0.865)	Dec	Reference	Aug	0.836(0.772, 0.906)
Sep	0.923(0.882, 0.967)			Sep	0.943(0.884, 1.005)
Oct	1.1(1.051, 1.152)			Oct	1.098(1.026, 1.176)
Nov	1.06(1.016, 1.105)			Nov	1.099(1.036, 1.166)
Dec	Reference			Dec	Reference

Table 7: Parameter Estimates for All Races in Upper Lag Range

Urban & Rural Upper Range		Urban Only Upper Range		Rural Only Upper Range	
<i>Parameter</i>	<i>Relative Risk (95% CI)</i>	<i>Parameter</i>	<i>Relative Risk (95% CI)</i>	<i>Parameter</i>	<i>Relative Risk (95% CI)</i>
Intercept	0(0, 0)	Intercept	0(0, 0)	NA	NA
ozone lag 12	1.002(1, 1.003)	ozone lag 10	1.006(1.003, 1.008)		
pm 2.5 lag 10	1.002(1, 1.004)	pm 2.5 lag 12	1.006(1.003, 1.009)		
% black Q4	1.97(1.531, 2.534)	% black Q4	3.682(2.682, 5.055)		
% black Q3	1.3(1.048, 1.612)	% black Q3	1.325(0.945, 1.859)		
% black Q2	1.109(0.9, 1.366)	% black Q2	1.459(0.986, 2.158)		
% black Q1	Reference	% black Q1	Reference		
% female Q4	1.509(1.152, 1.976)	Jan	0.946(0.886, 1.009)		
% female Q3	1.083(0.826, 1.421)	Feb	0.92(0.856, 0.989)		
% female Q2	1.323(0.984, 1.778)	Mar	0.862(0.798, 0.932)		
% female Q1	Reference	Apr	0.804(0.752, 0.86)		
Jan	0.967(0.922, 1.014)	May	0.709(0.659, 0.763)		
Feb	0.992(0.941, 1.046)	Jun	0.694(0.654, 0.736)		
Mar	0.916(0.867, 0.968)	Jul	0.645(0.602, 0.692)		
Apr	0.92(0.869, 0.975)	Aug	0.729(0.677, 0.785)		
May	0.782(0.741, 0.826)	Sep	0.84(0.791, 0.891)		
Jun	0.69(0.655, 0.727)	Oct	1.028(0.97, 1.09)		
Jul	0.658(0.624, 0.694)	Nov	0.973(0.93, 1.019)		
Aug	0.818(0.768, 0.872)	Dec	Reference		
Sep	0.932(0.886, 0.98)				
Oct	1.098(1.048, 1.15)				
Nov	1.056(1.012, 1.103)				
Dec	Reference				

Table 8: Parameter Estimates for Blacks Only in Lower Lag Range

Urban & Rural Lower Range		Urban Only Lower Range		Rural Only Lower Range	
<i>Parameter</i>	<i>Relative Risk (95% CI)</i>	<i>Parameter</i>	<i>Relative Risk (95% CI)</i>	<i>Parameter</i>	<i>Relative Risk (95% CI)</i>
NA	NA	Intercept	0(0, 0)	NA	NA
		ozone lag 4	1.004(1.001, 1.006)		
		pm 2.5 lag 2	1.003(1.001, 1.005)		
		Jan	0.904(0.835, 0.977)		
		Feb	0.882(0.82, 0.949)		
		Mar	0.869(0.79, 0.956)		
		Apr	0.822(0.754, 0.895)		
		May	0.723(0.659, 0.793)		
		Jun	0.715(0.659, 0.776)		
		Jul	0.676(0.616, 0.741)		
		Aug	0.773(0.708, 0.843)		
		Sep	0.895(0.829, 0.966)		
		Oct	1.127(1.057, 1.2)		
		Nov	1.025(0.972, 1.081)		
		Dec	Reference		

Table 9: Parameter Estimates for Blacks Only in Upper Lag Range

Urban & Rural Upper Range		Urban Only Upper Range		Rural Only Upper Range	
<i>Parameter</i>	<i>Relative Risk (95% CI)</i>	<i>Parameter</i>	<i>Relative Risk (95% CI)</i>	<i>Parameter</i>	<i>Relative Risk (95% CI)</i>
Intercept	0(0, 0)	Intercept	0(0, 0)	NA	NA
ozone lag 10	1.003(1.001, 1.005)	ozone lag 12	1.002(1, 1.004)		
pm 2.5 lag 13	1.003(1.001, 1.005)	pm 2.5 lag 12	1.006(1.003, 1.009)		
% female Q4	1.709(1.094, 2.672)	% female Q4	0.687(0.399, 1.183)		
% female Q3	1.087(0.71, 1.665)	% female Q3	0.489(0.287, 0.833)		
% female Q2	1.439(0.872, 2.372)	% female Q2	0.399(0.201, 0.791)		
% female Q1	Reference	% female Q1	Reference		
Jan	0.967(0.911, 1.028)	Jan	0.925(0.845, 1.013)		
Feb	0.948(0.892, 1.008)	Feb	0.891(0.831, 0.956)		
Mar	0.887(0.825, 0.955)	Mar	0.876(0.803, 0.955)		
Apr	0.912(0.85, 0.978)	Apr	0.84(0.776, 0.909)		
May	0.769(0.719, 0.823)	May	0.734(0.672, 0.801)		
Jun	0.691(0.646, 0.738)	Jun	0.722(0.67, 0.778)		
Jul	0.673(0.629, 0.721)	Jul	0.671(0.61, 0.739)		
Aug	0.835(0.769, 0.907)	Aug	0.757(0.691, 0.83)		
Sep	0.978(0.91, 1.05)	Sep	0.892(0.82, 0.972)		
Oct	1.196(1.12, 1.278)	Oct	1.12(1.052, 1.192)		
Nov	1.104(1.04, 1.171)	Nov	1.015(0.964, 1.068)		
Dec	Reference	Dec	Reference		

Table 10: Bayesian Model Parameter Estimates

Urban & Rural		Urban Only		Rural Only	
<i>Parameter</i>	<i>95% Credible Interval</i>	<i>Parameter</i>	<i>95% Credible Interval</i>	<i>Parameter</i>	<i>95% Credible Interval</i>
Intercept	0 (0,0)	Intercept	0 (0, 0)	Intercept	0 (0,0)
ozone lag 2	1.004 (1.003,1.006)	ozone lag 2	1.004 (1.002, 1.006)	ozone lag 2	1.004 (1.002,1.006)
pm 2.5 lag 4	1.003 (1.001,1.005)	pm 2.5 lag 5	1.004 (1.002, 1.007)	pm 2.5 lag 4	1.003 (1.000,1.006)
% black Q4	3.06 (2.505,3.792)	% black Q4	1.419 (1.015, 1.996)	% black Q4	2.987 (2.364,3.888)
% black Q3	2.096 (1.614,2.784)	% black Q3	1.185 (0.919, 1.543)	% black Q3	2.214 (1.805,2.773)
% black Q2	1.309 (1.089,1.511)	% black Q2	0.928 (0.766, 1.152)	% black Q2	1.529 (1.229,1.892)
% black Q1	Reference	% black Q1	Reference	% black Q1	Reference
% female Q4	0.677 (0.613,0.759)	Jan	0.925 (0.875, 0.981)	% female Q4	0.671 (0.479,0.827)
% female Q3	0.77 (0.675,0.87)	Feb	0.942 (0.888, 1.001)	% female Q3	0.987 (0.763,1.203)
% female Q2	1.03 (0.912,1.159)	Mar	0.867 (0.812, 0.926)	% female Q2	1.077 (0.854,1.294)
% female Q1	Reference	Apr	0.839 (0.783, 0.9)	% female Q1	Reference
Jan	0.952 (0.91,0.995)	May	0.729 (0.681, 0.78)	Jan	0.982 (0.919,1.053)
Feb	0.977 (0.935,1.021)	Jun	0.707 (0.661, 0.754)	Feb	1.044 (0.978,1.125)
Mar	0.879 (0.84,0.919)	Jul	0.669 (0.63, 0.716)	Mar	0.902 (0.843,0.974)
Apr	0.865 (0.822,0.909)	Aug	0.764 (0.714, 0.816)	Apr	0.918 (0.848,0.996)
May	0.736 (0.701,0.774)	Sep	0.858 (0.802, 0.915)	May	0.782 (0.725,0.846)
Jun	0.676 (0.645,0.708)	Oct	1.079 (1.018, 1.143)	Jun	0.638 (0.589,0.692)
Jul	0.641 (0.61,0.673)	Nov	1.001 (0.945, 1.06)	Jul	0.603 (0.557,0.653)
Aug	0.79 (0.756,0.827)	Dec	Reference	Aug	0.81 (0.755,0.872)
Sep	0.896 (0.857,0.937)	σ^2 Exchangeable	0.949 (0.811, 1.11)	Sep	0.915 (0.849,0.991)
Oct	1.078 (1.033,1.125)			Oct	1.06 (0.991,1.137)
Nov	1.059 (1.014,1.105)			Nov	1.098 (1.029,1.178)
Dec	Reference			Dec	Reference
σ^2 CAR	2.751 (2.573,2.941)			σ^2 CAR	2.69 (2.465,2.935)

Bibliography

- Agresti A, Booth JG, Hobert JP, Caffo B. 2000. Random-Effects Modeling of Categorical Response Data. *Sociological Methodology* 30: 27-80.
- American Lung Association. 2007. Trends in Asthma Morbidity and Mortality: Epidemiology & Statistics Unit Research and Program Services.
- Balmes JR. 1993. The Role of Ozone Exposure in the Epidemiology of Asthma. *Environmental Health Perspectives* 101: 219-224.
- Banerjee S, Carlin B, Gelfand A. 2004 Hierarchical Modeling and Analysis for Spatial Data. Boca Raton: Chapman and Hall/CRC.
- Besag J, Kooperberg CL. 1995. On conditional and intrinsic autoregressions *Biometrika* 82: 733-746.
- Besag J, York J, Mollié A. 1991. Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics* 43(1): 1-20.
- Breslow NE, Clayton DG. 1993. Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association* 88: 9-25.
- Cleveland WS, Grosse E, Shyu WM. 1992. Chapter 8 of *Statistical Models in S* Wadsworth & Brooks/Cole.
- Delfino RJ, Zeiger RS, Seltzer JM, Street DH, McLaren CE. 2002. Association of Asthma Symptoms with Peak Particulate Air Pollution and Effect Modification by Anti-Inflammatory Medication Use. *Environmental Health Perspectives* 110(10): A607-A617.
- Environmental Protection Agency. 2008. What Are the Six Common Air Pollutants? Available: <http://www.epa.gov/air/urbanair/> [accessed November 2008].
- Gelman A, Rubin DB. 1992. Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science* 7(4): 457-472.
- Gilmour MI, Jaakkola MS, London SJ, Nel AE, Rogers CA. 2006. How Exposure to Environmental Tobacco Smoke, Outdoor Air Pollutants, and Increased Pollen Burdens Influences the Incidence of Asthma. *Environmental Health Perspectives* 114(4): 627-633.
- Hodges JS, Carlin BP, Fan Q. 2003. On the Precision of the Conditionally Autoregressive Prior in Spatial Models. *Biometrics* 59(2): 317-322.
- Koren HS. 1995. Associations between criteria air pollutants and asthma. *Environmental Health Perspectives Supplements* 103: 235.

- Lewis TC, Robins TG, Dvornch JT, Keeler GJ, Yip FY, Mentz GB, et al. 2005. Air Pollution–Associated Changes in Lung Function among Asthmatic Children in Detroit. *Environmental Health Perspectives Supplements* 113(8): 1068–1075.
- Liang K-Y, Zeger SL. 1986. Longitudinal Data Analysis Using Generalized Linear Models. *Biometrika* 73(1): 13-22.
- McCullagh P, Nelder JA. 1989. *Generalized Linear Models*: Chapman & Hall/CRC.
- Schildcrout JS, Sheppard L, Lumley T, Slaughter JC, Koenig JQ, Shapiro GG. 2006. Ambient Air Pollution and Asthma Exacerbations in Children: An Eight-City Analysis. *Am J Epidemiol* 164(6): 505-517.
- Sheppard L, Levy D, Norris G, Larson TV, Koenig JQ. 1999. Effects of Ambient Air Pollution on Nonelderly Asthma Hospital Admissions in Seattle, Washington, 1987-1994. *Epidemiology* 10(1): 23-30.
- Spiegelhalter DJ, Best NG, Carlin BP, Linde Avd. 2002. Bayesian Measures of Model Complexity and Fit. *Journal of the Royal Statistical Society Series B (Statistical Methodology)* 64(4): 583-639.
- Spiegelhalter DJ, Thomas A, Best NG. 2003. *WinBUGS Version 1.4 User Manual*. MRC Biostatistics Unit.
- Sun H-L, Chou M-C, Lue K-H. 2006. The relationship of air pollution to ED visits for asthma differ between children and adults. *The American Journal of Emergency Medicine* 24(6): 709-713.
- Tolbert P, Mulholland J, MacIntosh D, Xu F, Daniels D, Devine O, et al. 2000. Air Quality and Pediatric Emergency Room Visits for Asthma in Atlanta, Georgia. *American Journal of Epidemiology* 151(8): 798-810.
- van Eeden SF, Tan WC, Suwa T, Mukae H, Terashima T, Fujii T, et al. 2001. Cytokines Involved in the Systemic Inflammatory Response Induced by Exposure to Particulate Matter Air Pollutants (PM10). *Am J Respir Crit Care Med* 164(5): 826-830.
- Waller LA, Gotway CA. 2004. *Applied Spatial Statistics for Public Health Data*: John Wiley & Sons.
- Weinhold B. 2008. The global sweep of pollution: satellite snapshots capture long-distance movement. *Environmental Health Perspectives* 116(8): A338-A345.
- Weisel CP, Cody RP, Liroy PJ. 1995. Relationship between Summertime Ambient Ozone Levels and Emergency Department Visits for Asthma in Central New Jersey. *Environmental Health Perspectives* 103: 97-102.

Appendix

Appendix 1: WinBUGS Model, Urban-Rural region with convolution prior

```

model{
  for(g in 1 : nGrids) {
    for(y in 1:nYears){
      for(m in 1:nMonths){
        for(d in 1:daysPerMonth[y,m]){

          ag_visits[m,d,y,g]~dpois(mu[m,d,y,g])

          log(mu[m,d,y,g]) <- log(cellPop[g, y]) + alpha0

          + alpha.o3_lag2 * o3_lag2[m,d,y,g]
          + alpha.pm25_lag4 * pm25_lag4[m,d,y,g]

          + alpha.m1 * m1[m]
          + alpha.m2 * m2[m]
          + alpha.m3 * m3[m]
          + alpha.m4 * m4[m]
          + alpha.m5 * m5[m]
          + alpha.m6 * m6[m]
          + alpha.m7 * m7[m]
          + alpha.m8 * m8[m]
          + alpha.m9 * m9[m]
          + alpha.m10 * m10[m]
          + alpha.m11 * m11[m]

          + alpha.black_q2 * black_q2[g, y]
          + alpha.black_q3 * black_q3[g, y]
          + alpha.black_q4 * black_q4[g, y]

          + alpha.fem_q2 * fem_q2[g,y]
          + alpha.fem_q3 * fem_q3[g,y]
          + alpha.fem_q4 * fem_q4[g,y]

          + b_grid[g]
          + b_CAR[g]

        } #end day iterator
      } #end month iterator
    } #end year iterator

    b_grid[g] ~ dnorm(0.0, tau.b_grid)

  } #end grid iterator and main model spec

```

```

#Spatial CAR Random Effect Prior
b_CAR[1:nGrids] ~ car.normal(adj[], weights[], numNeigh[], tau.b_CAR)
for(k in 1:sumNumNeigh){
  weights[k] <- 1
}

# priors:
#intercept must be dflat to account for CAR parameterized to sum to 0
alpha0 ~ dflat()
alpha.o3_lag2 ~ dnorm(0.0,1.0E-4)
alpha.pm25_lag4 ~ dnorm(0.0,1.0E-4);
alpha.m1 ~ dnorm(0.0,1.0E-4)
alpha.m2 ~ dnorm(0.0,1.0E-4)
alpha.m3 ~ dnorm(0.0,1.0E-4)
alpha.m4 ~ dnorm(0.0,1.0E-4)
alpha.m5 ~ dnorm(0.0,1.0E-4)
alpha.m6 ~ dnorm(0.0,1.0E-4)
alpha.m7 ~ dnorm(0.0,1.0E-4)
alpha.m8 ~ dnorm(0.0,1.0E-4)
alpha.m9 ~ dnorm(0.0,1.0E-4)
alpha.m10 ~ dnorm(0.0,1.0E-4)
alpha.m11 ~ dnorm(0.0,1.0E-4)

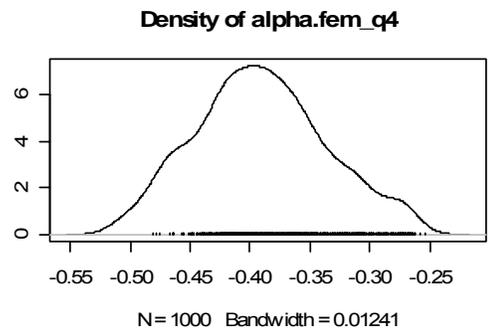
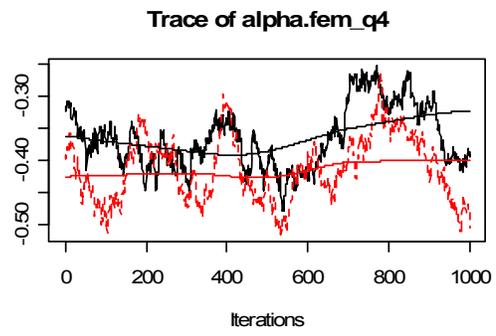
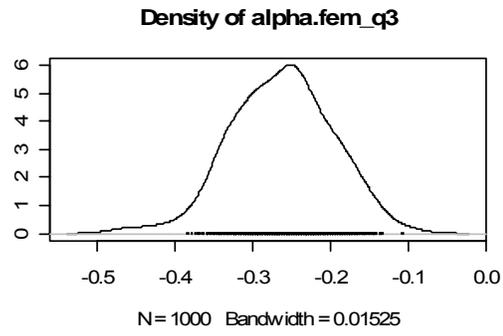
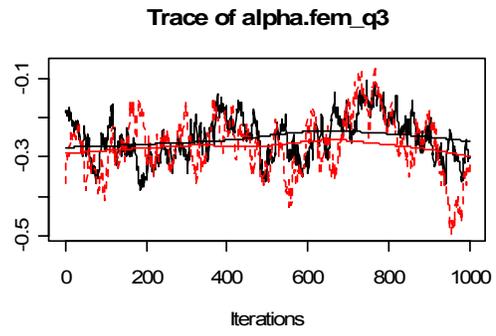
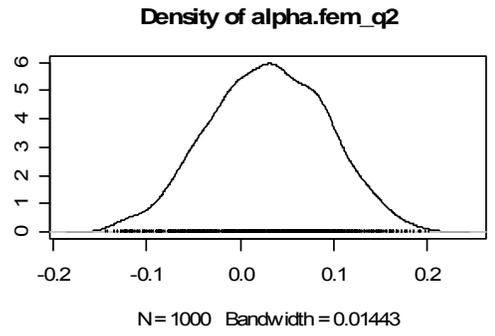
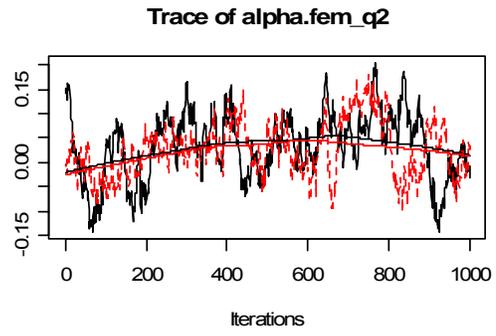
alpha.black_q2 ~ dnorm(0.0,1.0E-4)
alpha.black_q3 ~ dnorm(0.0,1.0E-4)
alpha.black_q4 ~ dnorm(0.0,1.0E-4)

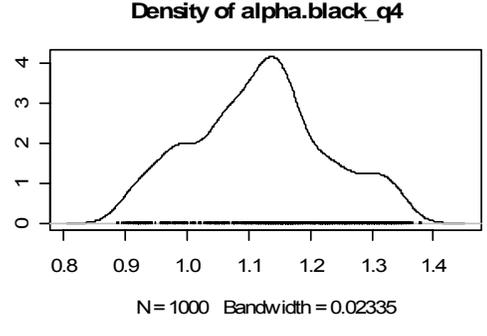
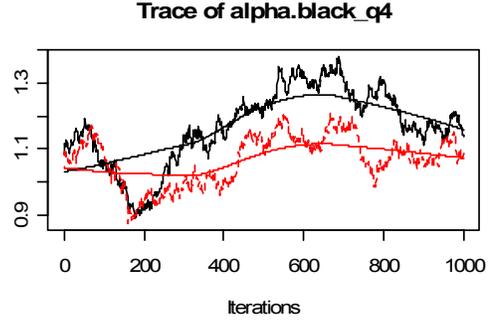
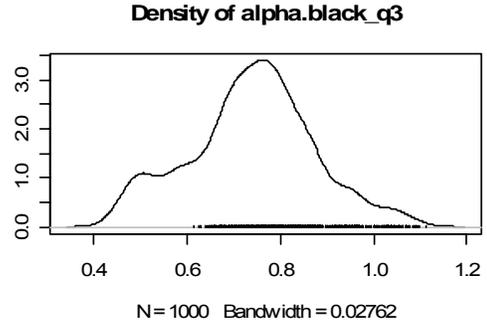
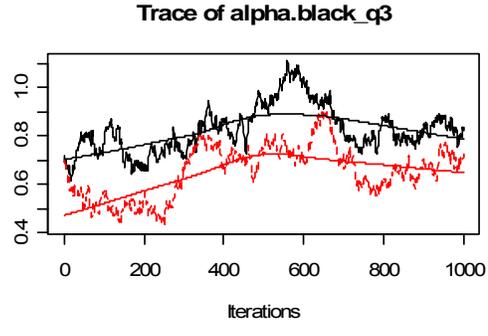
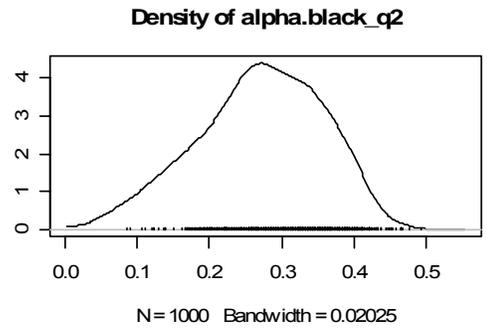
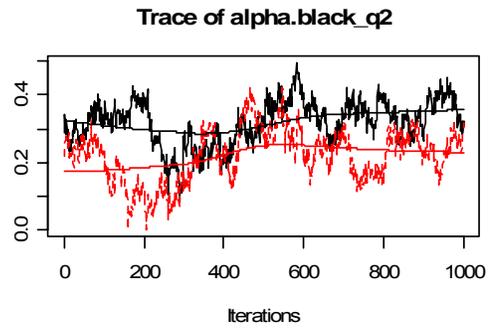
alpha.fem_q2 ~ dnorm(0.0,1.0E-4)
alpha.fem_q3 ~ dnorm(0.0,1.0E-4)
alpha.fem_q4 ~ dnorm(0.0,1.0E-4)

tau.b_grid ~ dgamma(1.0E-3,1.0E-3); sigma.b_grid <- sqrt(1.0/tau.b_grid)
tau.b_CAR ~ dgamma(1.0E-3,1.0E-3); sigma.b_CAR <- sqrt(1.0/tau.b_CAR)
}

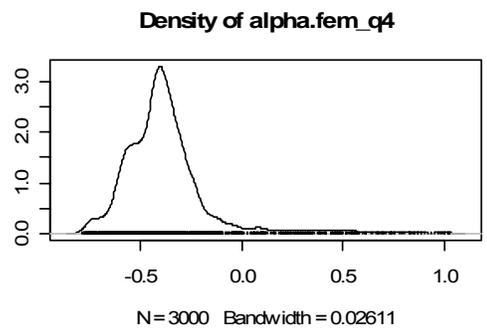
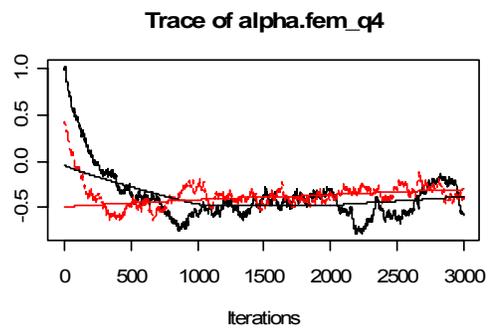
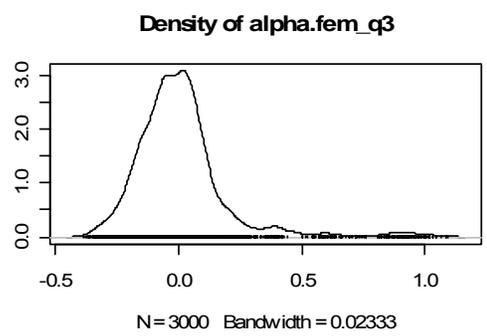
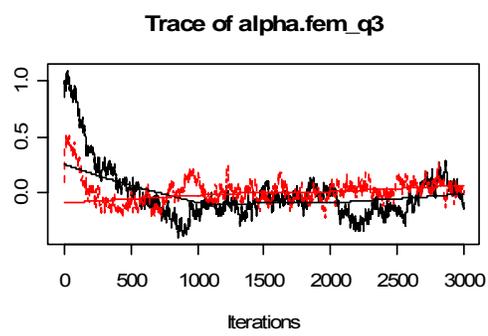
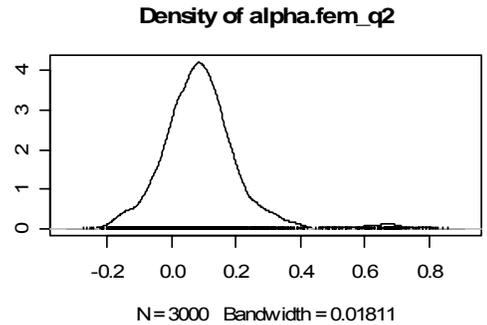
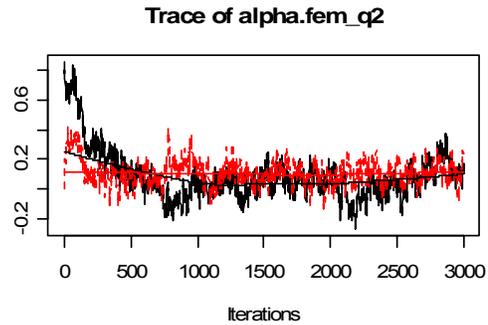
```

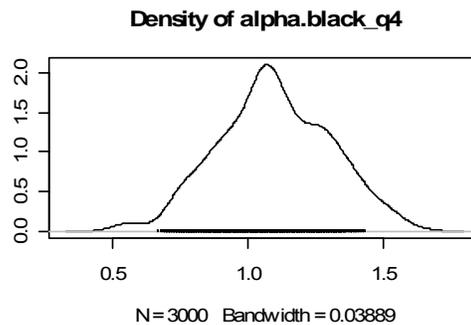
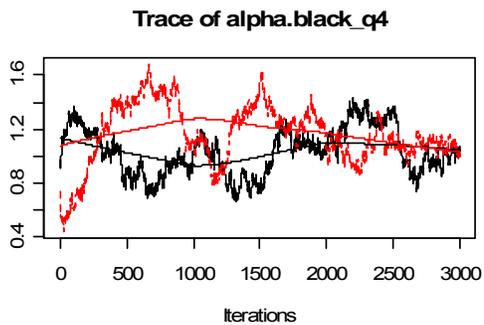
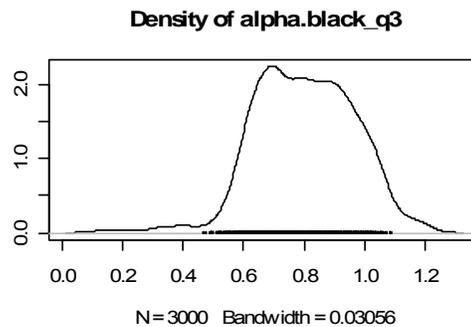
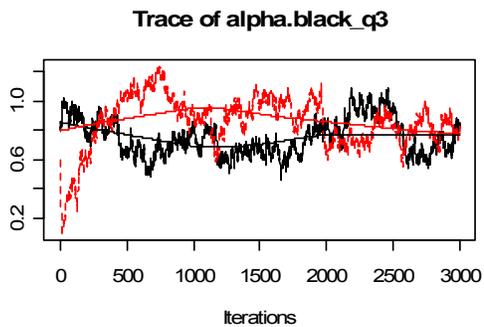
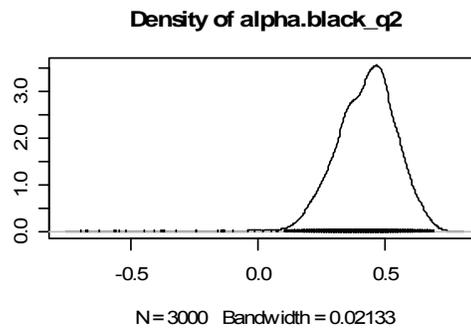
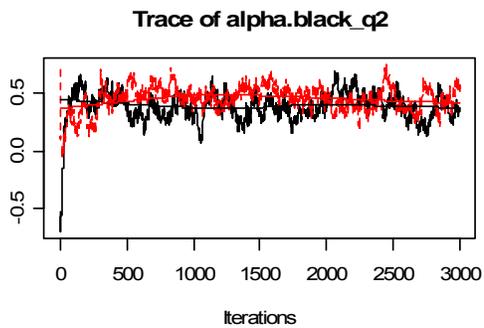
Appendix 2: Urban – Rural, Selected fixed effect MCMC chains and posterior distributions from CAR model





Appendix 3: Rural Only, Selected fixed effect MCMC chains and posterior distributions from CAR model





Appendix 4: Urban Only, Selected fixed effect MCMC chains and posterior distributions from Exchangeable model

