A CASE STUDY OF ENVIRONMENT JUSTICE –

AIR QUALITY IN HARRIS COUNTY, TEXAS

by

Ting Hsuan Yang

A directed research report submitted to the Geography Department of

Texas State University in partial fulfillment

of the requirements for the degree of

Master of Applied Geography

with a specialization in geographic information science

May 2021

Committee Members:

Dr. F. Benjamin Zhan

Dr. T. Edwin Chow

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Description
EJSCREEN	Environmental Justice Screening and Mapping Tool
EJ	Environmental Justice
GWR	Geographically Weighted Regression
NATA	National Air Toxics Assessment
PM	Particulate Matter
U.S. EPA	United States Environmental Protection Agency

ABSTRACT

This research integrates EJSCREEN data and geographically weighted regression analysis to explore the spatial variation in the relationships between environmental indicators and demographic factors in Harris County, Texas. The results show high levels of the goodness of fit for most of the models in PM 2.5, ozone, NATA respiratory hazard index, and NATA diesel particulate matter, except for traffic proximity. The results provide a clear representation that there is a high level of associations between a higher volume of air pollutants or a higher level of relevant health risk and a higher percentage of populations vulnerable to environmental risks.

I. INTRODUCTION

Environmental inequality is the phenomenon that environmental risks are geographically unequally distributed by race, age, socioeconomic status, and other demographic identifications. Hence it is a complex process with history, politics, economic development, and other human activities involved. Over the past few decades, with surging awareness of environmental change, for example, in climate, air quality, or ecosystem, and the fact that environmental risks are distributed unevenly, researchers started to discover that some of the population is exposed disproportionately to these risks than others in terms of race (Bell and Ebisu, 2012; Crowder and Downey, 2010; Downey and Hawkins, 2008), education (Ard, 2015; Branis and Linhartova, 2012), income(Bouvier, 2014; Branis and Linhartova, 2012;) and other characteristics (Bakhtsiyarava and Nawrotzki, 2017; Smiley, 2019; Downey et al.,2017). Previous research has also demonstrated inequalities in vulnerability, i.e., the ability to anticipate, cope with, or recover from the impact of natural hazards (Wisner et al., 2003), between different demographic groups when exposed to the same magnitude of environmental risk and hazards. (Brooks et al., 2005; Maantay and Maroko, 2009; Mohai et al., 2009)

Houston, the seat of Harris County, has a great diversity of population in age, ethnicity, the nation of origin, religion, language, and culture. As a dynamic mix of immigration, like all other major cities in the nation and in the world, there are certain patterns of population distribution in Houston and its nearby satellite cities and neighboring communities. This research explores the environmental inequality in Harris County in terms of air pollution by investigating the association between the demographic index and air pollution-related environmental indicators of EJSCREEN data.

The main objectives of this research report are to answer the following questions:

1. Where are the areas with disproportionate air pollutants in Harris County?
2. Are there associations between the demographic indicators and the air pollution indicators? If so, is there are significant differences between these associations?

To answer these questions, this research will employ EJSCRENN data to identify areas with minority populations and potential environmental quality issues and examine the data using geographically weighted regression (GWR) models to analyze the correlation between environmental and demographic indicators of interests and the spatial variations of relationships in Harris County in 2017. GWR analysis has been broadly applied on environmental justice topics to reveal the spatial variations of relationships between the relatively vulnerable population and the uneven distribution of pollutions and related human health risks (Gilbert and Chakraborty, 2011; Jephcote and Chen, 2012; Jin and Lu, 2017; Lee et al., 2017; Lersch and Hart, 2014; Mennis and Jordan,2008). However, there were limited studies that have employed this method to analyze the geographic variations of environmental injustice in the Houston / Harris County area.

By meeting this objective, the results provide a clear representation of whether an association between a higher level or volume of air pollutants and a higher percentage of minority populations and, if so, the strength of the associates.

II. LITERATURE REVIEW

According to U.S. EPA (n.d.), the goal of environmental justice is fair treatment for all people to enjoy the same degree of protection from environmental and health hazards, and the equal access to the decision-making process to have a healthy environment in which to live, learn, and work. Before this concept was generally known by the world, the term “Environmental racism” was initially coined by Benjamin Chavis in1982 (*Confronting Environmental Racism: Voices from the Grassroots*, 1993). Chavis reported a strong statistical correlation between communities of color and the location of hazardous wastes landfill sites in Afton, North Carolina. Due to discrimination by race, people of color had been neglected in policy development and deprived of the right to participate in public decision-making. However, soon the researchers realized that environmental racism is not definitive. Minority groups other than communities of color are also relatively vulnerable to environmental risks. (Cutter, 1995) As the environmental movement progressed, activist communities became to use environmental justice for it is a more inclusive term that is beyond race or ethnicity but embrace other socioeconomic deprived minority population, e.g. women, low income, unemployment, low education level, children, elders, etc. Furthermore, the focus of attention shifted from the preservation of faraway pristine habitats to a more localized environment enhancement of the living space of affected residents. Environmental movements were no longer limited to the white upper-class but diversely included local residents, working-class, and people of color. (Cutter, 1995)

Houston is the fourth most populous city in the nation since the 1990s (United States Census Bureau, n.d.) and is the largest in the southern U.S. and Texas (City of Houston, 2018). The growth progress of the city is inseparable from the development of the petroleum industry (Feagin, 1985, Qian, 2010). Crude oil has not only brought wealth to Houston, but it also has helped the whole business economy of the city to rapidly diversify into a broad range of industries, for example, banking and finance, technology, medical research, and health care, and it makes the immigration keep growing to date. From 2013 to 2018 (estimates as of July 1), the population of Houston has grown 5% from 2,134,707 to 2,325,502 people (U.S. Census Bureau, 2018). With its expanding economy and population, Houston is expected to become the third-largest United States city within a decade.

The major sources of air pollutants in Houston include tailpipe emissions from vehicles, emission of manufacturing facilities, emission of petroleum industry plants (e.g., crude oil refineries), petrochemical complex along the Houston Ship Channel and the Port of Houston, and other small operations (Bethel, 2005). Most of the oil refineries in Houston are located on the east side of the city along Buffalo Bayou, which is utilized as a ship channel, approximately15 miles away from the downtown. Hazardous Air Pollution emissions can be found from sources in oil and natural gas production, and natural gas transmission and storage (EPA, n.d.). Living or working close to air pollution resources has been related to asthma and long-term chronic respiratory infection (Brauer et al., 2007). Pollutants emitted by oil refineries can also be brought into the urban area by winds and perception. Nevertheless, Houston, as a highly developed city with a population of 2.3 million and counting, also suffers from heavy traffic and the poor air quality that comes with traffic.

Air pollution is a major threat to global public health (Bruce et al., 2000). A broad research base indicates an association between adverse health effects and ambient air pollutants, including ozone, nitrogen oxides, and particulate matter (Dockery et al., 1993; Gong et al., 2018; Mannucci and Franchini, 2017; Peel et al. 2006, Ritz and Wilhelm, 2017). Research on the impact of outdoor air pollution on the development of lung function and respiratory health had accelerated and is receiving a lot of interest (Duijts, 2012, Götschi et al., 2008, Miller and Marty, 2010, Sweileh et al., 2018). However, until today, despite the movement of environmental justice that has been developing for decades, Houston still has communities that are more deprived of equal rights to be protected from environmental risks of air pollutants.

III. RESEARCH METHODS

In order to determine the association between a higher level or volume of air pollutants and a higher percentage of minority populations, a multivariate regression analysis of data collected from the Environmental Justice Screening and Mapping Tool (EJSCREEN) developed by United States Environmental Protection Agency (U.S. EPA) was conducted. The EJSCREEN provides screening-level data based on nationally consistent data of environmental and demographic indicators. In EJSCREEN, the basic geographic resolution level is the Census block group, and the data is published yearly. EJCSREEN data is presented in three categories: environmental indicators, demographic indicators, and EJ indexes. The indicator/index names of EJCREEN data are:

1. Environmental indicators: PM 2.5, Ozone, NATA Diesel PM, NATA Cancer Risk, NATA Respiratory HI, Traffic Proximity, Lead Paint Indicator, Superfund Proximity, RMP Proximity, Hazardous Waste Proximity, and Wasted Discharge Indicator.
2. Demographic indicators: Minority (People of Color), Low Income, Linguistic Isolated, Less Than Highschool Education, Under Age 5, Over Age 64, and a Demographic Index (based on the average of two demographic indicators: Low-Income and Minority).
3. EJ Indexes: Demographic index, PM 2.5, Ozone, NATA Diesel PM, NATA Cancer Risk, NATA Respiratory HI, Traffic Proximity, Lead Paint Indicator, Superfund Proximity, RMP Proximity, Hazardous Waste Proximity, and Wasted Discharge Indicator.

EJ index is a combination of environmental and demographic information. Each of the EJ indexes was calculated by combining a single environmental indicator with demographic information. (U.S. EPA, 2019) It considers the extent to which the local demographics are above the national average. It does this by looking at the difference between the demographic composition of the block group, as measured by the Demographic Index, and the national average (approximately 35%). Mathematically, the EJ Index is constructed as the product of three items, multiplied together as follows:

$$
\begin{aligned}
& EJ\ Index \\
& = \ (Environmental\ Indicator) \\
& * (Demographic\ Index\ for\ Block\ Group - Demographic\ Index\ for\ US) \\
& * (Population\ count\ for\ Block\ Group)
\end{aligned}
$$

The study area consists of 2143 Census block groups in Harris County, Texas. Harris County has a population from various racial/ethnic backgrounds and a large growing international community, in part because of its world-leading academic institutions and powerful medical, energy, and aerospace industries. According to the population estimates of the U.S. Census 2019, compared to the entire population of the United States, Harris County has a higher percentage in the population of color, children under five years old, and households suffering from low-income and poverty (Table 1).

Table 1. Harris County and the United States population components estimates, 2019.

	Harris County	United States
Total Population Estimates	4,713,325	328,239,523
Race and Hispanic Origin		
White alone	69.6%	76.3%
Black or African American alone	20.0%	13.4%
American Indian and Alaska Native alone	1.1%	1.3%
Asian alone	7.3%	5.9%
Native Hawaiian and Other Pacific Islander alone	0.1%	0.2%
Two or More Races	2.0%	2.8%
Hispanic or Latino	43.7%	18.5%
White alone, not Hispanic or Latino	28.7%	60.1%
Age		
Persons under 5 years	7.4%	6.0%
Persons under 65 years and over	10.9%	16.5%
Income & Poverty		
Persons in poverty*	15.5%	10.5%

* Household income less than thresholds of poverty guidelines published by U.S. Census Bureau yearly. (U.S. Census, 2019)

This research applied a spatial statistical technique known as geographically weighted regression to explore the spatially varying relationships between different population identifications and environmental indicators in EJSCREEN data. Unlike commonly used statistical methods such as linear regression, which only provide global results that enclose the entire study area without explaining spatial variability, geographically weighted regression is capable of providing important spatial differences in relationships and prevent misrepresenting nuanced local processes within a study area.

There are six demographic indicators provided in EJSCREEN data, but multicollinearity among the predictor variables was found to be a problem as the indicators of Less than High School Education and Linguistic Isolated are highly correlated with some of the other demographic indicators. Therefore, these two demographic indicators were excluded from the analysis results. Among eleven environmental indicators provided in EJSCREEN data, this research selected five indicators that are for the analysis.

The input variables used in geographically weighted regression analysis in this research are:

- Independent variables:
    - MINORPCT: Percentage Minority Population,
    - LOWINCPCT: Percentage Low Income (< 2x poverty level),
    - OVER64PCT: Percentage Over Age 64, and
    - UNDER5PCT: Percentage Under Age 5.
- Dependent variables:
    - PM25: PM 2.5 Concentration Score,
    - OZONE: Ozone Concentration Score,
    - DSLPM: NATA Diesel Particulate Matter,
    - PTRAF: Traffic Proximity, and
    - RESP: NATA Respiratory Hazard Index.

This research applied each of the environmental indexes separately as the dependent variable and all four demographic indexes as the independent variables for GWR analyses. The fixed numbers of neighborhood approach based on golden search neighborhood selection method was utilized in this research. The model first finds the maximum and the minimum number of neighbors and test the Akaike Information Criterion (AICc) at various numbers of neighbors incrementally between them, then finally determines the best neighborhood size with the lowest AICc value.

IV. RESULTS

Summary statistics for all dependent variables utilized in this research are provided in Table 2. In EJSCREEN 2017 data, PM 2.5, ozone, and diesel PM indicators are the estimates of ambient levels of air pollutants. PM 2.5 is the estimate of the annual average ambient level of inhalable particles, with diameters that are generally 2.5 micrometers and smaller in microgram per cubic meter ($\mu g/m^3$); ozone is the summer seasonal average of daily maximum 8-hour concentration in air in parts per billion (ppb); NATA diesel PM is diesel particulate matter level in microgram per cubic meter ($\mu g/m^3$). Traffic proximity is the count of vehicles (AADT, average annual daily traffic) at major roads within 500 meters, divided by distance in meters. NATA respiratory hazard index is as calculated as the ratio of air toxics exposure concentration to health-based reference concentration.

Table 2. Descriptive statistics for dependent variable analyzed. (EJCREEN, 2017)

Dependent variable	N	Mean	Minimum	Maximum	SD
PM 2.5 ($\mu g/m^3$)	2143	10.94	9.89	11.31	0.31
Ozone (ppb)	2143	36.39	34.54	38.80	1.01
Traffic Proximity[a]	2143	761.41	0	15544.37	1371.29
NATA Respiratory Hazard Index[b, c]	2143	2.20	1.17	5.52	0.42
NATA Diesel PM ($\mu g/m^3$)[c]	2143	1.22	0.35	4.25	0.47

[a] Count of vehicles (AADT, avg. annual daily traffic) at major roads within 500 meters, divided by distance in meters.
[b] Ratio of exposure concentration to health-based reference concentration.
[c] Data year is 2014.

Spatial distribution maps of the dependent variables are displayed in Figure 1 to Figure 5. These environmental indicators differ greatly in what they indicate. As shown in Figure 1, higher levels of PM 2.5 can be found around the center of Houston City, and the values decrease as the distance from the downtown increases. The Ozone concentration level tends to elevate from southeast to northwest of the County (Figure 2). Traffic proximity level is greater along the main roads and interstate highways, such as I-10, I-45, I-69, I-610, and U.S. Route 290, the block group above Addicks where hosts major operations for many energy sector companies, and also the block groups near Jacinto City as there is a large amount of traffic in and out the petroleum industry facilities (Figure 3). Higher values of NATA respiratory hazard index can be observed near downtown Houston, along Gulf Freeway (the portion of Interstate 45 between downtown Houston and Galveston at the southeast), between Aldine and Spring where the George Bush International Airport locates, and from Baytown, Cedar Park to Jacinto City as there is a large amount of major petroleum industry plants located along the channel (Figure 4). Similar to the distribution pattern of traffic proximity and volume, a greater level of diesel PM can be found in block groups located near downtown Houston, along I-10, I-45, and I-610 (Figure 5).

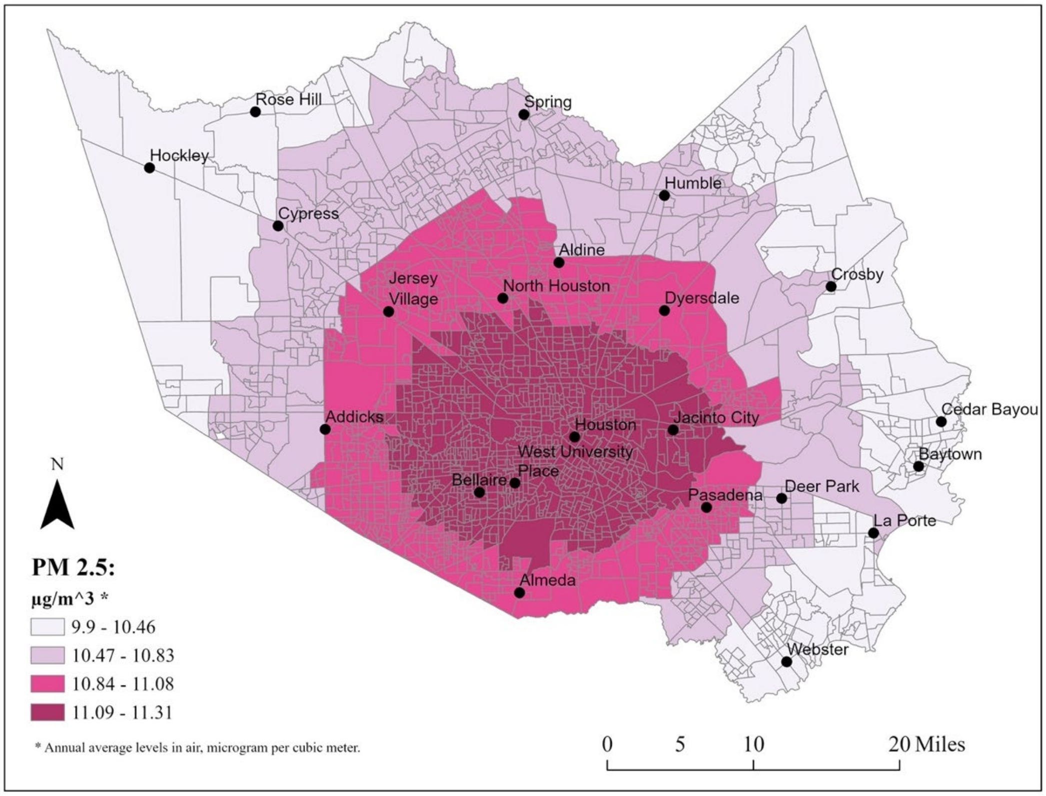

Figure 1. Spatial distribution of PM 2.5 annual average level in Harris County, 2017.

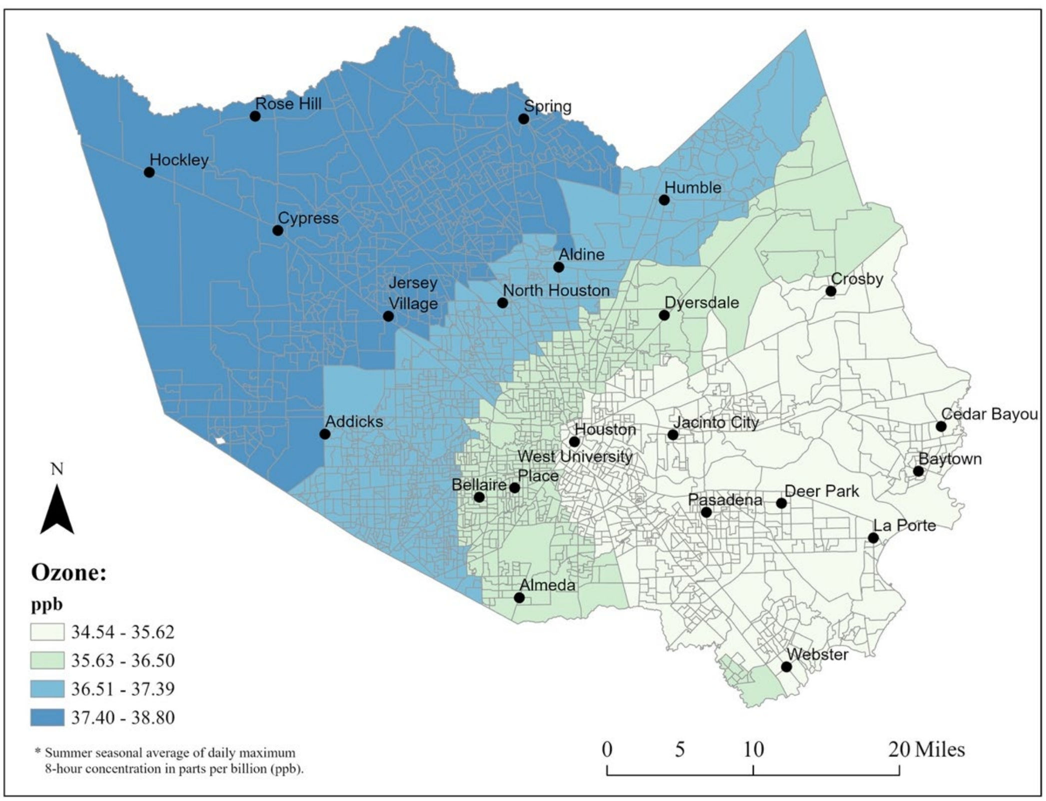

Figure 2. Spatial distribution of ozone summer seasonal average of daily maximum 8-hour concentration in Harris County, 2017.

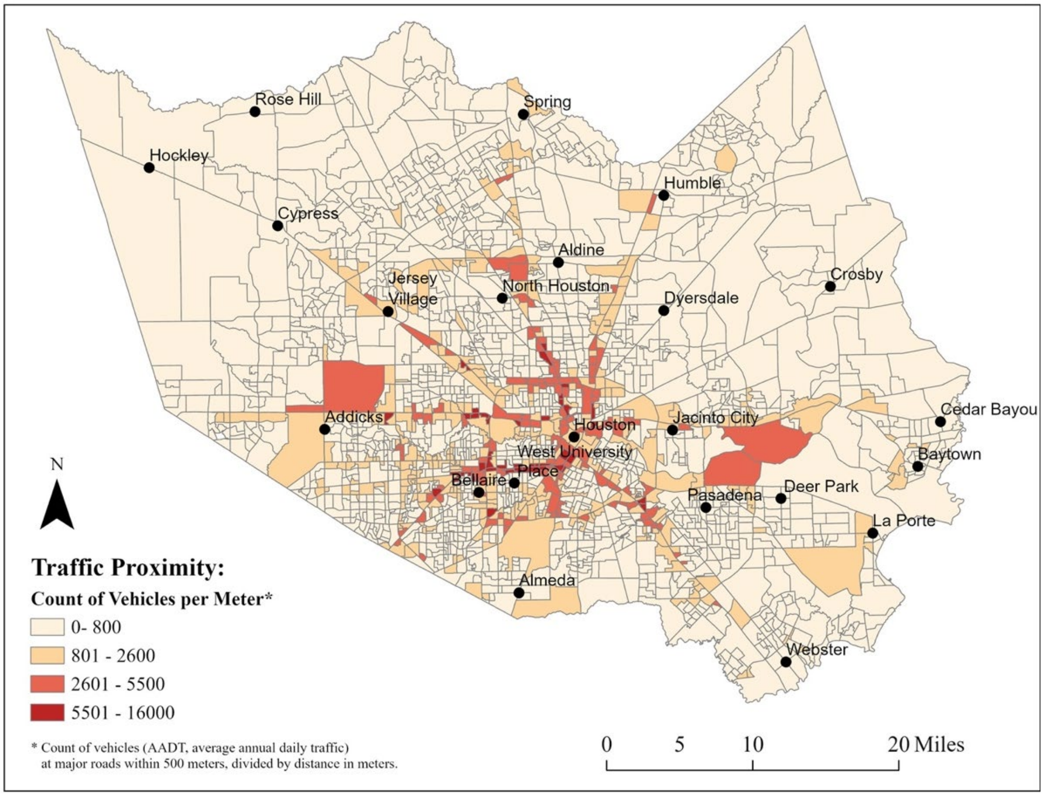

Figure 3. Spatial distribution of traffic proximity in Harris County, 2017.

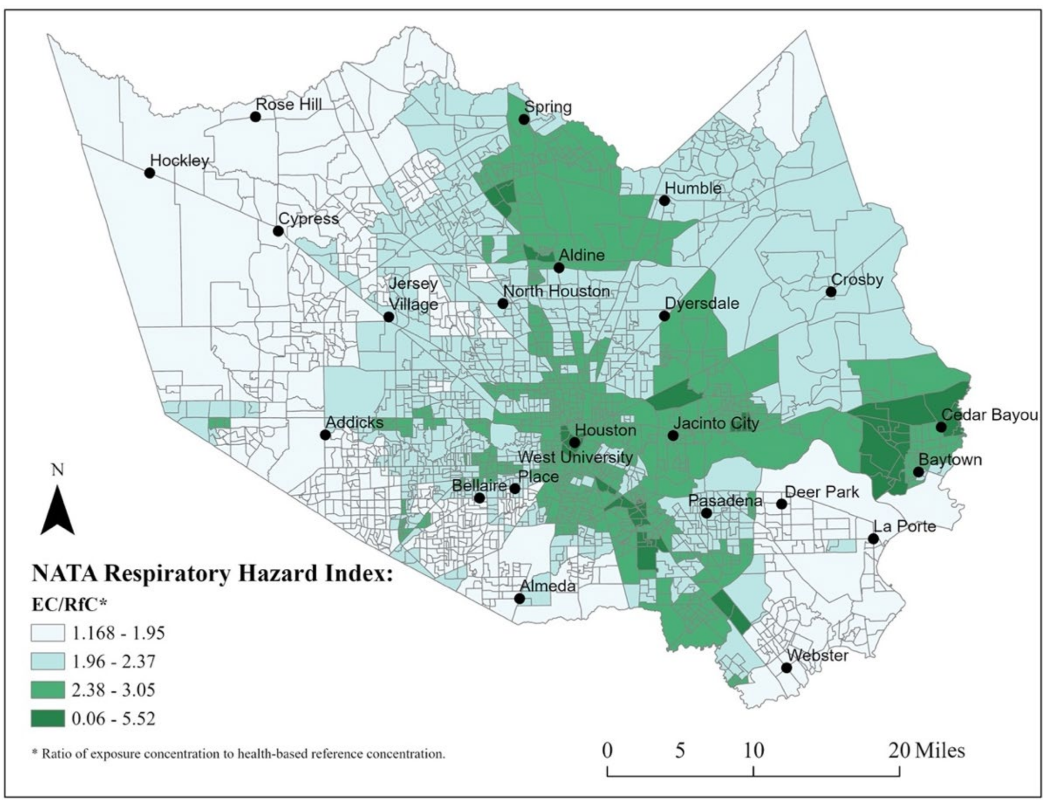

Figure 4. Spatial distribution of NATA respiratory hazard index in Harris County, 2014.

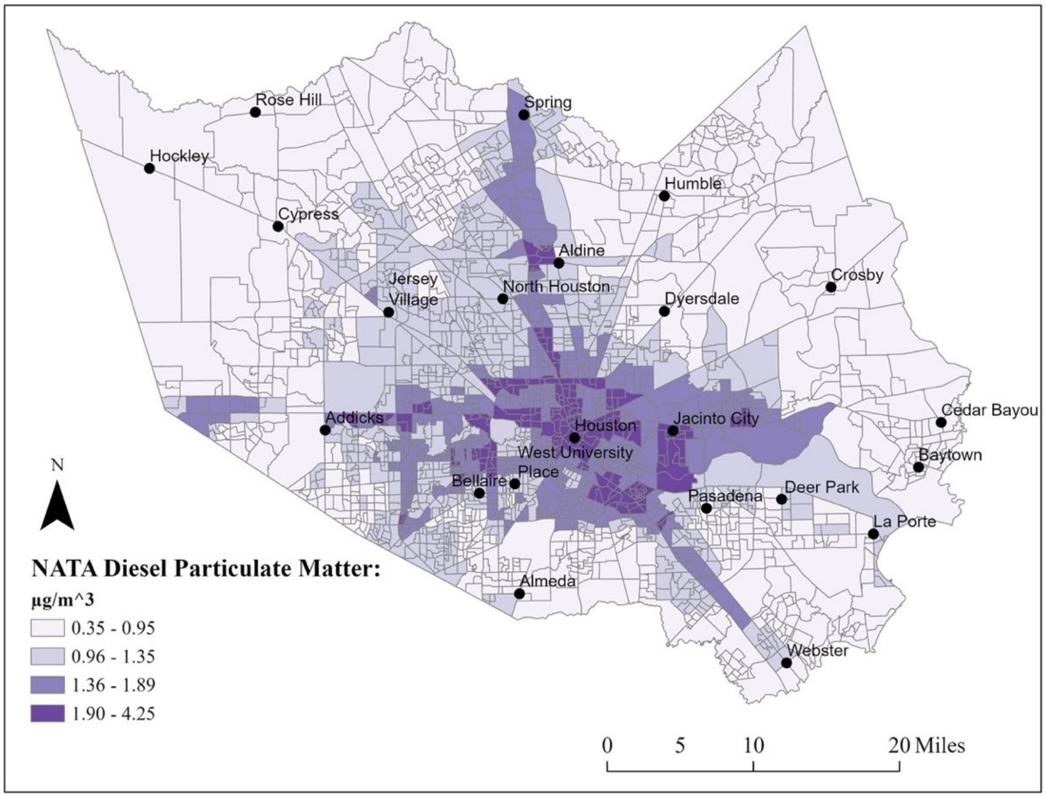

Figure 5. Spatial distribution of diesel particulate matter level in Harris County, 2014.

Geographically weighted Regression (GWR) tool in ArcGIS Pro 2.7 was used to analyze the EJCREEN data. The performance of GWR models using different environmental indicators as the dependent variable is presented in Table 1. As shown in the results, the considerable high R-squared values for models of PM 2.5 (0.9902), Ozone (0.9968), NATA Respiratory Hazard Index (0.8999), and NATA Diesel PM (0.888) indicate that the overall high goodness of fit. However, the local models did not fit traffic proximity very well (0.46).

Negative values of local R-squared can be found in the following results, except for ozone models, which may not be noticed if we only inspect the overall R-squared of the models for each environmental indicator in Table 3. These negative values of R-squared may result from the instability and inaccuracy caused by limited variation in some local areas, and thus the results should be examined with caution.

Table 3. Geographically Weighted Regression (GWR): summary of results and comparison between input environmental indicators.

	R^2	Adjusted R^2	Sigma-Squared	Sigma-Squared MLE
PM 2.5	0.9902	0.9843	0.0015	0.0009
Ozone	0.9968	0.9950	0.0051	0.0033
Traffic Proximity	0.4600	0.3000	1320099.05	1012045.66
NATA Respiratory Hazard Index	0.8999	0.8315	0.0301	0.0179
NATA Diesel PM	0.8880	0.8261	0.0383	0.0247

PM 2.5

The numerical results associated with the GWR analysis of PM 2.5 are summarized in Table 4. Spatial variations in these values demonstrate how the combined statistical effect of the independent variables on PM 2.5 differs across census block groups in Harris County.

Table 4. Geographically Weighted Regression (GWR) of PM 2.5.

	GWR coefficients			**Percent of census block group by significance (95% level) of t-statistic**		
	Min	**Median**	**Max**	**$t \leq -1.96$**	**$-1.96 < t < 1.96$**	**$t \geq 1.96$**
Minority	-0.66	0.01	1.30	4.01%	76.99%	18.99%
Low Income	-1.21	0.01	0.84	7.05%	85.11%	7.84%
Under Age 5	-1.55	-0.03	3.00	6.30%	88.10%	5.60%
Over Age 64	-1.28	0.02	4.17	3.73%	83.71%	12.55%
Intercept	9.72	10.99	11.56	0.00%	0.00%	100.00%
R-squared	-0.15	0.78	0.94			

The median local R-squared produced by the GWR is 0.78, representing a model that left only 21% of the variance in PM 2.5 unexplained. The spatial distribution of local

R-squared values generated by GWR analysis of PM 2.5 is depicted in Figure 6. The GWR models performing the best among block groups located near Crosby, North Houston, Cypress, and through the southern portion of the County. In other areas, such as block groups near downtown Houston and northwest of the County, PM 2.5 is not explained adequately by the independent variables.

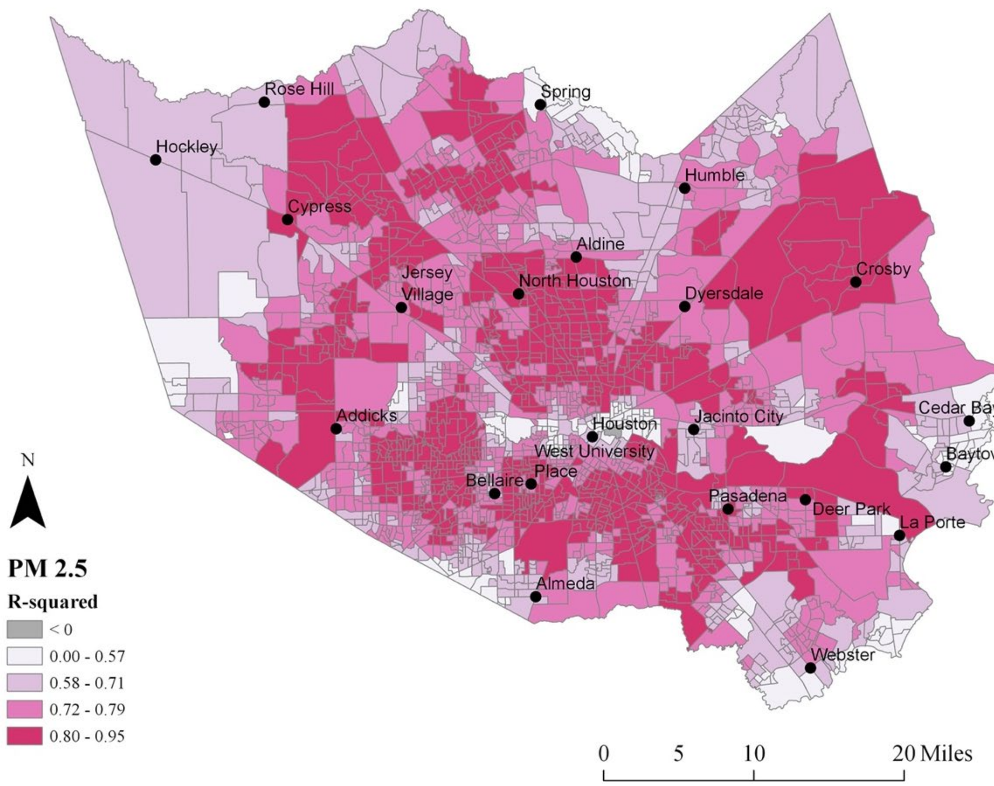

Figure 6. GWR model performance of PM 2.5: distribution of local R-squared values by block group.

Ozone

The model performance for ozone is similar to the results of PM 2.5 as the median local R-squared produced by the GWR is the same (0.78) shown in Table 5. As shown in Figure 7, the overall spatial variations pattern of the relationship between Ozone and the input independent variables is similar to results of PM 2.5 across census block groups in Harris County except for the block groups located along the Houston ship channel near

La Port, Deer Park and Pasadena which have lower R-squared values, and downtown Houston in where the model of Ozone fits better.

Table 5. Geographically Weighted Regression (GWR) of ozone.

	GWR coefficients			**Percent of census block group by significance (95% level) of t-statistic**		
	Min	**Median**	**Max**	**t ≤ -1.96**	**-1.96 < t < 1.96**	**t ≥ 1.96**
Minority	-2.65	-0.02	3.02	17.41%	72.70%	9.89%
Low Income	-2.87	-0.03	1.29	13.30%	77.18%	9.52%
Under Age 5	-6.81	0.12	4.05	5.51%	84.69%	9.80%
Over Age 64	-5.17	0.01	2.64	11.39%	76.95%	11.67%
Intercept	32.10	36.44	39.04	0.00%	0.00%	100.00%
R-squared	0.15	0.78	0.93			

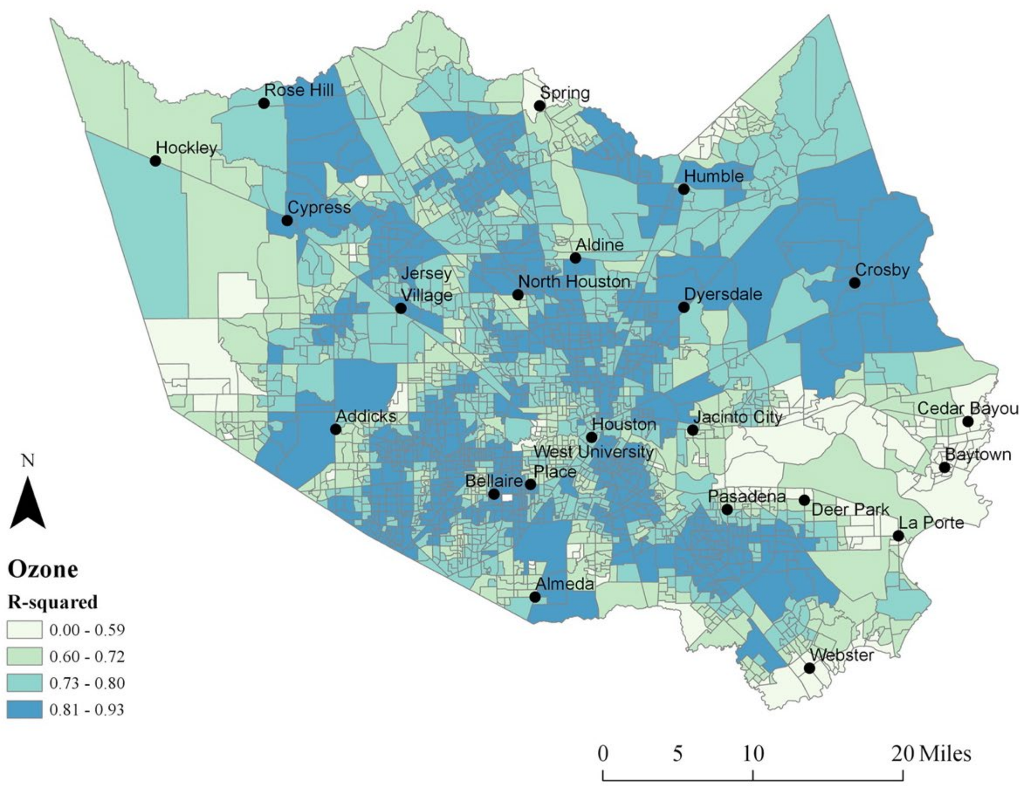

Figure 7. GWR model performance of Ozone: distribution of local R-squared values by census block group.

Traffic Proximity

As shown in Table 6, the coefficient for all independent variables is greater than 1 or less than -1, representing there are errors in the analysis. The original GWR analysis was completed with a warning message showing the final model didn't have the lowest AICc value. After conducting the GWR analysis again using the number of neighbors with the lowest AICc (provided in the original results), the model was still unable to generate a model with coefficient values between 1 and -1 (Table 7), and there is not much difference in the results. This failure may be triggered by extremely limited variation in the traffic proximity data itself. See the distribution of traffic proximity in Figure 3. Therefore, a clear correlation between traffic proximity and the independent variables could not be decided.

A map of the spatial distribution of local R-squared values generated by GWR analysis is presented in Figure 8, which shows a better performance in block groups located in neighborhoods just outside of downtown Houston, and block groups near Rose Hill and Webster. However, the relationships between traffic proximity and the independent variables remain doubtful due to the unreliable results.

Table 6. Geographically Weighted Regression (GWR) of Traffic Proximity using golden search method.

	GWR coefficients			Percent of census block group by significance (95% level) of t-statistic		
	Min	**Median**	**Max**	**t ≤ -1.96**	**-1.96 < t < 1.96**	**t ≥ 1.96**
Minority	-12913.90	214.72	9435.05	3.22%	90.99%	5.79%
Low Income	-7155.88	216.00	14596.13	1.82%	94.49%	3.69%
Under Age 5	-42296.14	-333.98	30247.68	7.93%	88.57%	3.50%
Over Age 64	-24240.77	-295.30	16255.02	7.42%	90.25%	2.33%
Intercept	-7416.56	298.84	11659.87	0.28%	86.75%	12.97%
R-squared	-1.82	0.31	0.72			

Table 7. Geographically Weighted Regression (GWR) of Traffic Proximity using the number of neighbors with lowest AICc.

	GWR coefficients			Percent of census block group by significance (95% level) of t-statistic		
	Min	**Median**	**Max**	**t ≤ -1.96**	**-1.96 < t < 1.96**	**t ≥ 1.96**
Minority	-13416.16	215.66	9762.59	3.27%	90.90%	5.83%
Low Income	-7198.28	211.07	14647.55	1.82%	94.49%	3.69%
Under Age 5	-42529.63	-330.84	30813.35	7.89%	88.66%	3.45%
Over Age 64	-24900.32	-297.52	16733.97	7.42%	90.15%	2.43%
Intercept	-7626.90	292.05	11885.96	0.23%	86.79%	12.97%
R-squared	-1.80	0.31	0.72			

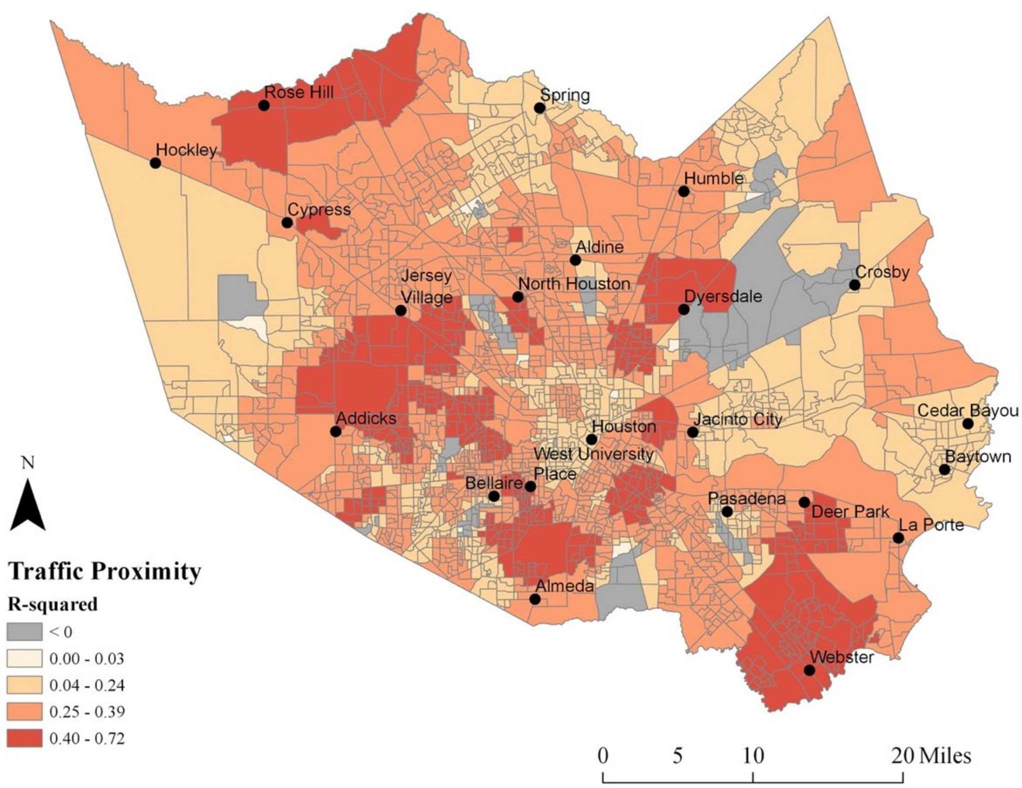

Figure 8. GWR model performance of Traffic Proximity: distribution of local R-squared values by census block group.

NATA Respiratory Hazard Index

The numerical results associated with the GWR analysis of NATA respiratory hazard index are summarized in Table 8. There are strong correlations between NATA respiratory hazard index with block groups with a higher percentage of people of color, low-income population, children under 5, and elders above 64. To examine the extent of improvement that GWR can achieve by applying local weighting to census block groups compared to the traditional multivariate regression based on the ordinary least squares (OLS) method, the same dataset was utilized to create the OLS regression model. The results are summarized in Table 9.

Table 8. Geographically Weighted Regression (GWR) of NATA Respiratory Hazard Index.

	GWR coefficients			Percent of census block group by significance (95% level) of t-statistic		
	Min	**Median**	**Max**	**t ≤ -1.96**	**-1.96 < t < 1.96**	**t ≥ 1.96**
Minority	-7.53	0.02	3.94	6.77%	85.58%	7.65%
Low Income	-1.71	0.04	2.55	4.11%	89.78%	6.11%
Under Age 5	-7.38	-0.17	7.33	5.65%	90.39%	3.97%
Over Age 64	-5.69	-0.09	10.39	8.17%	87.91%	3.92%
Intercept	-0.83	2.15	9.78	0.00%	5.23%	94.77%
R-squared	-3.75	0.60	0.97			

Table 9. Ordinary least squares regression of NATA Respiratory Hazard Index.

	Coefficient	**t-statistic**	
Minority	0.20	4.02	(p < 0.0001)
Low Income	0.22	3.77	(p = 0.0002)
Under Age 5	0.12	-0.51	(p = 0.6127)
Over Age 64	0.12	0.87	(p = 0.3857)
Intercept	1.97	54.29	(p < 0.0001)
R-squared		0.058	
Adj. R-squared		0.056	
N		2143	

It is important to note that traditional multivariate regression model is a global model which assumes the relationships between NATA respiratory hazard index and each independent variable are equal in all census block groups across Harris County and is unable to provide local variations in goodness of fit and model coefficients. While the traditional regression model left more than 94% of the variance in NATA respiratory hazard index unexplained (adjusted R-squared = 0.056), the median R-squared produced by the GWR model was 0.60, indicating a large improvement in explained variance. Local models of GWR associated with approximately 95% of census block groups in Harris County indicated an improvement over the adjusted R-squared of 0.056 from the global model of traditional multivariate regression.

The spatial distribution of local R-squared values generated by GWR analysis of respiratory hazard index (Figure 9) demonstrates an overall pattern of higher R-squared values in block groups located along main roads and highways across Harris County.

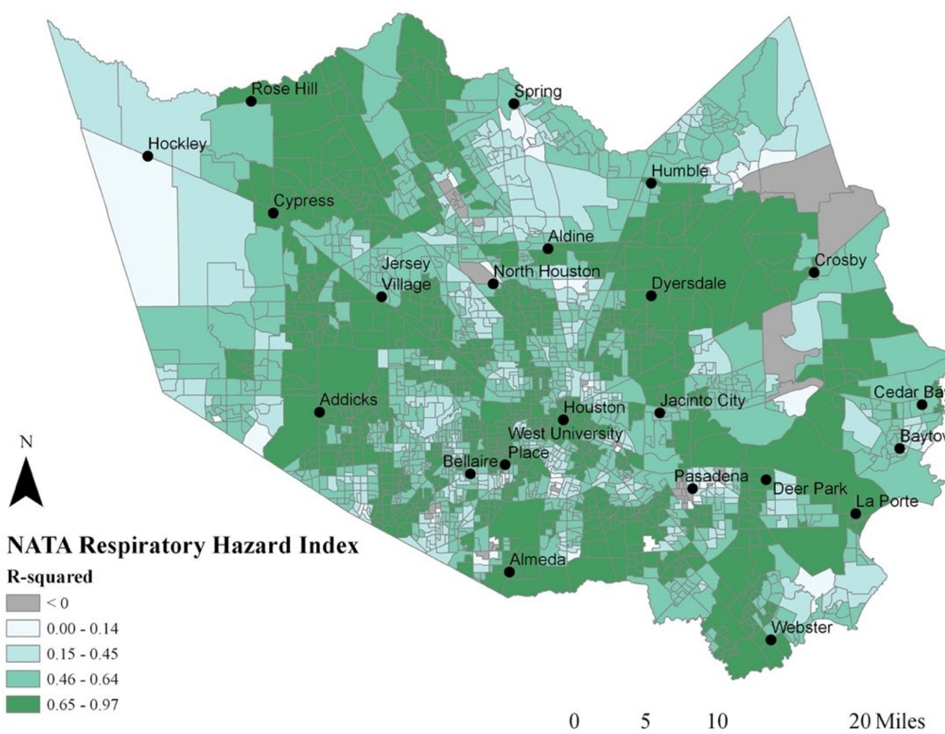

Figure 9. GWR model performance of NATA Respiratory Hazard Index: distribution of local R-squared values by census block group.

However, R-squared value only determines how well the model fits the data. It is unable to show if the relationships between dependent and independent variables are positive or negative. Figure 10 to Figure 13 provide a better understanding of model coefficients by providing the distribution and variation of the correlation coefficient values between NATA respiratory hazard index and four independent variables among local areas.

As shown in the Figure 10, there are stronger positive correlations between NATA respiratory hazard index and minority population in block groups located near Deer Park and La Porte, while the stronger negative correlations can be found in block groups located near Almeda, Aldine, and between Dyersdale and Jacinto City. For low income population (Figure 11), the overall correlations are weaker than minority as most of the block groups has a relatively low coefficient value than minority population. The spatial distribution of the weak correlations is similar to minority population correlations, but the spatial distribution of stronger positive and negative correlations is opposite to the correlations of minority population. In Figure 12, block groups with higher positive coefficient values of children under age 5 are located along main roads in Harris County, especially near the George Bush International Airport on the north and Cedar Park on the west, while the block groups with higher negative coefficient values are located near downtown, Baytown and outer suburban areas of Harris County. For elders over 64 year-old (Figure 13), the block groups located near the George Bush International Airport on the north, Cypress and Hockley on the northwest, and Katy on the west have the highest positive coefficient values while the block groups located near Jacinto City, Dyersdale and Cedar Bayou have the lowest negative coefficient values.

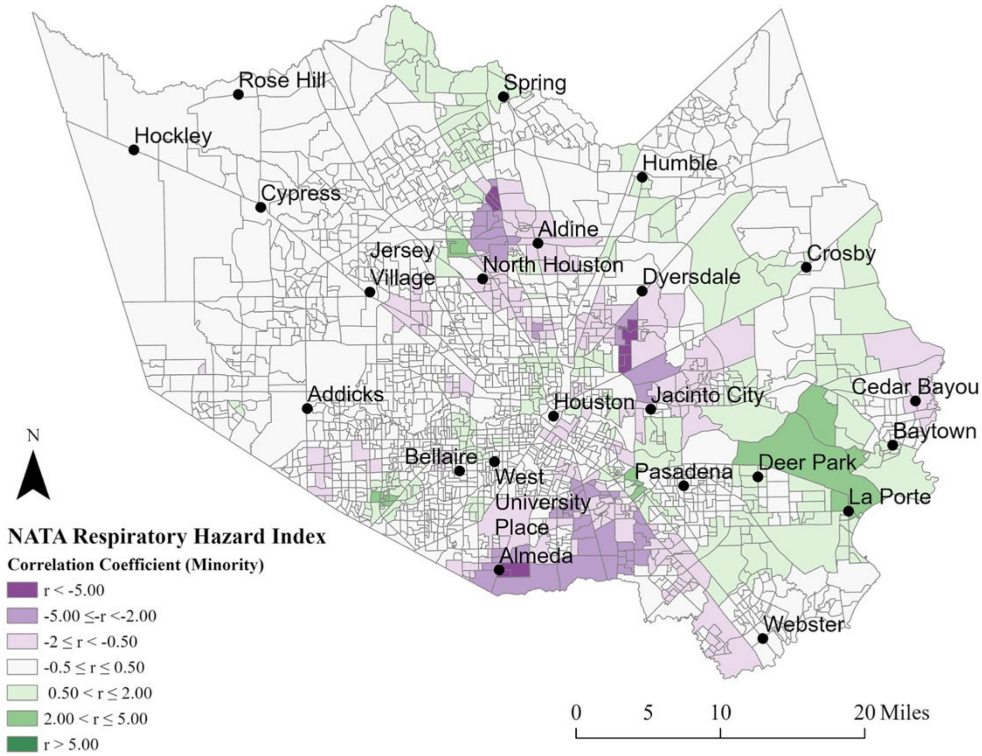

Figure 10. GWR model coefficients of NATA Respiratory Hazard Index: distribution of local correlation coefficient values of minority population by census block group.

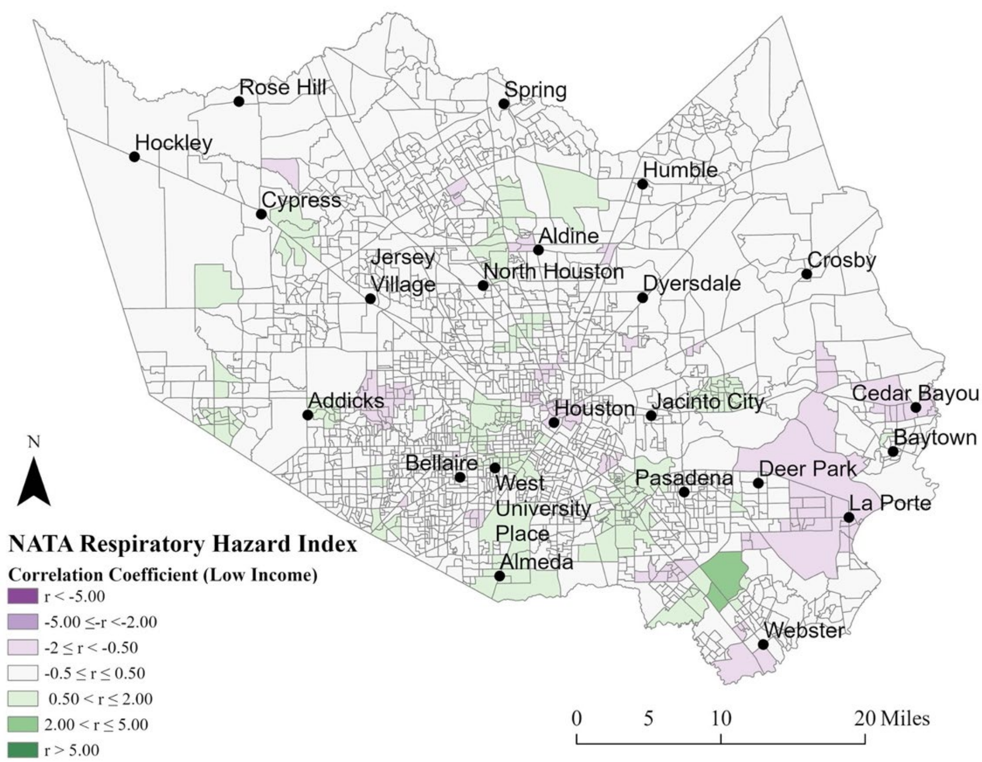

Figure 11. GWR model coefficients of NATA Respiratory Hazard Index: distribution of local correlation coefficient values of low income population by census block group.

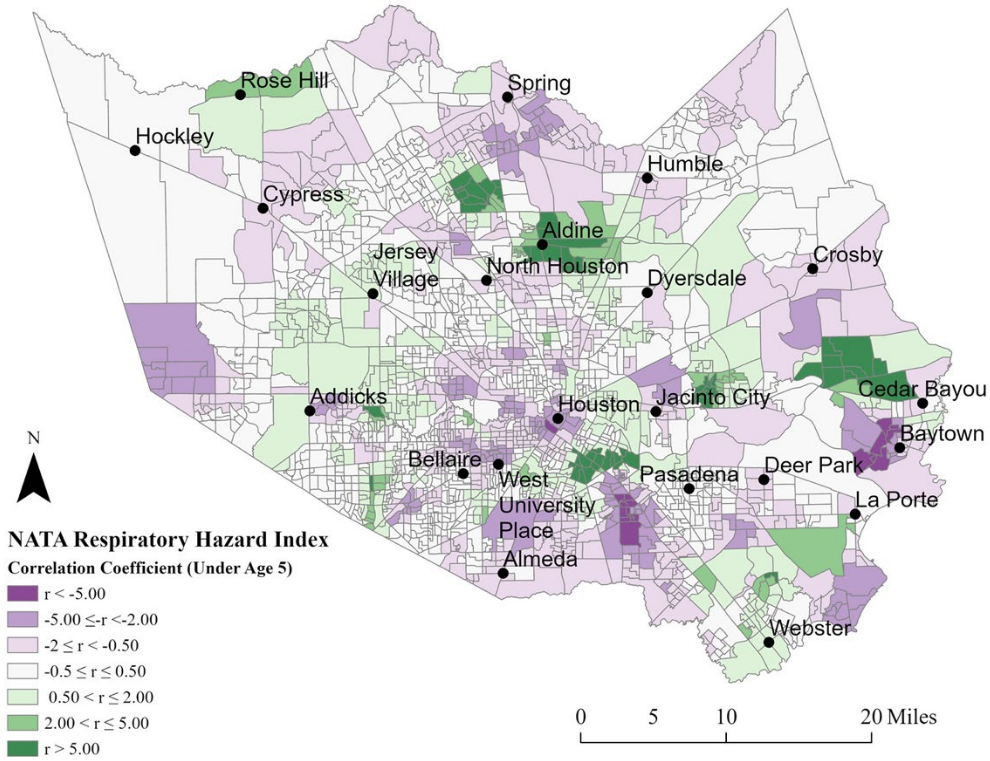

Figure 12. GWR model coefficients of NATA Respiratory Hazard Index: distribution of local correlation coefficient values of population under age 5 by census block group.

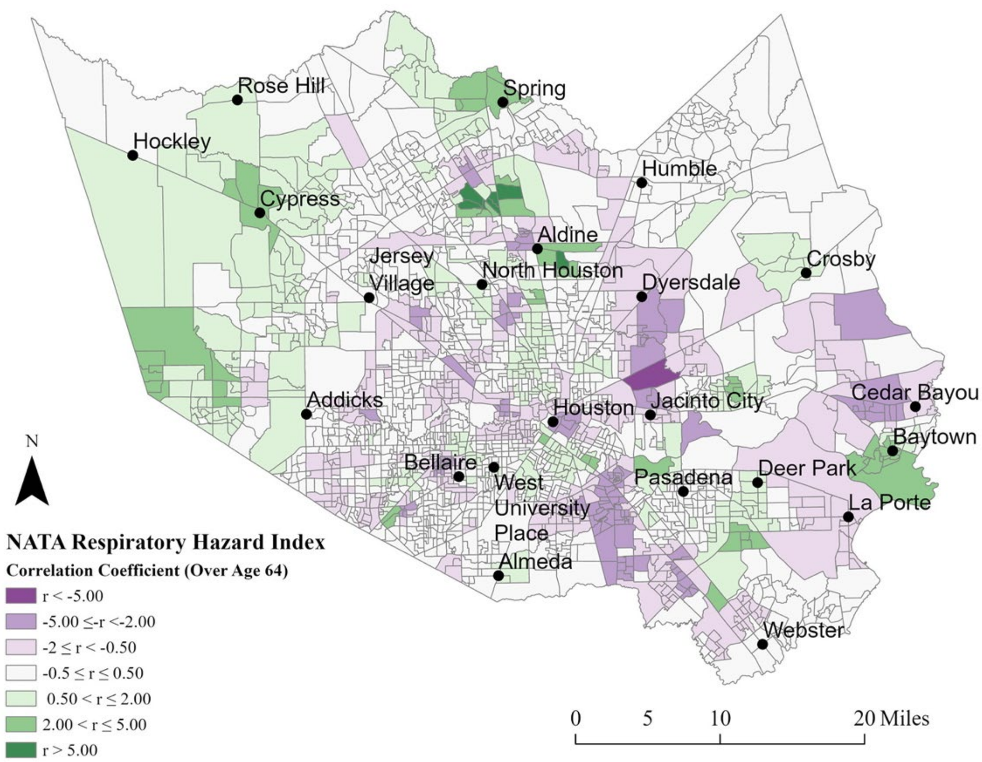

Figure 13. GWR model coefficients of NATA Respiratory Hazard Index: distribution of local correlation coefficient values of population over age 64 by census block group.

NATA Diesel Particulate Matter

The numerical results associated with the GWR analysis of NATA diesel particulate matter are summarized in Table 10. The performance of NATA diesel particulate matter GWR analysis model is slightly poorer than NATA respiratory hazard index as it has more block groups with lower R-squared. However, the spatial variations of local R-squared values generated by GWR analysis (Figure 14) show a distribution pattern similar to the results of NATA respiratory hazard index, higher R-squared values in block groups located along main roads and highways across Harris County.

Table 10. Geographically Weighted Regression (GWR) of NATA Diesel Particulate Matter.

	GWR coefficients			**Percent of census block group by significance (95% level) of t-statistic**		
	Min	**Median**	**Max**	**t ≤ -1.96**	**-1.96 < t < 1.96**	**t ≥ 1.96**
Minority	-5.72	0.11	3.87	5.04%	84.88%	10.08%
Low Income	-2.23	0.09	2.20	4.20%	84.51%	11.29%
Under Age 5	-6.44	-0.08	9.10	8.07%	87.49%	4.43%
Over Age 64	-6.63	-0.10	5.56	10.31%	86.33%	3.36%
Intercept	-1.97	1.06	7.11	0.28%	19.32%	80.40%
R-squared	-0.66	0.55	0.90			

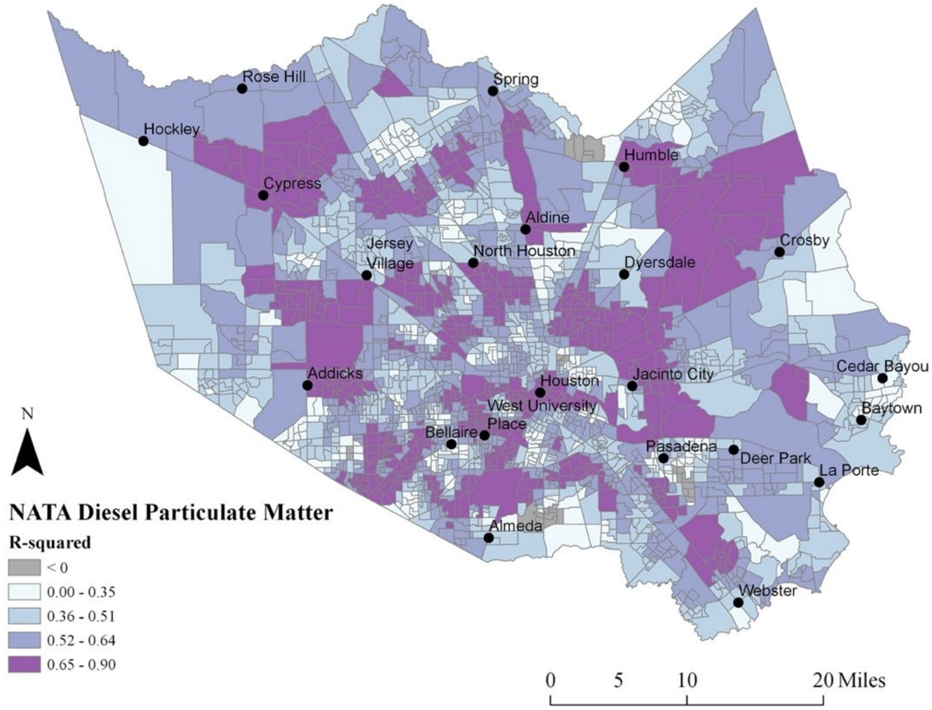

Figure 14. GWR model performance of NATA Diesel Particulate Matter: distribution of local R-squared values by census block group.

V. CONCLUSION

This research demonstrates the geographic variations in the relationship between EJSCREEN environmental indicators and various demographic and socioeconomic factors in Harris County, Texas. The results show high levels of the goodness of fit for most of the models except for traffic proximity due to extremely limited variation in the data. The improvement in R-square values of GWR local models over the performance of traditional global regression based on ordinary least squares also underlines the advantage of GWR as an exploratory data analysis tool for environmental justice assessment. It is important the readers should bear in mind that, however, there are some areas with very low variations after applying the weights in local models except for ozone. Therefore, the results should be interpreted and used with caution. Moreover, this research did not apply all of both environmental and demographical indicators provided in the EJSCREEN data due to limited study scope and multicollinearity. The indicators which were not been used in this research may provide better and more comprehensive explanations for the statistical and spatial variations in relationships between the demographic and environmental variables.

The results of this study have several implications for public policy and can be used to inform that improve the current status of environmental justice in Harris County. In addition to identifying where are the areas with disproportionate air pollution, this study reveals where air pollution is statistically and significantly related to the presence of particular demographic and socioeconomic groups. This information can guide state and local regulatory agencies to promote environmental justice and health protection through the constraint of emission sources such as industrial facility or road construction in areas where vulnerable population groups are currently exposed to disproportionate air pollution. With the capacity of demonstrating spatial differences in the relationship between exposure to air pollution and specific population groups, GWR can provide valuable insights into locally appropriate policies and solutions.

In conclusion, this study demonstrates the utility and advantage of GWR compared to traditional multivariate regression in environmental justice analysis. The results of this study provide a clear representation that there are a significant positive association between a higher volume of air pollutants or a higher level of relevant health risk and a higher percentage of minority populations and therefore we still have a long way to go to achieve environmental justice.

REFERENCES

Ard, K. (2015). Trends in exposure to industrial air toxins for different racial and socioeconomic groups: A spatial and temporal examination of environmental inequality in the U.S. from 1995 to 2004. *Soc Sci Res, 53*, 375-390. doi:10.1016/j.ssresearch.2015.06.019

Bakhtsiyarava, M., & Nawrotzki, R. J. (2017). Environmental Inequality and Pollution Advantage among Immigrants in the United States. *Appl Geogr, 81*, 60-69. doi:10.1016/j.apgeog.2017.02.013

Bell, M. L., & Ebisu, K. (2012). Environmental inequality in exposures to airborne particulate matter components in the United States. *Environ Health Perspect, 120*(12), 1699-1704. doi:10.1289/ehp.1205201

Bethel, H., Sexton, K., Linder, S., Delclos, G., Stock, T., Abramson, S., . . . Ward, J. (2005). *A Closer Look at Air Pollution in Houston: Identifying Priority Health Risks* Retrieved from

Bouvier, R. (2014). Distribution of income and toxic emissions in Maine, United States: Inequality in two dimensions. *Ecological Economics, 102*, 39-47. doi:10.1016/j.ecolecon.2014.03.005

Branis, M., & Linhartova, M. (2012). Association between unemployment, income, education level, population size and air pollution in Czech cities: evidence for environmental inequality? A pilot national scale analysis. *Health Place, 18*(5), 1110-1114. doi:10.1016/j.healthplace.2012.04.011

Brauer, M., Hoek, G., Smit, H. A., de Jongste, J. C., Gerritsen, J., Postma, D. S., . . . Brunekreef, B. (2007). Air pollution and development of asthma, allergy and infections in a birth cohort. *Eur Respir J, 29*(5), 879-888. doi:10.1183/09031936.00083406

Brooks, N., Adger, W. N., & Kelly, P. M. (2005). The determinants of vulnerability and adaptive capacity at the national level and the implications for adaptation. *Global Environmental Change, 15*(2), 151-163. doi:10.1016/j.gloenvcha.2004.12.006

Bruce, N., Perez-Padilla, R., & Albalak, R. (2000). Indoor air pollution in developing countries: a major environmental and public health challenge. *Bull World Health Organ., 78*(9), 1078-1092.

Confronting Environmental Racism: Voices from the Grassroots. (1993). (R. D. Bullard Ed.). Boston, Massachusetts: South End Press.

Crowder, K., & Downey, L. (2010). Interneighborhood Migration, Race, and Environmental Hazards: Modeling Microlevel Processes of Environmental Inequality. *American Journal of Sociology, 115*(4), 1110–1149. doi:10.1086/649576

Cutter, S. (1995). Race, class and environmental justice. *Progress in Human Geography, 19*(1), 111-122. doi:10.1177/030913259501900111

Dockery, D. W., Pope, C. A., Xu, X., Spengler, J. D., Ware, J. H., Fay, M. E., . . . Speizer, F. E. (1993). An Association between Air Pollution and Mortality in Six U.S. Cities. *The New England Journal of Medicine, 329*, 1753-1759. doi:10.1056/NEJM199312093292401

Downey, L., Crowder, K., & Kemp, R. J. (2017). Family Structure, Residential Mobility, and Environmental Inequality. *J Marriage Fam, 79*(2), 535-555. doi:10.1111/jomf.12355

Downey, L., & Hawkins, B. (2008). Race, Income, and Environmental Inequality in the United States. *Sociol Perspect, 51*(4), 759-781. doi:10.1525/sop.2008.51.4.759

Duijts, L. (2012). Fetal and infant origins of asthma. *Eur J Epidemiol, 27*(1), 5-14. doi:10.1007/s10654-012-9657-y

Feagin, J. (1985). The Global Context of Metropolitan Growth: Houston and the Oil Industry. *American Journal of Sociology, 90*(6), 1204-1230. Retrieved from https://www.jstor.org/stable/2779634

Gilbert, A., & Chakraborty, J. (2011). Using geographically weighted regression for environmental justice analysis: Cumulative cancer risks from air toxics in Florida. *Social Science Research, 40*(1), 273-286. doi:10.1016/j.ssresearch.2010.08.006

Gong, X., Lin, Y., Bell, M. L., & Zhan, F. B. (2018). Associations between maternal residential proximity to air emissions from industrial facilities and low birth weight in Texas, USA. *Environ Int, 120*, 181-198. doi:10.1016/j.envint.2018.07.045

Gotschi, T., Heinrich, J., Sunyer, J., & Kunzli, N. (2008). Long-term effects of ambient air pollution on lung function: a review. *Epidemiology, 19*(5), 690-701. doi:10.1097/EDE.0b013e318181650f

Jephcote, C., & Chen, H. (2012). Environmental injustices of children's exposure to air pollution from road-transport within the model British multicultural city of Leicester: 2000-09. *Sci Total Environ, 414*, 140-151. doi:10.1016/j.scitotenv.2011.11.040

Jin, H., & Lu, Y. (2017). The relationship between obesity and socioeconomic status among Texas school children and its spatial variation. *Applied Geography, 79*, 143-152. doi:10.1016/j.apgeog.2016.12.008

Lee, K. H., Dvorak, R. G., Schuett, M. A., & van Riper, C. J. (2017). Understanding spatial variation of physical inactivity across the continental United States. *Landscape and Urban Planning, 168*, 61-71. doi:10.1016/j.landurbplan.2017.09.020

Lersch, K. M., & Hart, T. C. (2014). Environmental Justice, Lead, and Crime: Exploring the Spatial Distribution and Impact of Industrial Facilities in Hillsborough County, Florida. *Sociological Spectrum, 34*(1), 1-21. doi:10.1080/02732173.2014.857184

Maantay, J., & Maroko, A. (2009). Mapping Urban Risk: Flood Hazards, Race, & Environmental Justice In New York". *Appl Geogr, 29*(1), 111-124. doi:10.1016/j.apgeog.2008.08.002

Mannucci, P. M., & Franchini, M. (2017). Health Effects of Ambient Air Pollution in Developing Countries. *Int J Environ Res Public Health, 14*(9). doi:10.3390/ijerph14091048

Mennis, J. L., & Jordan, L. (2005). The Distribution of Environmental Equity: Exploring Spatial Nonstationarity in Multivariate Models of Air Toxic Releases. *Annals of the Association of American Geographers, 95*(2), 249-268. doi:10.1111/j.1467-8306.2005.00459.x

Miller, M. D., & Marty, M. A. (2010). Impact of environmental chemicals on lung development. *Environ Health Perspect, 118*(8), 1155-1164. doi:10.1289/ehp.0901856

Mohai, P., Pellow, D. N., & Roberts, J. T. (2009). Environmental Justice. *Annual Review of Environment and Resources, 34*(1), 405-430. doi:10.1146/annurev-environ-082508-094348

Peel, J. L., Metzger, K. B., Klein, M., Flanders, W. D., Mulholland, J. A., & Tolbert, P. E. (2007). Ambient air pollution and cardiovascular emergency department visits in potentially sensitive groups. *Am J Epidemiol, 165*(6), 625-633. doi:10.1093/aje/kwk051

Qian, Z. (2010). Without zoning: Urban development and land use controls in Houston. *Cities, 27*(1), 31-41. doi:10.1016/j.cities.2009.11.006

Ritz, B., & Wilhelm, M. (2008). Ambient air pollution and adverse birth outcomes: methodologic issues in an emerging field. *Basic Clin Pharmacol Toxicol, 102*(2), 182-190. doi:10.1111/j.1742-7843.2007.00161.x

Smiley, K. T. (2019). A Polluting Creed: Religion and Environmental Inequality in the United States. *Sociology of Religion, 62*(6), 980-1000. doi:10.1177/0731121419862229

Sweileh, W. M., Al-Jabi, S. W., Zyoud, S. H., & Sawalha, A. F. (2018). Outdoor air pollution and respiratory health: a bibliometric analysis of publications in peer-reviewed journals (1900 - 2017). *Multidiscip Respir Med, 13*, 15. doi:10.1186/s40248-018-0128-5

Wisner, B., Blaikie, P., Blaikie, P. M., Cannon, T., & Davis, I. (2004). *At Risk: Natural Hazards, People's Vulnerability and Disasters*: Routledge.