Integration of health outcomes, air quality, and socio-economic data in Northwest Florida

Dr. Johan Liebens
Department of Environmental Studies
University of West Florida

Dr. Zhiyong Hu
Department of Environmental Studies
University of West Florida

Dr. K. Ranga Rao
Center for Environmental Diagnostics and Bioremediation
University of West Florida

August 15, 2009

Table of Contents

1 Introduction 1
2 Health Tracking Study 1
    2.1 Introduction 1
    2.2 Results 2
    2.3 Health Tracking Conclusions 5
3 Air Emitter Proximity Study 6
    3.1 General 6
    3.2 Cumulative Health Outcomes 6
    3.3 Specific Health Outcomes 8
    3.4 Environmental Inequity 11
    3.5 Proximity Study Conclusions 11
4 Air Toxics Modeling 11
    4.1 Introduction 11
    4.2 Results 12
        4.2.1 Stationary Point Sources 12
        4.2.2 Traffic Sources 14
        4.2.3 Acute Health Risks From Emissions From Crist Plant 14
    4.3 Air Toxics Modeling Conclusions 14
5 Geostatistical Study 15
    5.1 Introduction 15
    5.2 Data 15
    5.3 Mapping, Analyses And Modeling Results 15
        5.3.1 Mapping 16
        5.3.2 Exploratory Spatial Data Analysis (ESDA) 16
        5.3.3 Spatial Lag Model Of Asthma And Air Pollution 16
        5.3.4 Mortality Rates Of COPD, Stroke, And Lung Cancer Compared With Socio-Economic And Environmental Factors 18
            5.3.4.1 Focused Score Tests 18
            5.3.4.2 OLS Regression Analyses 19

5.3.5 Bayesian Hierarchical Modeling Of Stroke Mortality And Air Pollution, Income, And Greenness 21
5.3.6 Correlating MODIS Aerosol Optical Depth Data With Ground-Based PM2.5 Observations 22
5.3.7 Relationships Between Myocardial Infarction (MI) And AOD 24
5.3.8 Relationships Between Chronic Coronary Heart Disease (CCHD) And AOD 26
5.3.9 Association Of CCHD With PM2.5 28
6 Overall Conclusions 29
7 Acknowledgements 30
8 Outcomes And Outputs 30
8.1 Final Reports 30
8.2 Dissemination Of Results: Conference Presentations 30
8.3 Dissemination Of Results: Academic Publications 34
8.4 Dissemination Of Results: Media Reports 35
8.5 Graduate Students Trained At UWF 35
8.6 Outcomes 35

List of Tables

Table 1. Summary of results from mortality models. 2

Table 2. Summary of results from morbidity models. 4

Table 3. Spatial lag model of asthma and air pollution. 18

Table 4. Focused score tests of relationships between mortality rates of COPD, stroke, and lung cancer with socio-economic and environmental factors. 18

Table 5. Univariate OLS regression of disease rates on suspected factors. 20

Table 6. Multivariate OLS regression for COPD. 20

Table 7. Multivariate OLS regression for stroke. 21

Table 8. Multivariate OLS regression for lung cancer. 21

Table 9 Markov chain Monte Carlo results for Bayesian hierarchical modelling of stroke mortality vs. income, air pollution, and greenness. 22

Table 10. Correlation and geographically weighted regression results for the four PM monitoring sites in the study area. 23

Table 11. MI and AOD: spatial error model. 25

Table 12. Markov chain Monte Carlo simulation results for Bayesian hierarchical modeling of MI vs. AOD. 25

Table 13. Spatial error regression model. 28

Table 14. CCHD, AOD and PM2.5 data for Escambia and Santa Rosa Counties. 28

Table 15. Results of Bayesian hierarchical modeling. 29

List of Figures

Figure 1. Map of summary of mortality models. 3

Figure 2. Map of summary of morbidity models. 4

Figure 3. Comparison of birth defects in Escambia and Santa Rosa County ZIP codes and associated (matching) ZIP codes. 5

Figure 4a. Benzene weighted TRI proximity index for NW FL ZIP codes with better cumulative health outcomes vs. associated ZIP codes. 7

Figure 4b. Benzene weighted TRI proximity index for NW FL ZIP codes with worse cumulative health outcomes vs. associated ZIP codes. 8

Figure 5a. Proximity index for Zip codes with high or low incidence for specific causes of mortality: Comparison within northwest Florida. 9

Figure 5b. Proximity index for Zip codes with high or low incidence for specific causes of morbidity: Comparison within northwest Florida. 9

Figure 6a. Proximity index for Zip codes with high or low incidence for specific causes of mortality: Comparison with matching Zip codes. 10

Figure 6. Proximity index for Zip codes with high or low incidence for specific causes of morbidity: Comparison with matching Zip codes. 10

Figure 7. Elevated chronic cancer risk zones estimated by the RAIMI system. 13

Figure 8. Maps of point source polluters. 16

Figure 9. Zip-code level environmental exposure map. 17

Figure 10. Kernel estimates of the posterior density of the fixed effects in the Bayesian hierarchical model. 22

Figure 11. PM2.5 surface calculated by merging MODIS AOD and EPA PM2.5 ground measurements. 23

Figure 12. Raster surface of correlation between $PM_{2.5}$ and MODIS AOD. 24

Figure 13. R squares of geographically weighted regression of PM2.5 on MODIS AOD. 24

Figure 14. Coefficient kernel density curves, Bayesian hierarchical model

(SMR of MI vs. AOD) for eastern US. 26

Figure 15. CCHD rate and AOD: bivariate Moran's I scatter plot. 27

Figure 16. CCHD rate and AOD: local indicator of spatial association (LISA). 27

Figure 17. Kernel estimates of the posterior densities of the fixed effects in the Bayesian hierarchical model. 29

1 INTRODUCTION

The local incidence of some illnesses is perceived to be elevated by many citizens in Escambia and Santa Rosa Counties. Many assume that this perceived elevated incidence is due to the multitude of environmental pollution issues in the area, especially air pollution. To address this concern of the citizens, PERCH carried out a health tracking study that compares mortality and morbidity rates in the two-county area with those elsewhere in the state of Florida. To relate the results of this Zip code level health study to air pollution, PERCH evaluated relationships between the health outcomes and the proximity of the Zip codes to air emission sites. This proximity study pointed to possible connections between some specific health outcomes and the location of air emission sites, but was somewhat limited in its potential by the large spatial unit (Zip codes) used in the health tracking study. Therefore, PERCH further assessed the connections between air pollution and health outcomes with two other approaches, air toxics modeling and geostatistical modeling. Results of these four studies (health tracking, proximity analysis, air toxics modeling, and geostatistical modeling) are briefly summarized and related to each other in the remainder of this report.

2 HEALTH TRACKING STUDY

2.1 Introduction

There is considerable interest in being able to relate geographic patterns of exposure to air pollution to variation in the health status of populations. These spatially defined associations between environmental hazards and population health are referred to as environmental health tracking studies. In its health tracking study, PERCH adopted a strategy to investigate if specific Zip codes in Escambia and Santa Rosa Counties had rates of health indicators that were higher (at a statistically significant level) than matched comparison Zip codes (http://www.uwf.edu/CEDB/Perch_USF_EPA_April04.pdf). Patterns of such differences across multiple indicators and age/race strata could suggest potential Zip codes in the region that should be targeted for study as more detailed environmental data becomes available.

The investigators of the health outcome study (PERCH project collaborators from the University of South Florida) identified a list of health indicators that were thought to be sensitive to increased exposure to airborne environmental hazards. Both mortality and morbidity indicators were included. The mortality indicators, which might best be characterized as health conditions that would be affected by long-term exposure to environmental hazards, included deaths from: (1) all cancers, (2) lung cancers alone, (3) cardiovascular diseases, (4) any respiratory disease, (5) birth defects, and (6) all causes of death to infants. The morbidity indicators, which might best be characterized as being sensitive to short-term exposure to environmental hazards, included hospitalizations for (1) asthma, (2) cardiovascular disease, and (3) respiratory disease. Two additional morbidity indicators, numbers of live births with very low birth weight and the number of live births with low birth weight were added. While these indicators have not been directly associated with airborne environmental exposures they have been widely used as indicators of poor population health.

A cross-sectional observational study design was employed to compare the selected health outcome measures for Zip codes within the two counties with Zip codes having similar demographic and socio-economic characteristics from the remainder of Florida (matches). Escambia and Santa Rosa Zip codes were matched using propensity scores that were calculated through a series of logistic regression equations that included the following independent variables: percent female, percent of the population over 65 years of age, percent of the population that is black, percent of the population that is Hispanic, total population, the percent of households earning $15,000 or less, and per capita income.

A series of generalized linear (Poisson regression) models were developed to test whether standardized 5-year mortality/morbidity ratios for each Escambia and Santa Rosa County Zip code were different than those from the matched comparison Zip codes at a statistically significant level. Specifically, least squares means for each health indicator studied, for five years of data, were compared after adjusting for age, gender, percent of the population that is black, percent of the population over the age of 65 and the percent of households earning $15,000 per year or less. Five years of data were included in the analysis to measure impact over a reasonable time period and to increase the power of the statistical tests.

2.2 Results

The majority of statistically significant differences for mortality related to birth defect and infant mortality. Only one Zip code (32570) had significantly higher rates across more than one category of disease (Table 1, Figure 1). Table 1 summarizes the results of the models for mortality by listing the number of health outcomes that had a statistically significantly higher incidence in the given Zip code than in the matching Zip codes. These results are graphically depicted in Figure 1.

Table 1. Summary of results from mortality models.

Zip code	County	Blacks		Whites	
		All ages	Over 65	All ages	Over 65
32570	SR	3s/1w*	2s/1w	1s/1w	1s/1w
32566/ 32561	SR	2w	2w	-	3s/1w
32534	ES	2s/1w	1s/1w	-	-
32501	ES	1s/2w	1s/1w	1s	-
32577	ES	1s/2w	-	2s	-
32533	ES	1s/1w	-	-	-
32503	ES	1s/2w	-	-	-

* s indicates strong statistical evidence, w indicates weak statistical evidence.

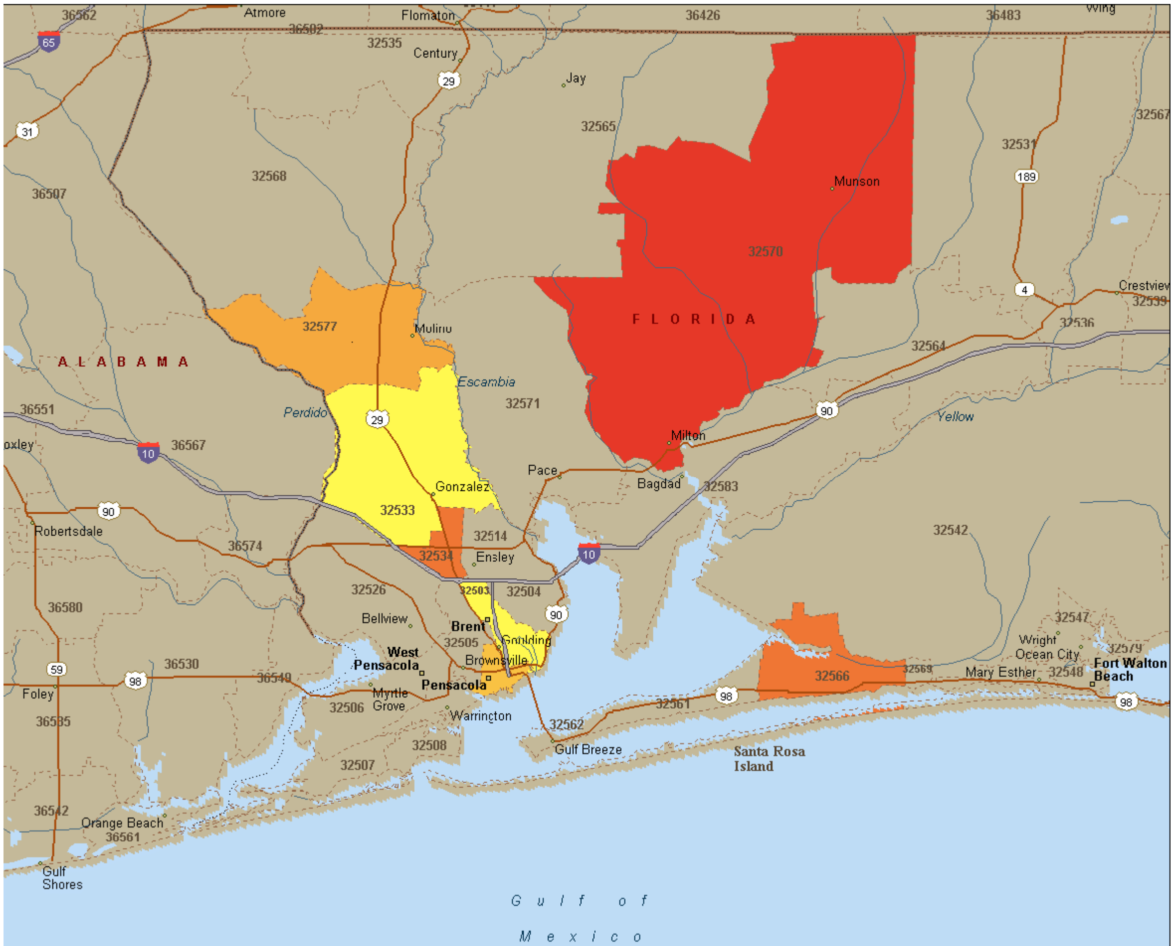


Figure 1. Map of summary of mortality models. Darker colors indicate a greater burden of disease.

Fewer statistically significant differences in which Zip codes had higher rates of disease than the matching Zip codes were found in the morbidity models than in the mortality models. The morbidity health indicators, unlike the mortality indicators, fall primarily into one disease group (cardio-respiratory). There were relatively few statistically significant results for models based on the total population or on the total white or black population. More consistent patterns were found in the models for those over the age of 65 (Table 2). Table 2 summarizes the results of the models for morbidity by listing the number of health outcomes that had a statistically significantly higher incidence in the given Zip code than in the matching Zip codes. These results are graphically described Figure 2.

Table 2. Summary of results from morbidity models.

Zip code	County	Blacks		Whites	
		All ages	Over 65	All ages	Over 65
32570	SR	1s*	4s	2s	1s
32535/ 32565	ES	2s	3s	3s	1w
32566/ 32561	SR	-	1s	-	3s
32583/ 32530	SR	-	1w	1s	1w
32571	SR	-	1s	-	1w
32533	ES	1s	1w	-	-
32504	ES	-	2s	-	-
32501	ES	-	1w	-	-

* s indicates strong statistical evidence, w indicates weak statistical evidence.

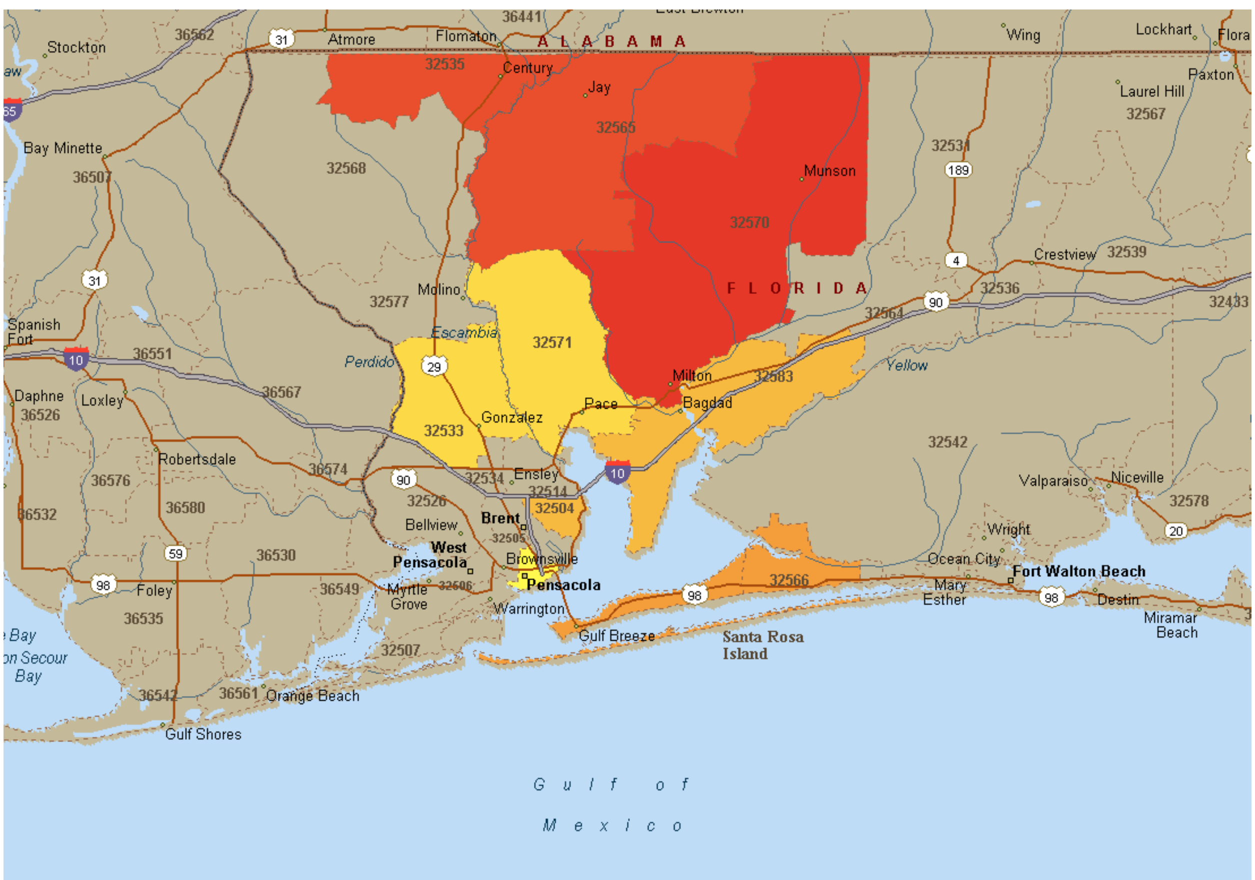


Figure 2. Map of summary of morbidity models. Darker colors indicate a greater burden of disease.

2.3 Health Tracking Conclusions

In general, health outcomes for ZIP codes in Escambia and Santa Rosa Counties differ spatially and some ZIP codes have significantly higher or lower levels of adverse health outcomes than matching ZIP codes elsewhere in Florida (see Figure 3 for example). This indicates that for some specific health outcomes and some specific Zip codes the citizen's concerns about high rates may be justified, but this health tracking study did not find evidence that the overall health of the population of Escambia and Santa Rosa Counties is significantly different from that in socio-economically and demographically comparable areas in the remainder of Florida.

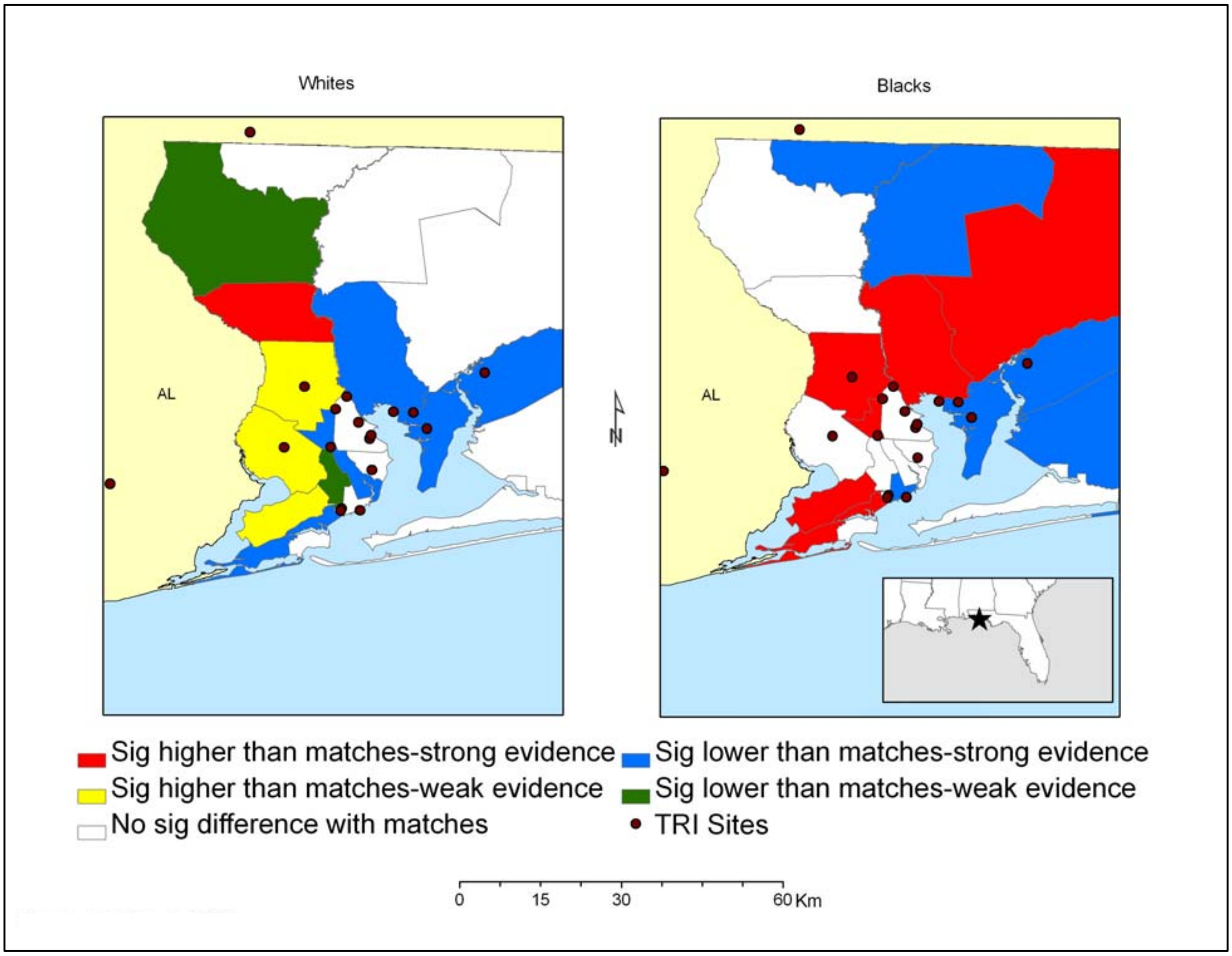


Figure 3. Comparison of birth defects in Escambia and Santa Rosa County ZIP codes and associated (matching) ZIP codes.

3 AIR EMITTER PROXIMITY STUDY

3.1 General

To evaluate if the results of the health outcomes study may be influenced by air pollution we compared the results with the geographical distribution of air emitters in the area. For this task we collected various types of spatially referenced data for air emitters. The data included: (1) Year 2000 Toxic Release Inventory (TRI) air pollution data including name, location, and emission data; (2) State of Florida permitted minor source emitters from 2002 with their name, address, and type of permit; and (3) A year 2000 dataset from the local Florida DEP office that had the location of 108 major and minor permitted air emitters in Northwest Florida but did not have emission data.

As a first step in the spatial comparison of health outcomes and air pollution data we developed a distance index that represents the proximity of a ZIP code to emission sites. Because of the large size of some ZIP codes and the heterogeneity of the population in some ZIP codes the index was initially determined for census block groups. For each block group the distance index was calculated for emission sites within 10 km from the centroid of the block group as follows:

$$\text{Proximity index for block centroid i } = \sum_{j=\#sites}^{j=1} \left(\log\left(d_{i,j}+1\right)\right)^{-1}$$

where $d_{i,j}$ is the distance from block centroid i to emission site j.

The indexes were summed by block group and the resulting total indexes for the block groups were averaged by ZIP code. Benzene-equivalents for 2002 TRI site air emissions were collected from the Environmental Defense website (http://www.scorecard.org/env-releases/def/tep_cancer.htm). The TRI total emissions and benzene-equivalent emissions were used to weight the proximity index resulting in three indexes, i.e. an unweighted index, a total emission weighted index and a benzene-equivalent weighted index. Benzene-equivalents were not available for the other two emitter data sets (State of Florida and local FL DEP), and only unweighted and/or total emission weighted indexes could be calculated for these data sets.

The average proximity indexes for each ZIP code were statistically compared to cumulative health outcomes and to specific health outcomes. In both cases, ZIP codes within NW FL were compared with each other and with the socio-economically and demographically matching ZIP codes elsewhere in the state. The hypothesis of this study was that if morbidity and/or mortality in the area are linked to air pollution, Zip codes with worse health outcomes than their matching Zip codes should be located closer to air emission sites.

3.2 *Cumulative* Health Outcomes

In this part of the study the proximity indexes were evaluated for two groups of NW FL ZIP codes: ZIP codes identified by the health tracking study as having *cumulative* evidence for better health outcomes than their respective matching ZIP codes and ZIP codes having *cumulative* evidence for worse health outcomes.

The ZIP codes in NW FL with worse (resp. better) cumulative health outcomes do not systematically have higher (resp. lower) proximity indexes than their associated ZIP codes (e.g. Figure 4a and 4b). Difference of means testing showed a significant difference ($P < 0.05$) between NW FL and associated ZIP codes only for the benzene weighted TRI proximity index and worse outcomes (Figure 4b), but not for the other proximity indexes. Consequently, these results for *cumulative* health outcomes do not show a strong relationship between proximity to emission sites and health outcomes in the study area.

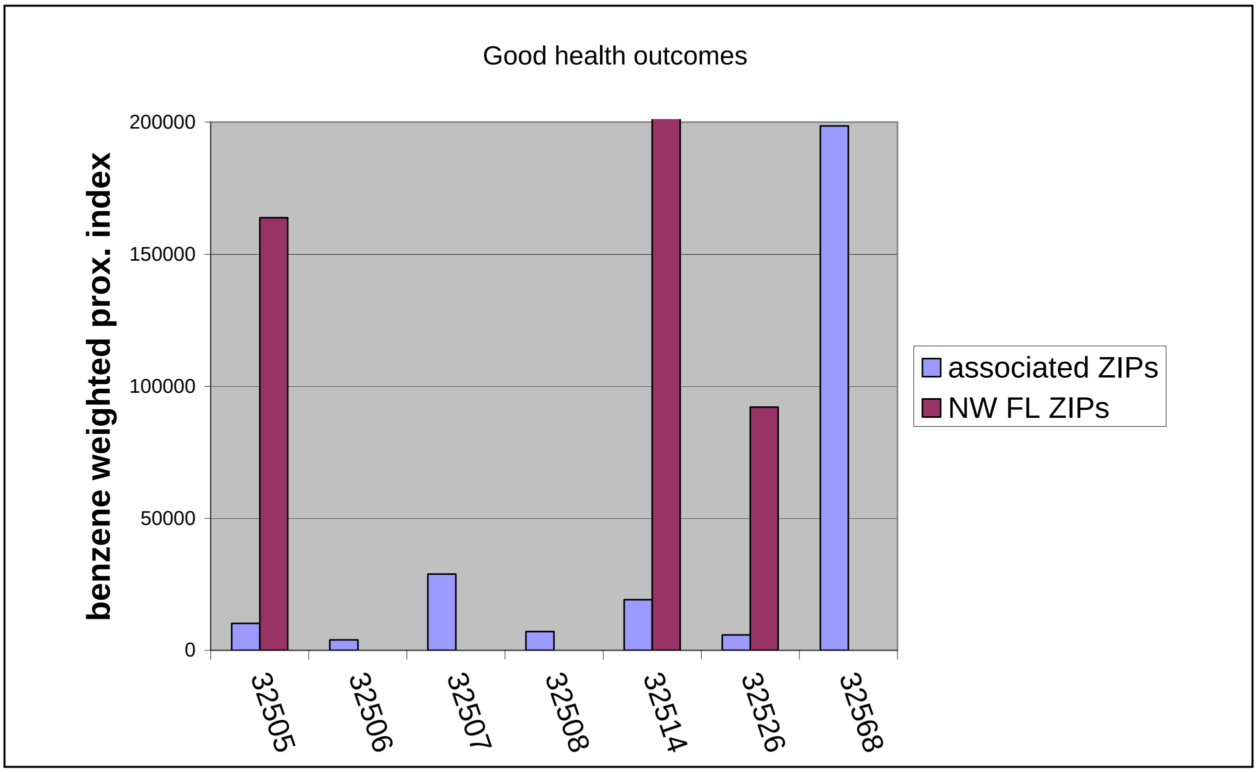


Figure 4a. Benzene weighted TRI proximity index for NW FL ZIP codes with better cumulative health outcomes vs. associated ZIP codes.

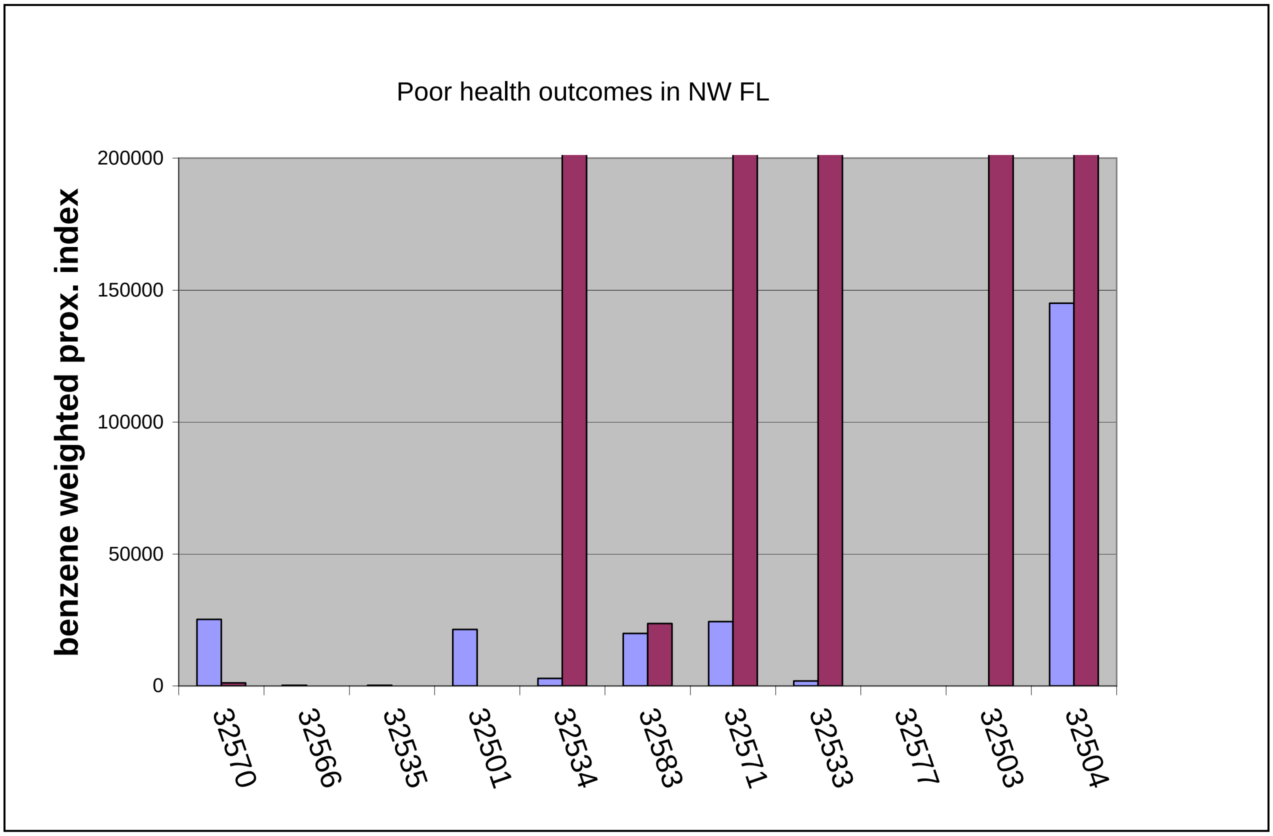


Figure 4b. Benzene weighted TRI proximity index for NW FL ZIP codes with worse cumulative health outcomes vs. associated ZIP codes. See Figure 4a for legend.

3.3 *Specific* Health Outcomes

ZIP codes with a high incidence of some *specific* health outcomes have a higher proximity index than ZIP codes with a low incidence (Figure 5a, b), suggesting that there is a link between proximity to emission sites and the incidence of these specific health outcomes. The specific health outcomes for which this relationship holds true varies only slightly between the various proximity indexes and are:

- Mortality: white, >65, cardiac; black, >65, lung cancer; black, birth defects
- Morbidity: black, all ages, asthma; black, >65, cardiac; white, >65, pneumonia

For these 6 specific health outcomes NW FL ZIP codes were also compared to their respective associated ZIP codes. Graphs shows that NW FL Zip codes with a higher incidence of these 6 specific health outcomes have a higher proximity index than their matching Zip codes and that NW FL Zip codes with a lower incidence have a lower proximity index than their matching Zip codes (Figure 6a, b). This observation corroborates the contention that a link exists between the proximity to emission sites and the incidence of these specific health outcomes. Weaker indications for this link can be observed for mortality in blacks, >65, due to all cancers and morbidity in blacks, >65, due to respiratory illnesses (Figure 5a, b; Figure 6a, b). Statistical analysis were not performed because of the low number of cases (ZIP codes) in each category. These observations for *specific* health outcomes, both within NW FL and compared to matching

Zip codes, suggest that there is an influence of proximity to emission sites on the incidence of some of the specific health outcomes. These observations are preliminary and have to be confirmed by other analysis using more robust statistics, but they suggest that such further evaluation is warranted.

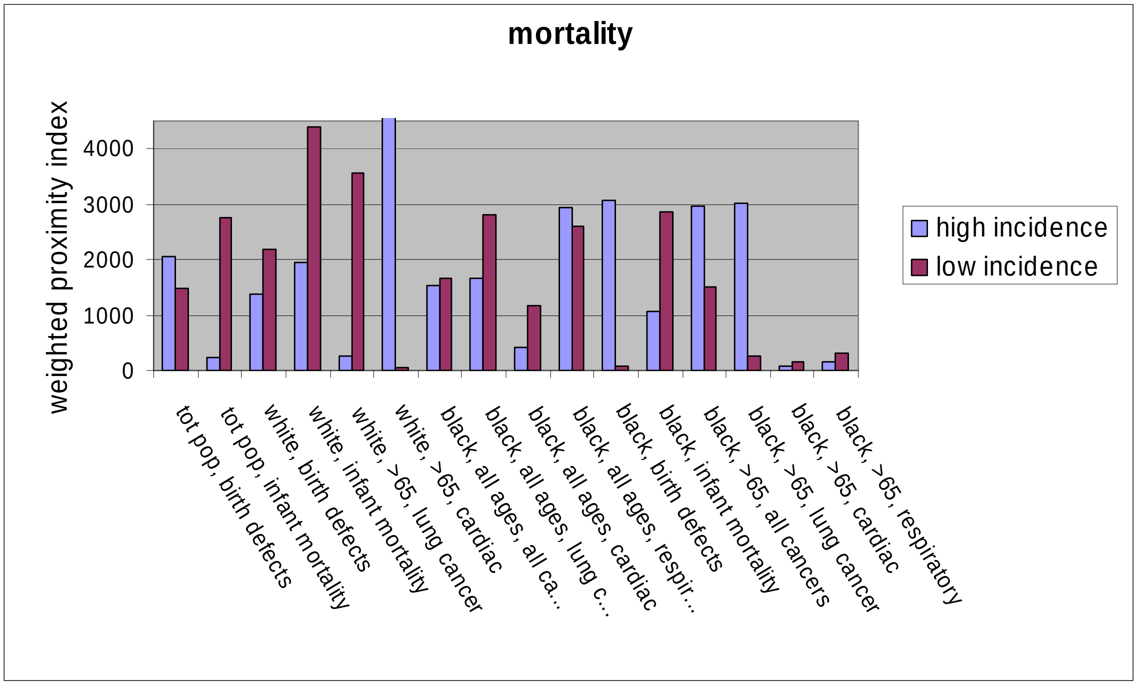


Figure 5a. Proximity index for Zip codes with high or low incidence for specific causes of mortality: Comparison within northwest Florida. Based on statewide database.

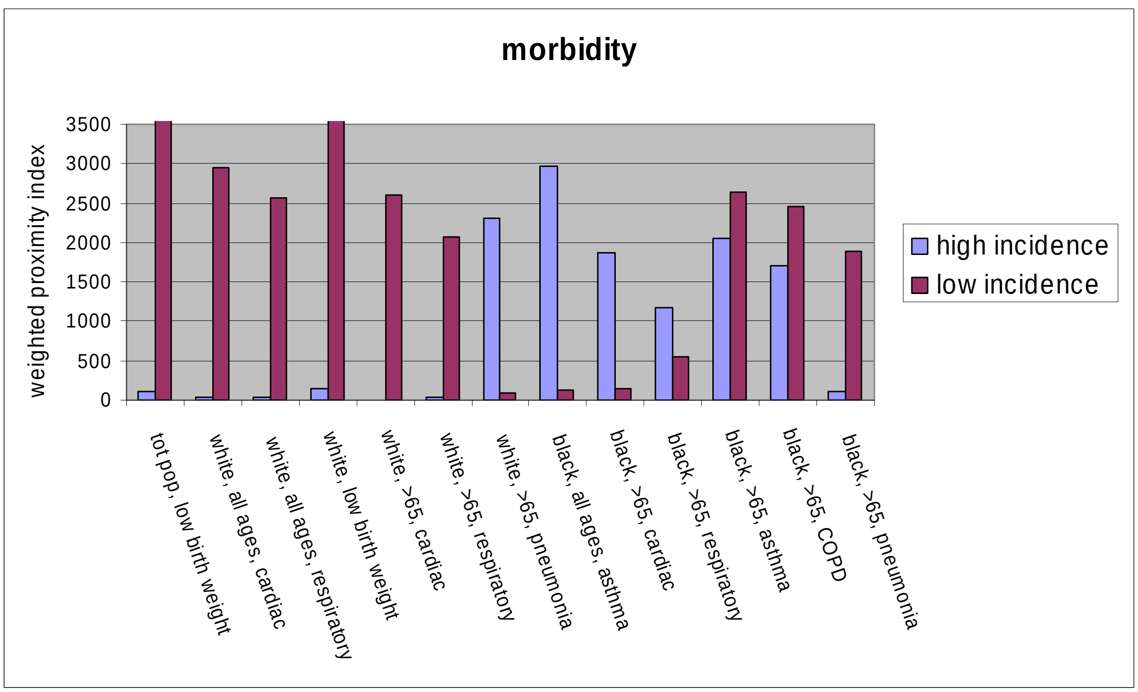


Figure 5b. Proximity index for Zip codes with high or low incidence for specific causes of morbidity: Comparison within northwest Florida. Based on statewide database.

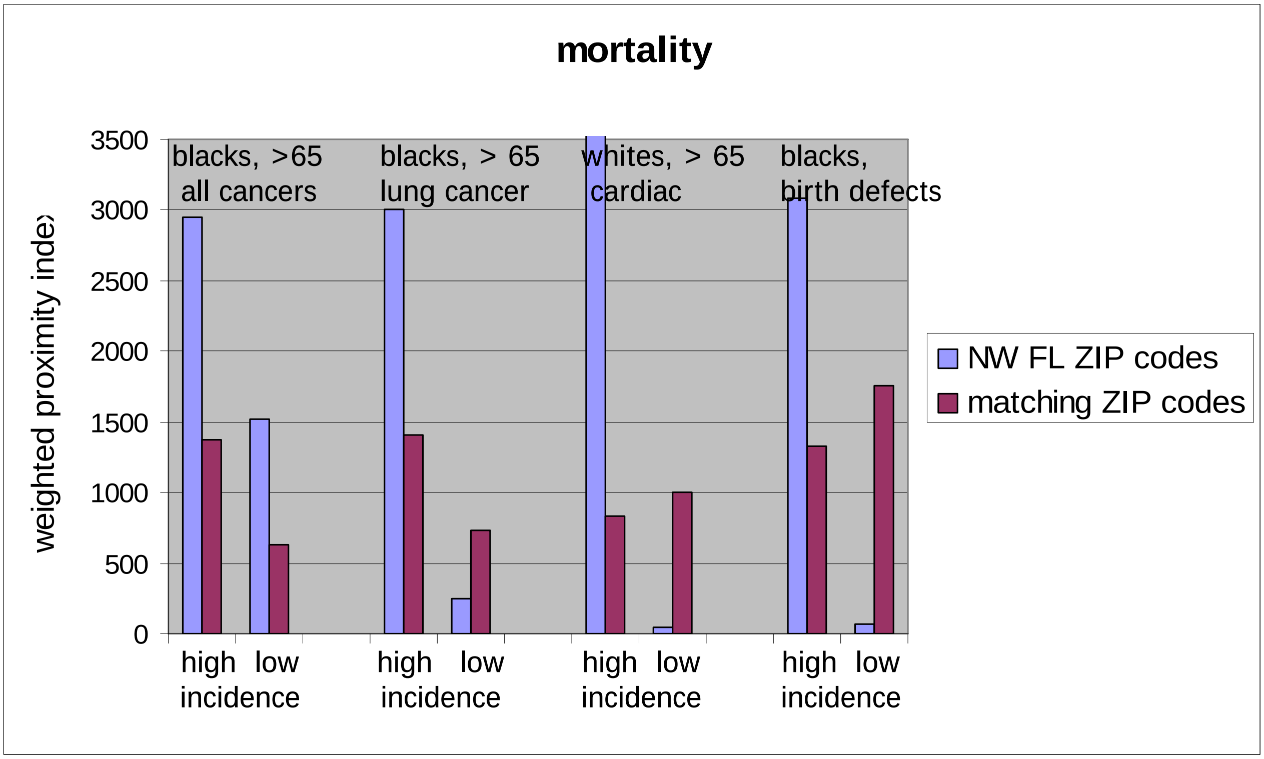


Figure 6a. Proximity index for Zip codes with high or low incidence for specific causes of mortality: Comparison with matching Zip codes. Based on statewide database.

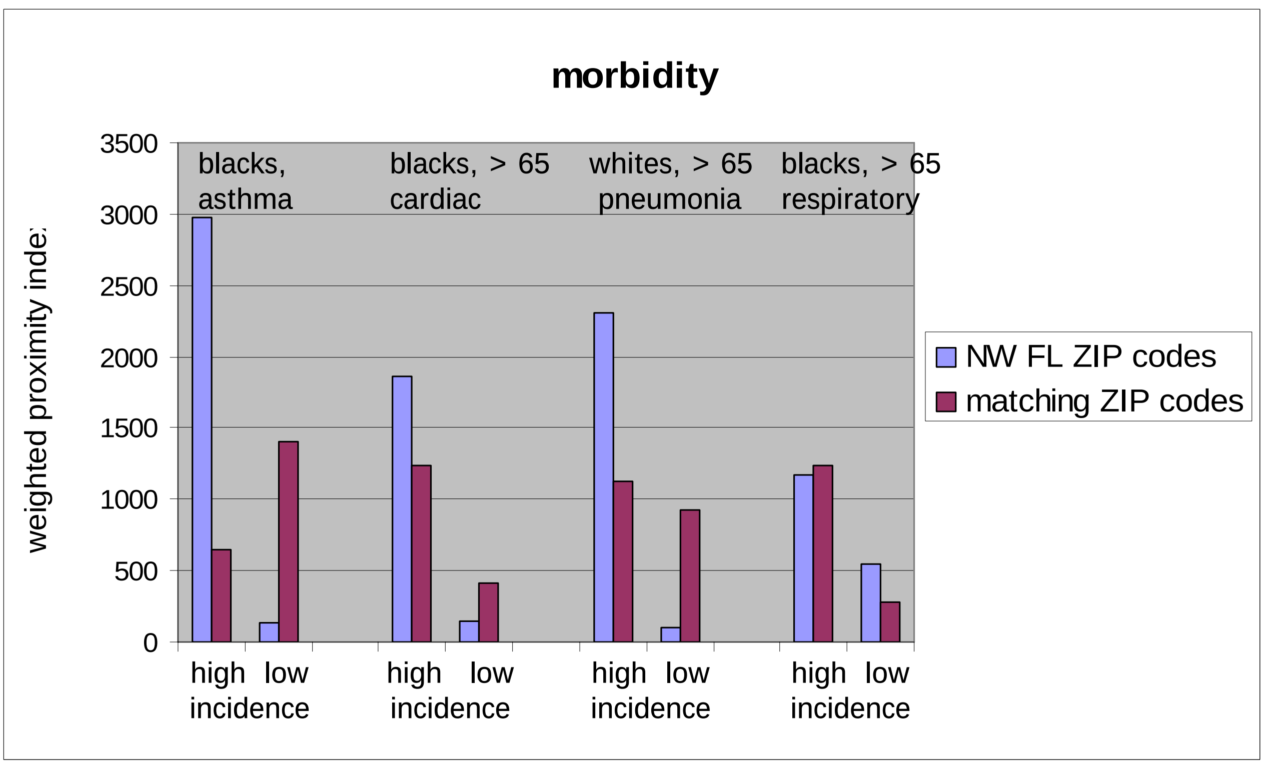


Figure 6b. Proximity index for Zip codes with high or low incidence for specific causes of morbidity: Comparison with matching Zip codes. Based on statewide database.

3.4 Environmental Inequity

To assess environmental equity the TRI-based proximity indexes were averaged at the census tract level and statistically correlated with population density, percent whites, percent non-white, poverty rate, industrial employment, and educational attainment. The unweighted proximity index showed moderately strong and significant correlation with population density and industrial employment. The two emission weighted indexes did not have a strong correlation with the demographic and socio-economic variables. These observations indicate that a spatial relationship exists between the location of TRI sites and two of the variables (i.e. population density and industrial employment) but this does not seem to lead to greater exposure of any racial group to emissions.

To further explore potential connections between proximity to emission sites and the demographic and socio-economic variables multiple-linear-regressions were run. The R^2 value was very low for the two weighted TRI proximity indexes (0.09) but somewhat higher (0.24) for the unweighted proximity index. Non-linear models yielded comparable R^2 values. Population density and industrial employment had a statistically significant effect in the regression for the unweighted proximity index. These results are consistent with the results for the correlation coefficients and also fail to show evidence for environmental inequity in exposure to TRI site emissions.

3.5 Proximity Study Conclusions

The air emitter proximity study did not find clear evidence for an influence of proximity to emission sites on ZIP code level *cumulative* health outcomes. Some of the *specific* health outcomes are directly related to proximity to emission sites as evidenced by relationships at the Zip code level within NW Florida and comparisons between NW Florida and similar areas elsewhere in the state. Preliminary statistical analysis at the census tract level of the proximity indexes and demographic and socio-economic data does not indicate environmental inequity in exposure to emission sites in NW Florida. PERCH considered this proximity study only a first step in evaluating relationships between health outcomes and air pollution in Escambia and Santa Rosa Counties because it was hampered by the unavoidable use of spatial units (Zip codes) that are not ideally suited for this type of analysis. Therefore, PERCH further assessed connections between air pollution and health outcomes with two other approaches, air toxics modeling and raster-based geostatistical modeling.

4 AIR TOXICS MODELING

4.1 Introduction

As part of its effort to assess all aspects of the environment in Escambia and Santa Rosa Counties PERCH carried out a comprehensive evaluation of air pollution in the area, conducted by PERCH project collaborators from the Georgia Institute of Technology (http://www.uwf.edu/CEDB/Perch_Air_Quality_Studies.cfm). Based on a review of ambient monitoring data, available information regarding emissions, other studies, and discussions with various stakeholders, three classes of air pollutants were found to be of particular concern in the area: ground level ozone, fine particulate matter (PM2.5), and air toxics. Health cost estimates associated with these pollutants indicated that PM2.5 imposed the highest per person per year costs, followed by ozone and then air toxics. Particulate matter likely presents the greatest air quality risk to human health in the region. Sulfate is a large fraction of the observed ambient PM2.5 loading, with high concentrations most often associated with northerly air flow. Additionally, organic carbon was found to be a large fraction of the ambient PM2.5 loading. There is community concern regarding air toxics based on TRI discharges from point sources. To further try to relate observations of the health tracking study to air quality measures a multipronged modeling and analysis approach of air toxics was performed. This air toxics modeling study assessed the health risk associated with stationary point sources, traffic sources, and one specific emitter (Gulf Power's Plant Crist power plant). The primary objective of the modeling was to quantify risk levels in the two-county study area to advise policy- and decision-makers about the existence of elevated risks for adverse health outcomes due to toxic air pollution. For its modeling, the study employed the Regional Air Impact Modeling Initiative (RAIMI) system, which consists of a set of tools designed to evaluate the potential for health impacts as a result of exposure to multiple contaminants from multiple sources, at a community level of resolution. RAIMI integrates emission data, meteorological data, a dispersion model, and risk estimation in a GIS environment and allows estimation and representation of cancer and non-cancer risks via inhalation.

4.2 Results

4.2.1 Stationary Point Sources

Application of the RAIMI system with the 1999 National Emission Inventory (NEI) for the study area indicated four concentrated hotspots of potentially elevated cancer risk related to point sources (Figure 7).

Risk Zone 1 – Northern Santa Rosa County. Risk Zone 1 is in northern Santa Rosa County in the vicinity of three emission sources: a petroleum/natural gas extraction operation, a natural gas pipeline compressor station, and a landfill. A maximum cumulative risk of 48 in a million was predicted by RAIMI. The peak risk was attributed almost entirely to formaldehyde emissions from the natural gas compressor station. This risk zone overlaps mostly with Zip code 32531, in which the health study did not find significant differences with the matching Zip codes.

Risk Zone 2 – Northern Santa Rosa County. Risk Zone 2 is also in northern Santa Rosa County in the vicinity of two emission sources: a petroleum/natural gas extraction operation and a landfill. A maximum cumulative risk of 23 in a million was predicted by RAIMI. The peak risk was attributed almost entirely to formaldehyde and toluene emissions from the petroleum/natural gas extraction operation.. This risk zone is located in a Zip code (32565) that was merged with another one in Escambia County (32535) by the health study due to low population numbers. This large combined Zip code was among the worst in the morbidity models of the health study, but all of these models were for non-cancer related morbidity. This suggests that the spatial overlap of the risk zone with the Zip code with poor health outcomes is coincidental.

Risk Zone 3 – Pace Community in Santa Rosa County. Risk Zone 3 is near the Pace community in Santa Rosa County in the vicinity of six emission sources: four industrial plants and two landfills. A maximum cumulative risk of 709 in a million was predicted by RAIMI. The peak risk was attributed almost entirely to acrylonitrile emissions from the acrylic fiber manufacturing operation. This risk zones partially overlaps with Zip codes 32571 and 32583, which have worse morbidity than their matching Zip codes for some non-cancer health outcomes but not for cancers, indicating again that a causal relationship between risk and observed health outcomes can not be demonstrated.

Risk Zone 4 – Cantonment Community in Escambia County. Risk Zone 4 is near Cantonment in Escambia County about 10 km northwest of Downtown Pensacola in the vicinity of a large pulp and paper manufacturing operation. A maximum cumulative risk of 5.4 in a million was mostly attributed to methanol, acetaldehyde, benzene and xylene, which are used as chemical solvents in the pulping operation. This risk zone partly overlaps with Zip codes 32533 and 32534. Zip code 32534 has worse rates than its matching Zip codes for some cancer-related causes of mortality in African Americans. Both Zip codes have worse mortality due to birth defects.

Results for non-cancer risk from stationary point sources showed two very small areas with elevated non-cancer risk on the premises of the emitting facilities (Solutia Inc. and Sterling Fibers). Given their very small size, these risk zones do not provide an explanation for the findings of the health tracking study.

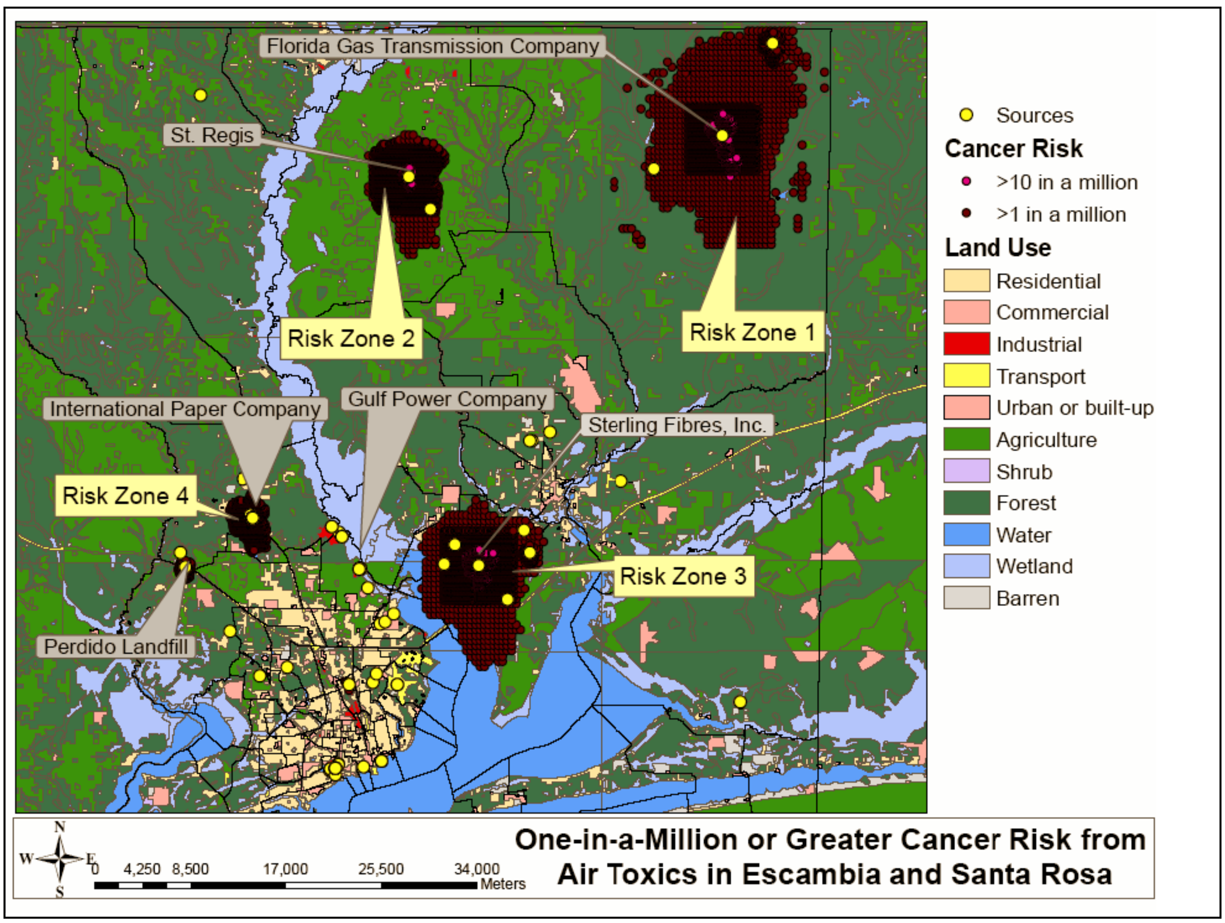


Figure 7. Elevated chronic cancer risk zones estimated by the RAIMI system.

4.2.2 Traffic Sources

Almost all the regions around the modeled roads in both counties are subject to a cancer risk of 1 in a million or greater. Large parts of Escambia County and a few regions close to main roadways in Santa Rosa County are subject to 10 in a million greater cancer risk. Many parts of urban Escambia are subject to estimated cancer risks of more than 100 in a million. In Santa Rosa, 100 in a million or greater cancer risk is mainly concentrated along Interstate 10 and US 98 roadways. As in the case of cancer risks, almost all locations are subject to a hazard index (HI) of more than one for non-cancer risks. Higher values of HI (10-100 range) are concentrated in the urbanized areas of Escambia and along I-10 and US 98 in Santa Rosa. The risks diminish by several orders of magnitude a few hundred meters off the roadway. Given the difference in spatial units, these results can not be compared to those for the health tracking study.

4.2.3 Acute Health Risks From Emissions From Crist Plant

One of the emitters in the study area that is often referred to in informal conversations among citizens and in the media is the highly visible Crist Plant, a coal-fired power plant. All HCl (7,559 tons) and HF (153 tons) emissions in the two-county study area are reported to come from this facility. Though these compounds are not carcinogenic, PERCH assessed if the sizable emissions could be a source of short-term health risks. For assessing the acute health risks related to the significant emissions of HCl and HF from Plant Crist, PERCH used an approach developed by the Georgia Environmental Protection Division (EPD). The Georgia EPD approach is convenient for modeling applications because it provides, with an appropriate margin of safety, a conversion of occupational exposure safety thresholds (typically 8 hours) into more relevant averaging periods (e.g., 24 hours) for assumption of continuous exposures. Using risk-based criteria an acceptable ambient concentration (AAC) can be calculated for each pollutant to represent acceptable risk levels for acute (15-minute and 24-hour average) time periods as well as chronic exposures (annual average). The modeling protocol followed a similar approach as the cancer risk assessment in RAIMI by using the same dispersion model (ISCST3) and meteorological data set. Results of the modeling show that HCl and HF ambient impacts from Plant Crist are 90 to 98% below the risk-based AAC, and thus they do not appear to present a significant acute health risk via inhalation.

4.3 Air Toxics Modeling Conclusions

Air toxics modeling identified four zones in the study area with elevated cancer risk due to inhalation of air emissions from stationary point sources. Three of these risk zones overlap with Zip codes that were not found by the health tracking study to have significantly elevated cancer rates, compared to matching Zip codes elsewhere in the state. It is noteworthy that Zip code 32570, which is worse than its matching Zip codes for several cancer related health outcomes is surrounded by but not covered by these three cancer risk zones. The fourth risk zone, just north of Pensacola, partially overlaps with Zip codes where the health tracking study found comparatively high rates of some cancers in some sections of the population. Further detailed study is required to determine if this spatial association of risk and health outcome is causal.

5 GEOSTATISTICAL STUDY

5.1 Introduction

This geostatistical component of PERCH investigated the associations between air pollution and health outcomes in Escambia and Santa Rosa counties using three approaches: mapping spatial patterns of health outcomes and air pollution, exploratory statistical analyses, and statistical modeling. The project was carried out progressively from simple mapping and modeling, to more refined mapping using satellite imagery and advanced spatially extended modeling.

5.2 Data

Health outcome data were obtained from the University of South Florida CATCH (Comprehensive Assessment for Tracking Community Health) data warehouse, the Florida Department of Health CHARTS (Community Health Assessment Resource Tool Set) database, and the US Centers for Disease Control and Prevention (CDC) WONDER (Wide-ranging Online Data for Epidemiologic Research) records. The downloaded health outcome data at the Zip code level included hospitalizations due to cardiac, respiratory problems, COPD, asthma, low birth weight, and mortality due to all cancers, lung cancer, cardiac, respiratory problems, birth defects, and infant death. Finer resolution health outcome data at the census tract level included mortalities due to COPD, stroke, and lung cancer. Data at the county level included myocardial infarction (MI) and chronic coronary heart disease (CCHD).

Socio-economic/demographic data at the census tract level for total population, male, female, white, black, Asian, Hispanic, 2+ races, 65 and 65+ years old, poverty status, income below poverty, and median household income were extracted from the 2000 Census Summary File.

Point source air pollution data (Figure 8) were collected for US EPA Toxic Release Inventory (TRI) sites, Superfund sites, and Florida Department of Environmental Protection monitored solid waste sites, sewer treatment sites, and brown field sites. For mobile source data traffic counts with emission estimates were obtained from the Florida Department of Transportation. Density surfaces were derived from all point and mobile source air quality data.

A Normalized Difference Vegetation Index (NDVI) raster surface and a greenness surface representing the amount of green space were calculated from a cloud-free Landsat 7 Enhanced Thematic Mapper Plus (ETM+) imagery. Moderate Resolution Imaging Spectrometer (MODIS) daily level 2 (2003-2004) aerosol optical depth (AOD) data in Hierarchical Data Format (HDF) were obtained from the NASA Level 1 and Atmosphere Archive and Distribution System (LAADS Web) at http://ladsweb.nascom.nasa.gov/. MODIS Level 2 data are produced at the spatial resolution of a 10×10 1-km (at nadir)-pixel array.

PM2.5 ground data was obtained from the EPA Air Quality System (AQS) online Data Mart at http://www.epa.gov/ttn/airs/aqsdatamart/index.htm. PM2.5 values measured within 1 hour of the MODIS imaging time were retrieved for the year 2004. Annual statistical summary PM2.5 data for 2003 and 2004 were also obtained for monitoring sites covering the conterminous land. Calculation of annual statistics included exceptional air pollution events.

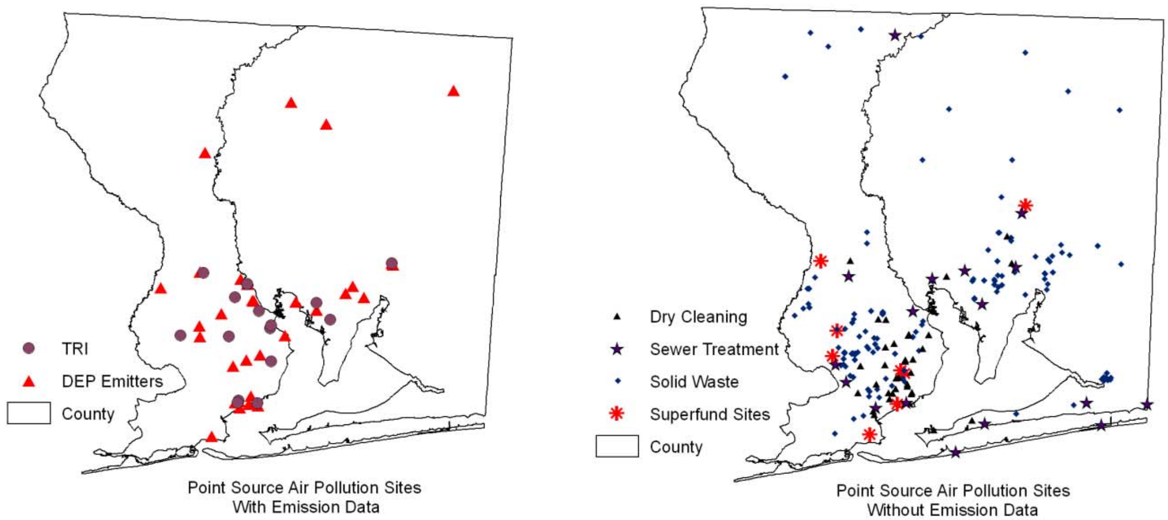


Figure 8. Maps of point source polluters.

5.3 Mapping, Analyses And Modeling Results

5.3.1 Mapping

At the census tract level, three characteristic patterns stand out from environmental exposure maps (Figure 9): (1) Zip code 32514 (largest circle in Figure 9b), where Gulf Power's Crist Plant is located, has the largest emissions (total air or benzene-equivalent) but has low mortality rates (Figures 9a, 9b); (2) all other point source pollutants are concentrated within the urban extent (Figures 9c, 9d); and (3) traffic volume has the highest density in the City of Pensacola, along Interstate 10, and highways 90, 98 and 29 (Figures 9e, 9f).

5.3.2 Exploratory Spatial Data Analysis (ESDA)

A parallel coordinate plot was created to show values for Mortality Respiratory Total Population (MTPR) and six environmental exposure variables (TRI total air emission, TRI benzene-equivalent emission, DEP emitters emission, all other point sources, traffic density, and NDVI). Again, Zip code 32514 area stands out as highlighted in the plot and the map due to its high emission values.

A Moran's I plot was generated for total population respiratory hospital admission rates. The Moran's I value is 0.47, indicating strong spatial autocorrelation. Zip codes 32577 and 32568 have much lower respiratory hospital admission rates than their neighbors.

5.3.3 Spatial Lag Model Of Asthma And Air Pollution

The model shows relationships between the total population asthma hospitalization rate and air pollution (positive, $0.08 < p < 0.26$) and 'greenness' (negative, $p = 0.23$) (Table 3).

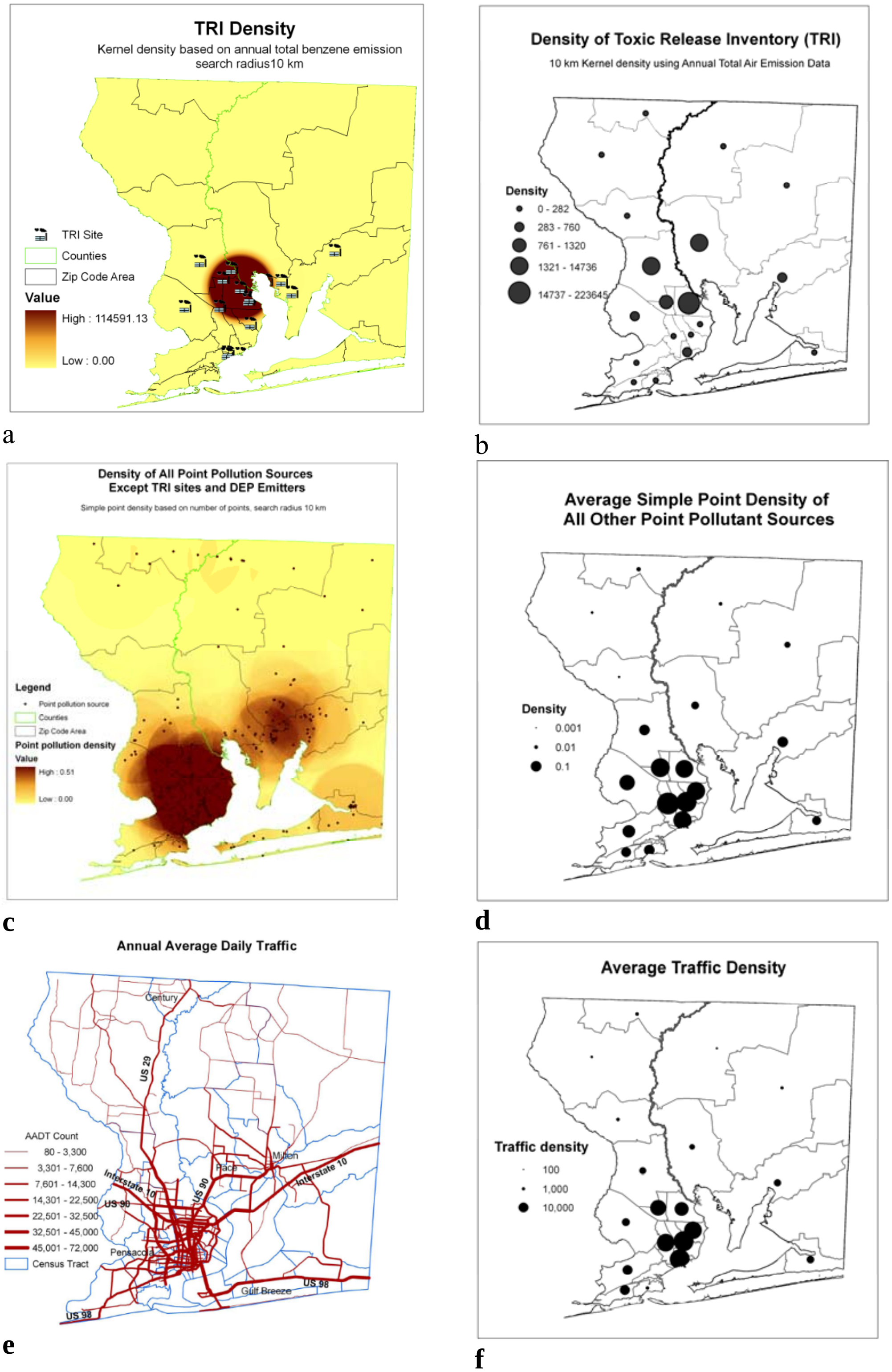


Figure 9. Zip-code level environmental exposure maps.

Table 3. Spatial lag model of asthma and air pollution.

Variable	Coefficient	Probability
ρ	0.3645	0.087
Constant	0.7242	0.019
Traffic	1.6192e-005	0.110
Greenness	-0.221	0.230
TRI benzene	8.6748e-006	0.256
TRI Total	1.6497e-006	0.147
DEPEMIT	0.0021	0.085
Other point pollution	1.2563	0.115

5.3.4 Mortality Rates Of COPD, Stroke, And Lung Cancer Compared With Socio-Economic And Environmental Factors

5.3.4.1 Focused Score Tests

Maps reveal linkages between some of these health outcomes and the environmental and socio-demographic factors. High COPD mortality tends to occur in areas with high poverty rates. Scatter plots show relationships between death rates and suspected factors. Mortality rates of all three diseases show positive relationships with proportion blacks, population age 65 and above, and the poor, as well as mobile and point source air pollution, and negative relationships with median household income, percentage males, and greenness.

Table 4 shows the results of focused score tests. The table indicates significant focused clustering of the deaths of COPD, stroke, and lung cancer around traffic pollution and point source pollution. The amount of greenness does not show significant relationships with the deaths.

Table 4. Focused score tests of relationships between mortality rates of COPD, stroke, and lung cancer with socio-economic and environmental factors.

Disease	Foci	T^*_{sc}	p value
COPD	Traffic	3.3415	0.00042
	Point source	3.1594	0.00079
	Greenness	0.2946	0.38414
Stroke	Traffic	13.6522	<0.0001
	Point source	9.0468	<0.0001
	Greenness	-7.6334	1.00000
Lung Cancer	Traffic	4.8486	<0.0001
	Point source	3.4729	0.00026
	Greenness	0.2577	0.39834

5.3.4.2 OLS Regression Analyses

Table 5 shows the OLS univariate regression results. Higher COPD, stroke, and lung cancer mortality rates occurred in census tracts with lower median household income, percentage of male population, and greenness. However, the relationship between stroke and percent male and the relationships between all the three health outcomes and greenness are not significant. All the mortality rates are significantly positively associated with proportion blacks, people age 65 or above, poverty rate, and air pollution from both mobile and point sources. Regression of COPD has the lowest fits with R-square values ranging from 0.01 (greenness) to 0.15 (poverty rate). The models fit best for stroke with R-square values between 0.02 (percent males) and 0.56 (percent blacks). AADT explains 37.84% of the variability in stroke death rate. For lung cancer, the proportion of people age 65 and above has the largest R-square value (0.23).

Tables 6-8 show OLS multivariate regression results for each disease. At the level of 0.10, COPD is significantly positively associated with air pollution from both mobile and point sources, and negatively related to median household income and percent of male population. Stroke shows significant positive relationship with population with age 65 or above and air pollution, and negative relationship with median household income. The same relationships are found for lung cancer death rate. Lung cancer also shows a negative relationship with the percent male population. The stroke model fits best (R^2 = 0.53). The lung cancer model has an R^2 value of 0.49. The COPD model has the weakest fit (R^2 = 0.26).

Table 5. Univariate OLS regression of disease rates on suspected factors.

Dependent variable	Independent variable	Slope	*p*-value	*R*-square
COPD	Income	-1.58e-008	0.00450	0.1026
	Black	0.00075	0.00690	0.0932
	Population >65	0.00400	0.00214	0.1187
	Male	-643.033	0.00305	0.1111
	Poor	0.002226	0.00047	0.1513
	Traffic	1.16e-008	0.00323	0.1099
	Greenness	-5.89e-006	0.43225	0.0082
	Point source	0.003351	0.00525	0.0993
Stroke	Income	-1.29E-007	<0.0001	0.2470
	Black	0.00966	0.00000	0.5558
	Population >65	0.03265	<0.0001	0.2830
	Male	-0.0063	0.26835	0.0163
	Poor	0.01607	<0.0001	0.2821
	Traffic	1.136E-007	<0.0001	0.3784
	Greenness	-0.00013	0.00045	0.1522
	Point source	0.01721	0.00680	0.0936
Lung cancer	Income	-3.894e-008	<0.0001	0.1935
	Black	0.00164	0.00085	0.1386
	Population >65	0.00994	<0.0001	0.2272
	Male	-0.00738	<0.0001	0.1910
	Poor	0.00457	<0.0001	0.1975
	Traffic	2.136e-008	0.00247	0.1158
	Greenness	-5.885e-006	0.66302	0.0025
	Point source	0.00650	0.00247	0.1157

Table 6. Multivariate OLS regression for COPD.

```
Dependent Variable  :    COPD rate
R-squared           :    0.258502  F-statistic          :     4.95042
Adjusted R-squared  :    0.206284  Prob(F-statistic)    : 0.000605213
Sum squared residual:1.75527e-005  Log likelihood       :     479.565
Sigma-square        :2.47221e-007  Akaike info criterion :     -947.13
S.E. of regression  : 0.000497214  Schwarz criterion    :    -933.067
Sigma-square ML     :2.27957e-007
S.E of regression ML: 0.000477449
----------------------------------------------------------------------
Variable                 Coefficient     Probability
----------------------------------------------------------------------
CONSTANT                  0.003582364      0.0000160
Income                   -1.599449e-08     0.0188372
Population >65            0.00205754       0.1554837
MALE                     -0.001912757      0.0997262
Traffic                   5.886609e-09     0.0057732
Point source              0.002795998      0.0682218
```

Table 7. Multivariate OLS regression for stroke.

Dependent Variable	**Stroke rate**		
R-squared	0.533386	F-statistic	16.232
Adjusted R-squared	0.500526	Prob(F-statistic)	1.19759e-010
Sum squared residual	0.000308776	Log likelihood	369.17
Sigma-square	4.34896e-006	Akaike info criterion	-726.339
S.E. of regression	0.00208542	Schwarz criterion	-712.276
Sigma-square ML	4.01008e-006		
S.E of regression ML	0.00200252		

Variable	**Coefficient**	**Probability**
CONSTANT	0.002217874	0.0999130
Income	-5.678549e-08	0.0442312
Population >65	0.02213693	0.0001146
Traffic	6.209707e-08	0.0353378
Green	-4.30978e-05	0.1851318
Point source	0.004164622	0.0524035

Table 8. Multivariate OLS regression for lung cancer.

Dependent Variable	**Lung cancer rate**		
R-squared	0.493561	F-statistic	13.8389
Adjusted R-squared	0.457896	Prob(F-statistic)	1.96394e-009
Sum squared residual	3.872e-005	Log likelihood	449.106
Sigma-square	5.45352e-007	Akaike info criterion	-886.211
S.E. of regression	0.00073848	Schwarz criterion	-872.148
Sigma-square ML	5.02857e-007		
S.E of regression ML	0.000709124		

Variable	**Coefficient**	**Probability**
CONSTANT	0.006729455	0.0000001
Income	-4.961127e-08	0.0000037
Population >65	0.006191974	0.0048375
MALE	-0.005135011	0.0035637
Traffic	2.799079e-08	0.0041351
Point source	0.006985843	0.0026527

5.3.5 Bayesian Hierarchical Modeling Of Stroke Mortality And Air Pollution, Income, And Greenness

An ecological geographical approach was adopted using census tract level stroke data and a Bayesian hierarchical model. The mean age-adjusted stroke death rates were 8.39 times the average age-adjusted stroke rate in the US South. Table 9 provides the estimated posterior mean, median, and associated 95% credible set for each of the fixed effects. Figure 10 provides kernel estimates of the corresponding posterior densities. Table 9 and Figure 10 reveal strong negative effects of income and greenness (the posterior densities of β_1 and β_5 primarily covers negative values) and positive effects of both mobile and point source air pollution (the 95% credible sets cover positive values). High risk of stroke mortality was found in areas with low income level, high air pollution level, and low level of exposure to green space.

Table 9 Markov chain Monte Carlo results for Bayesian hierarchical modelling of stroke mortality vs. income, air pollution, and greenness.*

Fixed Effects	Posterior Mean	Posterior Median	Standard Deviation	MC Error	95% Credible Set
β_0	1.829	1.832	0.083	0.004	(1.661, 1.986)
β_1	-0.193	-0.193	0.047	0.003	(-0.286, -0.097)
β_2	0.089	0.089	0.028	0.001	(0.034, 0.144)
β_3	0.937	0.932	0.276	0.010	(0.419, 1.495)
β_4	0.974	0.980	0.290	0.012	(0.413, 1.522)
β_5	-0.161	-0.161	0.067	0.002	(-0.289,-0.031)

* Posterior means, medians, and 95% credible sets are based on 5,000 postconvergence iterations (from 5,001 to 10,000). Fixed effects are: β_0 - intercept, β_1 - income effect, β_2 -traffic air pollution effect, β_3 - effect of EPA and Florida DEP monitored point source air emission, β_4 - effect of non-monitored point source air pollution, and β_5 - greenness.

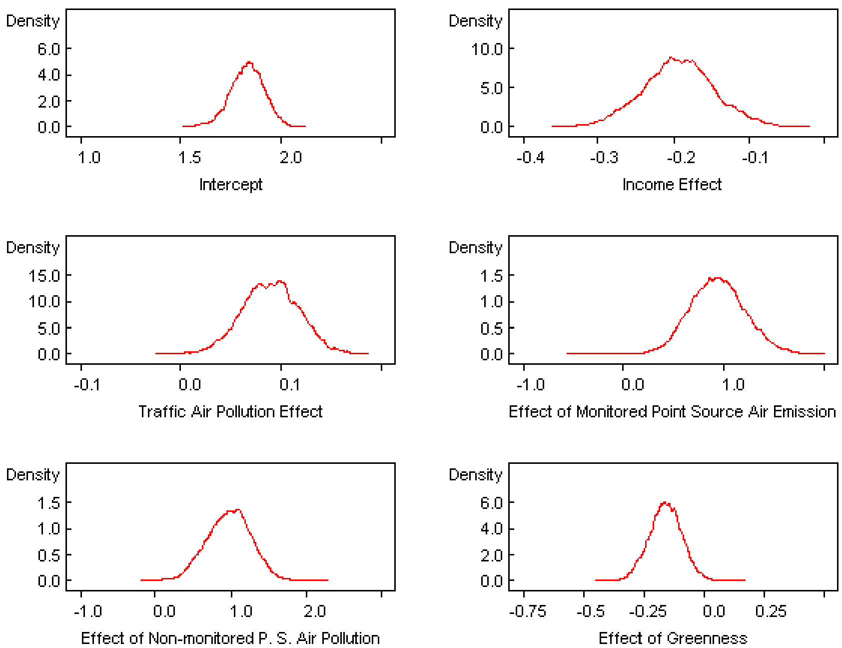


Figure 10. Kernel estimates of the posterior density of the fixed effects in the Bayesian hierarchical model.

5.3.6 Correlating MODIS Aerosol Optical Depth Data With Ground-Based PM2.5 Observations

Years 2003 and 2004 daily MODIS Level 2 AOD (aerosol optical depth) images were collated with US EPA PM2.5 data covering the conterminous USA. Pearson's correlation analysis and geographically weighted regression (GWR) found that the relationship between PM2.5 and AOD is not spatially consistent across the conterminous states. The average correlation is 0.67 in the east and 0.22 in the west. GWR predicts well in the east and poorly in the west. The GWR model was used to derive a PM2.5 grid surface (Figure 11) using the mean AOD raster calculated using the daily AOD data (RMSE = 1.67 $\mu g/m^3$).

There are 4 PM monitors in Escambia and Santa Rosa counties. Table 10 shows EPA monitor source IDs, coordinates, correlations between PM2.5 and MODIS AOD, GWR R squares, and GWR constants and AOD coefficients. Figures 12 and 13 show the correlation surface and R squares of geographically weighted regression of PM2.5 on MODIS AOD.

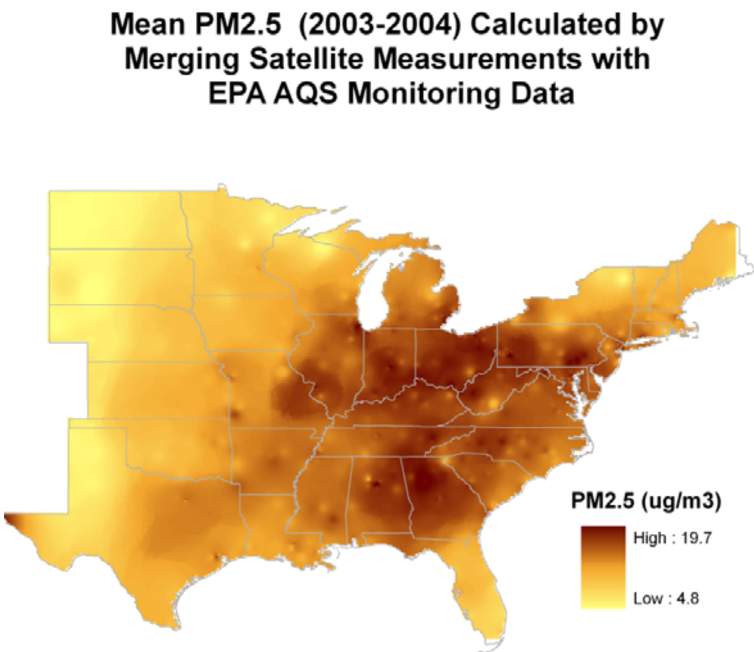


Figure 11. PM2.5 surface calculated by merging MODIS AOD and EPA PM2.5 ground measurements.

Table 10. Correlation and geographically weighted regression results for the four PM monitoring sites in the study area.

EPA Source ID	Latitude	Longitude	Correlation	GWR R^2	GWR AOD Coefficient	GWR Constant
120330026881011	30.550	-87.376	0.738	0.53	0.023	9.374
120330004881011	30.525	-87.204	0.705	0.509	0.023	9.424
120330025881011	30.437	-87.256	0.647	0.507	0.023	9.386
121130014881011	30.408	-86.890	0.691	0.519	0.023	9.374

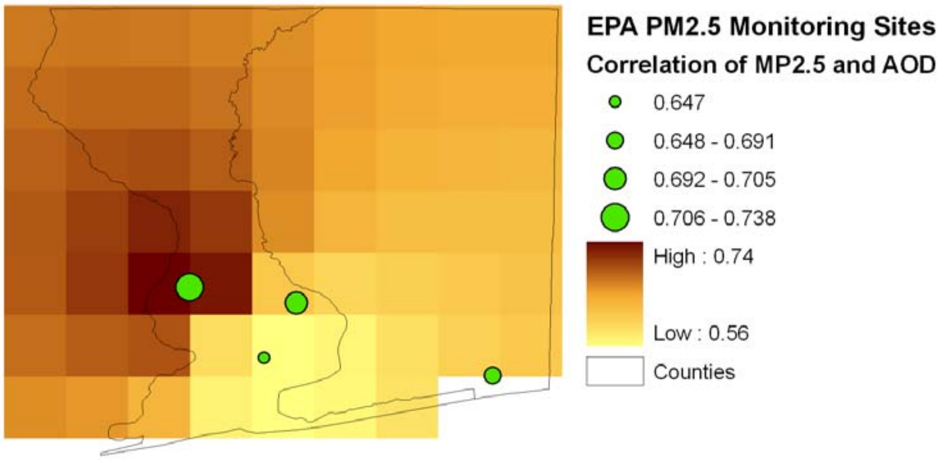


Figure 12. Raster surface of correlation between $PM_{2.5}$ and MODIS AOD.

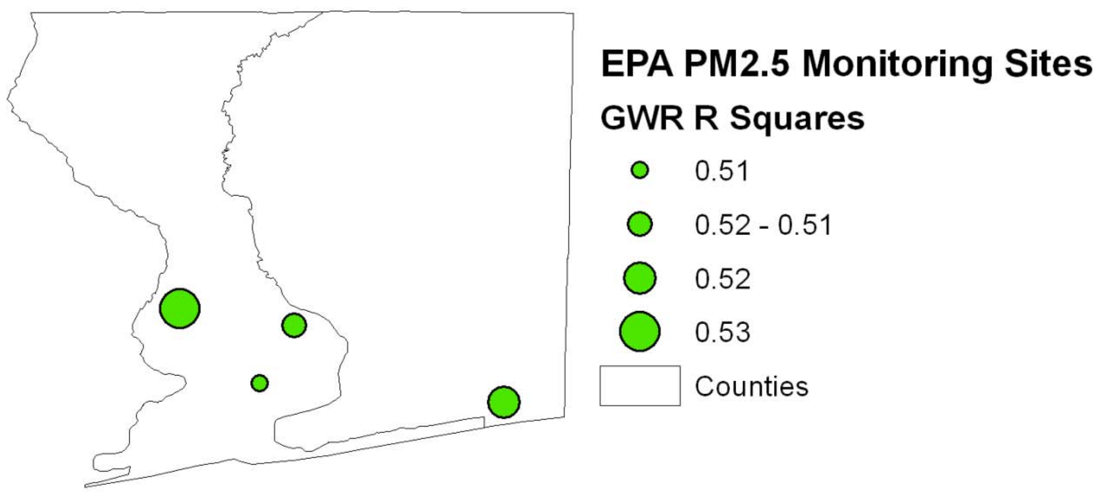


Figure 13. R squares of geographically weighted regression of PM2.5 on MODIS AOD.

5.3.7 Relationships Between Myocardial Infarction (MI) And AOD

Two models were fitted to examine the effect of AOD on MI. One is a spatial error model and the other is a Bayesian hierarchical model. The spatial error model used SMR as the dependent variable and AOD as the explanatory variable. Model results are shown in Table 11. The coefficient on the spatially correlated errors (λ) has a positive effect and is highly significant. AOD has a significantly positive coefficient (coefficient = 2.3173, p<0.001). Higher risk of MI is associated with higher aerosol optical depth.

A Bayesian hierarchical model was also fitted with a convolution prior to considering a county specific covariate – AOD. Table 12 and Figure 14 show the model results. The AOD coefficient kernel density curve reveals a positive effect of AOD on MI. The 95% credible set is (0.3208, 0.4708) for the AOD effect. Similar to the spatial error model above, the Bayesian model showed positive association between MI and AOD.

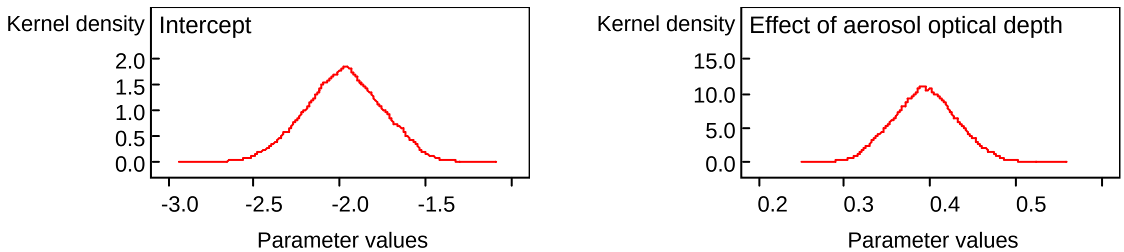
Table 11. MI and AOD: spatial error model.

```
SUMMARY OF OUTPUT: SPATIAL ERROR MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Spatial Weight        : cook_weight.GWT
Dependent Variable    :          SMR  Number of Observations:  589
Mean dependent var    :     0.929524  Number of Variables   :    2
S.D. dependent var    :     0.164615  Degree of Freedom     :  587
Lag coeff. (Lambda)   :     0.561258

R-squared             :     0.298463  R-squared (BUSE)      : -
Sq. Correlation       : -             Log likelihood        :  324.058852
Sigma-square          :     0.019010  Akaike info criterion :    -644.118
S.E of regression     :     0.137878  Schwarz criterion     : -635.360851

-----------------------------------------------------------------------
    Variable    Coefficient     Std.Error    z-value      Probability
-----------------------------------------------------------------------
    CONSTANT    0.5530958     0.04827483       11.45723    0.0000000
         AOD     2.317266      0.2843838       8.148373    0.0000000
      LAMBDA    0.5612576     0.06995007        8.02369    0.0000000
-----------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                        DF      VALUE          PROB
Breusch-Pagan test                          1        1.926539      0.1651376

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : cook_weight.GWT
TEST                                        DF      VALUE          PROB
Likelihood Ratio Test                       1        51.29031      0.0000000
```

Table 12**.** Markov chain Monte Carlo simulation results for Bayesian hierarchical modeling of MI vs. AOD.

Node	Mean	Standard Deviation	MC Error	2.5%	Median	97.5%	Start Iteration No.	Number of Samples
a_0	-1.977	0.2309	0.007435	2.437	-1.976	-1.536	10000	180002
a_1	0.3943	0.03839	0.001234	0.3208	0.394	0.4708	10000	180002

a_0 – Intercept, a_1 - Effect of aerosol optical depth (AOD)



Kernel estimates of the posterior densities of the fixed effects in the Bayesian hierarchical model

Figure 14. Coefficient kernel density curves, Bayesian hierarchical model (SMR of MI vs. AOD) for eastern US.

5.3.8 Relationships Between Chronic Coronary Heart Disease (CCHD) And AOD

Race and age standardized mortality rate (SMR) of CCHD was computed for each of the 2306 counties in the eastern USA for the time period 2003-2004. A mean AOD raster grid for the same period was derived from Moderate Resolution Imaging Spectrometer (MODIS) aerosol data and the average AOD was calculated for each county. A bivariate Moran's I scatter plot, a map of local indicator of spatial association (LISA) clusters, and three regression models (ordinary least square, spatial lag, and spatial error) were used to analyze the relationships between AOD and CCHD SMR. The global Moran's I (Figure 15) value is 0.2673 (p =0.001), indicating an overall positive spatial correlation of CCHD SMR and AOD. The entire study area is dominated by spatial clusters of AOD against SMR (high AOD and high SMR in the east, and low AOD and low SMR in the west) (permutations = 999, p=0.05) (Figure 16). Of the three regression models, the spatial error model achieved the best fit (R^2 =0.28). The effect of AOD is positive and significant (beta = 0.7774, p= 0.01) (Table 13).

Aerosol particle pollution has an adverse effect on CCHD mortality risk in the eastern US. High risk of CCHD mortality was found in areas with elevated levels of outdoor aerosol pollution, as indicated by satellite derived AOD. Escambia and Santa Rosa counties of Florida have relatively low rates of CCHD (SMR< 1) (Table 14). The LISA map shows that the two counties have high AOD values but low SMRs (Figure 16).

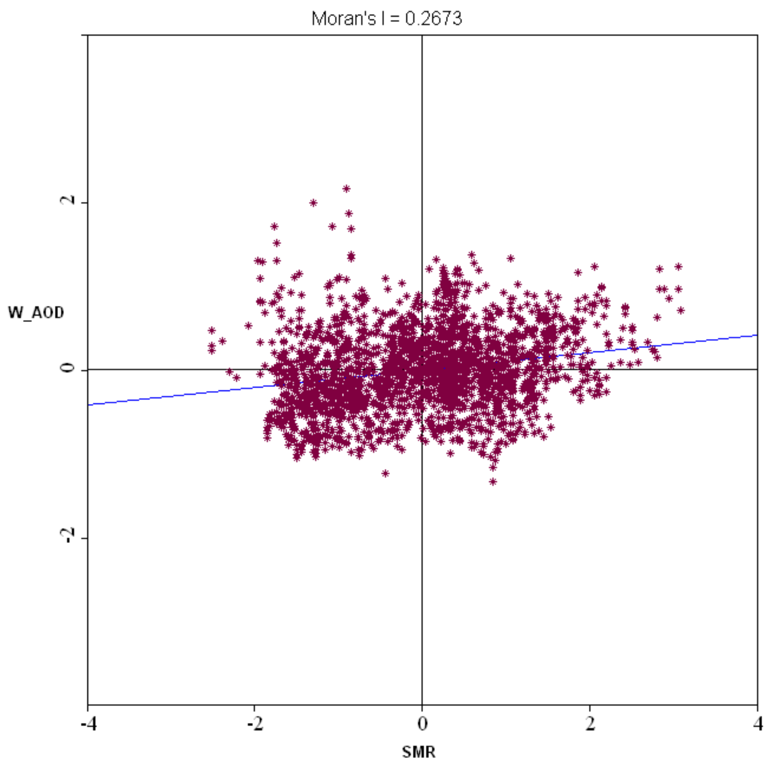


Figure 15. CCHD rate and AOD: bivariate Moran's I scatter plot.

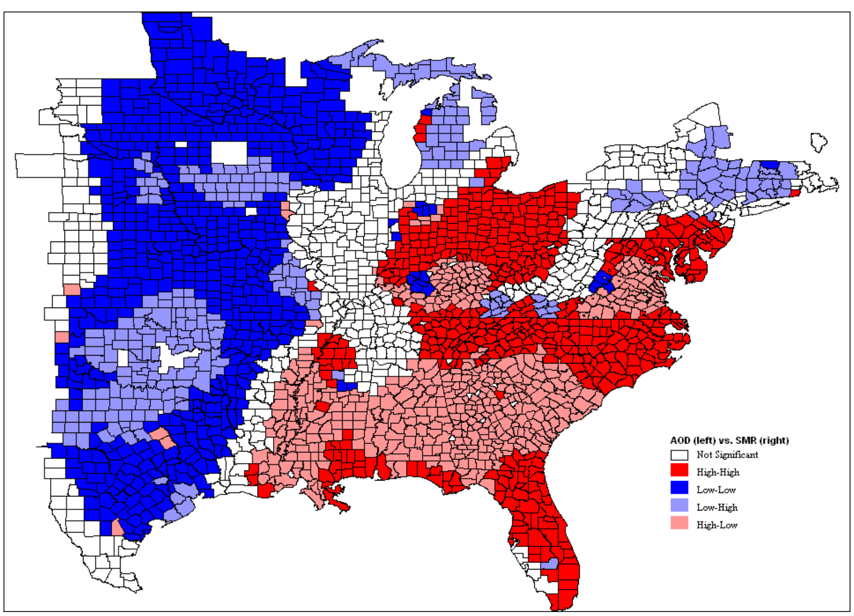


Figure 16. CCHD rate and AOD: local indicator of spatial association (LISA).

Table 13. Spatial error regression model.

Model description					
	Y	No. of variables	No. of observations	Degrees of freedom	
	SMR	2	2306	2304	
Model fit					
		R^2	Log likelihood	AIC	
		0.2800	-201.834	407.669	
Model estimation					
	Variable	Coefficient	Std. Error	t-Statistic	*p*
	CONSTANT	0.7506	0.0509	14.759	0.0000
	AOD	0.7774	0.3022	2.573	0.0101
	λ	0.5678	0.0231	24.600	0.0000
Diagnostic tests					
	Tests		DF	Value	*p*
Heteroskedasticity	Breusch-Pagan		1	0.067	0.7954
Spatial dependence	Likelihood Ratio		1	501.856	0.0000

Table 14. CCHD, AOD and PM2.5 data for Escambia and Santa Rosa Counties.

County	Observed CCHD	Expected CCHD	SMR	AOD	PM2.5
Escambia	1584	1675	0.946	0.197	12.334
Santa Rosa	531	542	0.980	0.194	12.507

5.3.9 Association Of CCHD With PM2.5

A Bayesian hierarchical model was employed to link PM2.5 predicted with the GWR model with age-race standardized mortality rates (SMRs) of chronic coronary heart disease (CCHD). The study found that chronic coronary heart disease mortality rate increases with exposure to PM2.5. (Figure 17 and Table 15). High risk of CCHD mortality was found in areas with elevated levels of fine particulate air pollution. The association between CCHD mortality and PM2.5 justifies the need of further toxicological studies of the influence of fine particulate air pollution on the heart. The evidence of raised CCHD mortality risk in high pollution areas supports targeting policy interventions on such areas to reduce pollution levels. Aerosol remote sensing like that used in the present study could help fill pervasive data gaps that impede efforts to study air pollution and protect public health.

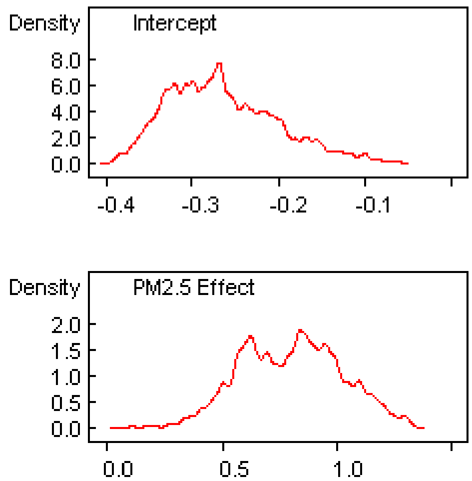


Figure 17. Kernel estimates of the posterior densities of the fixed effects in the Bayesian hierarchical model.

Table 15. Results of Bayesian hierarchical modeling.

Fixed effects	Posterior mean	Posterior median	Standard deviation	MC error	95% Credible set
β_0	-0.264	-0.273	0.064	0.003	(-0.366, -0.117)
β_1	0.802	0.812	0.223	0.010	(0.386, 1.225)

* Posterior means, medians, and 95% credible sets are based on 20,000 post-convergence iterations (from 60,001 to 80,000). Fixed effects are: β_0 - intercept, β_1 – effect of $PM_{2.5}$.

6 OVERALL CONCLUSIONS

The health tracking study carried out by PERCH indicates that some Zip codes in Escambia and Santa Rosa Counties have worse health outcomes than socio-economically and demographically matching Zip codes elsewhere in Florida, but other Zip codes in the area have better health outcomes than their matching Zip codes. These variations in health outcomes do not show clear spatial trends nor are they consistent for any one specific health outcome. An initial study of air pollution suggests that the proximity to stationary air emission sites may influence the incidence of some of the specific health outcomes, but not overall health. Geostatistical modeling corroborates this observation. Other exposures and life style decisions not evaluated in the PERCH studies may also affect the incidence of these specific health outcomes. A PERCH study that examined risk associated with inhalation of air pollution identified four zones with elevated cancer risk. One of these zones, the smallest one, coincides spatially with parts of Zip codes that have an elevated incidence of some health outcomes, but a direct cause and effect between risk and incidence was not established.

7 ACKNOWLEDGEMENTS

This report is a component of the "Assessment of Environmental Pollution and Community Health in Northwest Florida" supported by a USEPA Cooperative Agreement award X-9745502 to The University of West Florida (Project Director: Dr. K. Ranga Rao). The contents of this report are solely the responsibility of the authors and do not necessarily represent the official views of the USEPA. The report summarizes the PERCH studies that focused on air quality and its potential to adversely affect human health. The studies were undertaken because of an increasing concern in the community for environmental pollution and potential impacts on human health in Northwest Florida. Many students, listed in the original reports of the studies, assisted in various phases of the work. Their assistance has been invaluable.

8 OUTCOMES AND OUTPUTS

8.1 Final Reports

Health Tracking Study: http://www.uwf.edu/CEDB/Perch_USF_EPA_April04.pdf
Air Toxics Study: http://www.uwf.edu/CEDB/Perch_Air_Quality_Studies.cfm

8.2 Dissemination Of Results: Conference Presentations

2009 Hu, Z. and K. R. Rao. "Evaluating the relationship between chronic ischemic heart disease and air pollution using satellite aerosol optical depth data." Annual meeting of the Association of American Geographers, Las Vegas, Nevada.

> Abstract: This project first examined the relationship between U.S. EPA ground monitored PM2.5 concentration values and MODIS (Moderate Resolution Imaging Spectrometer) aerosol optical depth data for the conterminous U.S. MODIS Level 2 images were compared with PM2.5 data for the year 2004 with imaging time collated with PM2.5 measurement time within one hour. The Pearson's correlation value was calculated for each point containing at least 8 pairs of data values. Significant positive correlations between PM2.5 and AOD were found to be to the east of the -100° longitude line. For the year 2004, a geographically weighted regression (GWR) model was also fitted to examine the relationship between PM2.5 and AOD using annual mean PM2.5 and MODIS Level 3 AOD. It can be seen from the map of local R square that GWR predicts well in the eastern U.S. and poorly in the west. The coefficient raster surface AOD exhibits regional variation. The relationship between PM2.5 and AOD is not spatially consistent (stationary) across the conterminous states. Eastern U.S. shows higher AOD coefficient values, while values in the west are lower. Ecological spatial regression models and a Bayesian hierarchical simulation model using disease data at the county level for the eastern U.S. revealed a significant positive relationship between chronic ischemic heart disease mortality and AOD. The strong correlation between PM2.5 and AOD in the east suggests that AOD could be used as an air quality indicator to assess health effect of air pollution.

2008 Hu, Z., J. Liebens and K. R. Rao. "Assessing heath effect of aerosol particles using MODIS AOD remote sensing data." The 31st International Geographical Congress. Tunis, Tunisia. August 2008.

Abstract: Although most regulations of air pollution focus on gases, aerosol particles cause possibly more health problems than do gases. $PM_{2.5}$ causes the most severe health problems. For the use of public health assessment, particulate matter ground monitoring data often lacks spatially complete coverage. Some studies have found that satellite remotely sensed aerosol optical depth (AOD) is positively correlated to ambient PM concentration. This study assesses MODIS AOD and fine AOD data against spatio-temporally collocated EPA ground PM2.5 monitoring data for the conterminous U.S. for the year 2004. It was found that MODIS level 2 hourly AOD has a high positive correlation with $PM_{2.5}$ east of the -100 degree longitude line while western region shows no correlation. MODIS level 3 yearly mean fine AOD shows overall significant positive relationship with $PM_{2.5}$. A spatial lag model using aggregate data at county level shows that low birth weight rate is positively associated with fine AOD. Satellite measurement of fine AOD could directly be used as an air pollution indicator for public health effect assessment.

2008 Hu, Z. and K. R. Rao. "Extraction of particulate matter surface from MODIS Data for linking stroke mortality with air pollution in Northwest Florida." Annual Meeting of the Association of American Geographers, Boston, MA.

Abstract: Using Northwest Florida as the study area, this project intends to determine if there is association between stroke mortality and particulate matter concentration. Stroke death count data at the census tract level was obtained from Florida Vital Statistics in a 5-year (1998-2002) aggregate. Expected counts were calculated by adjusting ages using the National Vital Statistical System data for the US south as standard population. EPA's Air Quality System (AQS) ground-based PM2.5 measurements were obtained for the year 2000. However, the number of monitor stations is too limited to be useful for reliable interpolation for creating a PM2.5 surface. Previous studies have demonstrated that NASA Moderate Resolution Imaging Spectrometer (MODIS) satellite Aerosol Optical Depth (AOD) has positive relationship with PM2.5 during the warm season (April - September). Therefore, Terra MODIS AOD data for April through September of 2000 were grouped and the data values for the monitor sites were regressed on the measured PM2.5 concentration values. The regression equation was used to calculate a PM2.5 surface. The average PM2.5 concentration for each census tract was calculated using a GIS zonal statistic function. A Bayesian hierarchical model allowing for a convolution prior for the random effects was fitted using the Markov chain Monte Carlo (MCMC) simulation method. A 5,000 update burn in followed by a further 20,000 updates yielded the parameter for PM2.5 with the mean equal to 1.675 and a 95% credible set of (1.469, 1.875), which shows the strong positive relationship between stroke mortality and PM2.5.

2008 Hu, Z., J. Liebens and K. R. Rao. “Spatial associations between air emissions and health outcomes.” Department of Mathematics University of West Florida Colloquium Series.

No abstract.

2008 Hu, Z., J. Liebens and K. R. Rao. “Remote sensing of air quality and Bayesian hierarchical modeling of relationship between air pollution and disease.” Division of Environmental Health Environmental Public Health Tracking Florida Department of Health.

No abstract.

2008 Liebens, J., Z. Hu and K.R. Rao. “Spatial associations between air emissions and health outcomes.” Colloquium Series, Department of Mathematics, University of West Florida, Pensacola, FL.

No abstract

2007 Liebens, J. and K. Flanders. “Associations between spatial patterns of air emissions and morbidity.” Annual Meeting of the Association of American Geographers, San Francisco, CA.

Abstract: The incidence of some health outcomes is statistically significantly higher in NW Florida than in demographically, economically, and socially similar areas elsewhere in the state. Northwest Florida also has high rates of water and air pollution. Health outcomes have been shown elsewhere to be affected by environmental factors. This study examined if the spatial patterns of some of the health outcomes with high incidence in NW Florida are spatially associated with patterns of air emissions. The study also assessed the sensitivity of the results to the inclusion in the analysis of various types of air emission sources. Emission patterns were linked to the health outcomes with an index for the proximity of census blocks to emission sources. The proximity index was used unweighted, weighted with total source strength emission data, and with benzene equivalent emissions for Toxic Release Inventory (TRI) sites. The resulting three indexes for the census blocks were summarized by census tract or ZIP code as required by the pre-existing health data. Results show that various national and state government emission databases are inconsistent and have spatial information of greatly varying quality. Morbidity for specific respiratory illnesses such as pneumonia and asthma are associated with patterns of air emissions for some of the emission database/proximity index combinations. Cancer morbidity shows little influence from emissions. Racial inequity in exposure to air emissions is very small in the study area and does not affect health outcomes appreciably. Supported by EPA Cooperative Agreement X-9745500.

2007 Liebens, J. and K. Flanders. "Associations between spatial patterns of air emissions and health outcomes." NW Florida Regional Environmental Symposium, University of West Florida, Pensacola, FL.

Abstract: The incidence of some health outcomes is statistically significantly higher in NW Florida than in demographically, economically, and socially similar areas elsewhere in the state. Elsewhere, health outcomes have been shown to be affected by environmental factors. This study examined if the spatial patterns of some of the health outcomes with high incidence in NW Florida are spatially associated with patterns of air emissions. Emission patterns were linked to the health outcomes with an index for the proximity of census blocks to emission sources. The proximity index was used unweighted, weighted with total source strength emission data, and with benzene equivalent emissions for Toxic Release Inventory (TRI) sites. The resulting three indexes for the census blocks were summarized by census tract or ZIP code as required by the pre-existing health data. Results show that various national and state government emission databases are inconsistent and have spatial information of greatly varying quality. Morbidity for specific respiratory illnesses such as pneumonia and asthma are associated with patterns of air emissions for some of the emission database/proximity index combinations. Cancer morbidity shows little influence from emissions. Supported by EPA Cooperative Agreement X-9745500.

2007 Hu, Z., J. Liebens and K. R. Rao. "Exploring relationship between asthma and air pollution: a geospatial methodology using dasymetric mapping, GIS analysis and spatial statistics." The 15th International Conference on Geoinformatics, Nanjing, China, 2007.

Abstract: This paper presents methodology using dasymetric mapping from remotely sensed imagery, geographic information system (GIS), spatial analysis and spatial statistics to explore relationship between asthma and air pollution in the Pensacola metropolitan region of Florida. Health outcome indicators thought to be sensitive to increased exposure of airborne environmental hazards are mortality and morbidity rates for total population asthma patients. Environmental data for the time around the year 1999 include point source pollution sites and emissions, traffic count with emission estimates, and a Landsat ETM+ image. Standardized mortality/morbidity ratios (SMRs) were used as dependent variables for the analysis. A centroid map was created from the zip code map with each centroid assigned the corresponding SMR values. Then spatial interpolation using the Kriging method was used to generate continuous SMR surfaces. An emission or point count based kernel density raster map was created from each of the air pollution maps. A raster layer 'greenness' was extracted using tasseled cap transformation from the Landsat ETM+ image. The dasymetric mapping technique was employed to limit the analysis and modeling to the area where human activities occur. The ETM+ image was classified into a thematic land use/cover map and the developed area extracted. A road network was combined with the developed area to generate a buffer (buffer distance = 1.5 km). A random sample with enough number of points was generated across the study area and 505 points were found within the developed area and the buffer. Data values at these sample points were extracted and used for statistical modeling. Two spatial autoregressive models (spatial error and spatial lag) were fitted. Both models show relationship between the asthmas outcome indicators and air pollution (positive) and 'greenness' (negative).

2005 Worley A. and J. Liebens. "Relationships between health outcomes and air pollution in Northwest Florida." (poster, A. Worley as first author). Annual Meeting of the Southeastern Division of the Association of American Geographers, Palm Beach, FL.

Abstract: Morbidity and mortality in Northwest Florida vary spatially at the ZIP code level and, for some ZIP codes, are significantly different from those in socially and demographically similar ZIP codes elsewhere in the state. We examined if in Northwest Florida statistical and spatial relationships exist between morbidity and mortality, demography and air pollution, and compared these results with those for the similar ZIP codes elsewhere in the state. To this end, we calculated proximity indexes for permitted air emitters within 10 km of Census 2000 block centroids for all ZIP codes in an existing health outcome study for the region. The indexes were weighted separately with TRI site total emissions, benzene equivalent pound emissions from TRI sites, and with a combination of FDEP major and minor air emitters. Results indicate that differences in mortality and morbidity within Northwest Florida are not statistically related to the weighted proximity indexes. Proximity indexes for Northwest Florida ZIP codes and the socially and demographically similar ZIP codes are comparable. Statistical analysis does not identify environmental inequity in proximity to TRI sites in Northwest Florida. Supported by EPA Cooperative Agreement X-9745500.

8.3 Dissemination Of Results: Academic Publications

Hu, Z., J. Liebens and K. R. Rao. (under review). Merging satellite measurement with ground-based air quality monitoring data to assess health effects of fine particulate matter pollution. In J. Maantay & S. McLafferty (Eds.), *Geospatial Analysis of Environmental Health*. Springer Verlag.

Hu, Z. and K. R. Rao. 2009. Particulate air pollution and chronic ischemic heart disease in the eastern United States: a county level ecological study using satellite aerosol data. *Environmental Health*, 8:26.

Hu, Z. 2009. Spatial analysis of MODIS aerosol optical depth, $PM_{2.5}$, and chronic coronary heart disease. *International Journal of Health Geographics*, 8:27.

Hu, Z., J. Liebens and K. R. Rao. 2008. Linking stroke mortality with air pollution, income, and greenness in northwest Florida: an ecological geographical study. *International Journal of Health Geographics*, 7-20.

Hu, Z., J. Liebens, and K. R. Rao, 2008. Assessing heath effects of aerosol particles using MODIS AOD remote sensing data." *Proceedings of the 31st International Geographical Congress*. Tunis, Tunisia. August 2008.

Hu, Z., J. Liebens, and K. R. Rao, 2007. Exploring relationship between asthma and air pollution: a geospatial methodology using dasymetric mapping, GIS analysis and spatial statistics. *Geoinformatics 2007: Geospatial Information Science. Proceedings of SPIE* Vol. 6753, 67532T. SPIE: The International Society for Optical Engineering.

8.4 Dissemination Of Results: Media Reports

1. Independent News (newspaper): Air pollution and disease, 2/28/2008
2. ABC WEAR-TV News interview of Z. Hu

8.5 Graduate Students Trained At UWF

Angela Worley, Kristal Flanders, Brail Stephens, Michael Somerville.

8.6 Outcomes

PERCH project studies provided, for the first time, a detailed analysis of air quality and the potential impacts on human health in Northwest Florida. These analyses included novel methodologies (e.g., proximity index analysis, geostatistical analysis) that could have wider applications nationally and abroad. RAIMI system, first developed and applied by EPA Region 6, has been successfully applied for analysis of air toxics impacts in Northwest Florida—first such application in the geographic area covered by EPA Region 4.

The results of our analysis have been communicated to the scientific community through presentations at professional meetings and publications in peer-reviewed journals and conference proceedings, and thus contributing new knowledge to the areas of investigation. As a result of our findings, additional collaborative work with Florida Department of Health is being done to assist in ongoing environmental public health tracking.

Our results have also been widely publicized through local news media (news papers, TV coverage) and by postings at our web site. As a result, there is an increased awareness of our study outcomes among the public, and also recognition of issues of concern by the city, county, and state governmental agencies. This should enable them to facilitate measures that would further enhance air quality and protection of public health.