11 The Spatial Imperatives of Environmental Justice

Trevor Fuller, Department of Geography, Geology, & Anthropology, Indiana State University, Terre Haute, Indiana, USA

Jay D. Gatrell, Department of Geography, Geology, & Anthropology, Indiana State University, Terre Haute, Indiana, USA

E. LaFary, School of Geography & Environmental Sciences, University of Auckland

The nexus between environmental justice and geo-technologies is an evolving one. That is to say, geographic information systems, remote sensing, and other technologies have the capacity to locate and situate the politics and place-based dangers of environmental risk within a broader conceptual and policy framework. Conceptually, GIScience has the capacity to chart new geographies of environmental risk across the urban and rural landscape. Empirically, GIScience has the capacity to map heretofore disparate datasets in an attempt to unlock the socio-economic determinants of "who's at risk and where?" In this paper, we build on the earlier work of Buzzelli to explore the socio-spatial dynamics of environmental risk in Terre Haute, Vigo County, Indiana. Using GIS, remote sensing, census, and environmental data, the paper presents a framework for unlocking the spatial dynamics of socioeconomic status and environmental risk across urban and rural neighborhoods in Vigo County.

11.1 Study Area

The study area is located in the state of Indiana located within the United States. Situated on the banks of the Wabash River, Terre Haute, Indiana is the county seat of Vigo County (Fig. 1). Terre Haute had a 2000 population of 69,614 with an observed county wide median income of $33,184 and a median housing value of $72,500 (U.S. Census, 2002). There is considerable variety in the land use encountered in Terre Haute and Vigo County with dense and mixed urban, parks, suburban, and rural/agricultural regions present. In this respect, Terre Haute and Vigo County are typical of moderate Midwestern metropolitan areas.

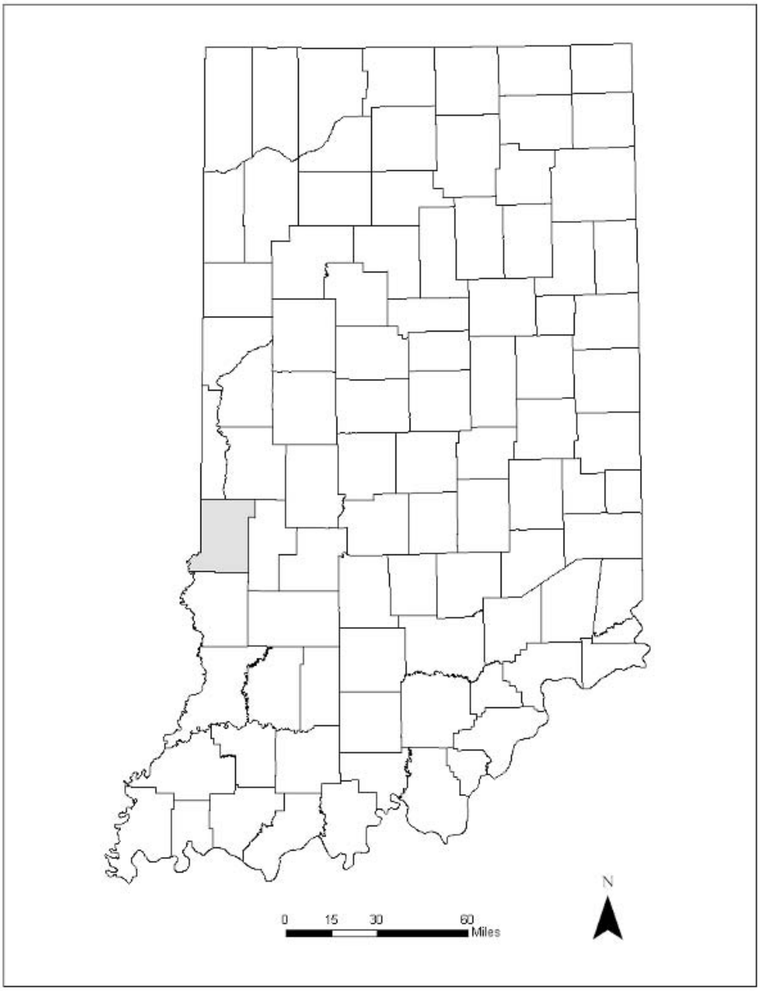

Fig. 1. Vigo County, Indiana

11.2 Placing & Scaling Environmental Justice

Environmental justice has been defined many ways over the last 30 years. The dominant narrative suggests that specific populations, particularly marginalized groups, are being subjected to a disproportionate amount of risk from environmental disamenities. Disamenities, as used here, refer to commonly used indicators of environmental quality, such as the location of certain types of facilities and spills or releases to the environment. Most often in such research the differentiation among the population occurs by means of socioeconomic/demographic characteristics. The United States Environmental Protection Agency (U.S. EPA) defines environmental justice as "the fair treatment and meaningful involvement of all people regardless of race, color, national origin, or income with respect to the development, implementation, and enforcement of environmental laws, regulations, and policies." (U.S. EPA Office of Environmental Justice 2006).

The wide variability in the strength of correlation between socioeconomic status and environmental disamenities further fuels the controversy as to whether certain populations do indeed bear a disproportionate amount of environmental risk (Cutter, et al. 2001). Such variability is exemplified in research that has indicated people of color were disproportionately exposed, especially working-class Latinos (Pulido 2000), while other research demonstrated a strong relationship between environmental risk and dwelling value, as well as lone-parent families (Buzzelli, et al. 2003). Such mixed results have been revealed over the years, with significant shifts in the relative role of demographic and socioeconomic conditions in determining disproportionate environmental risk (Cutter, et al. 2001). Some research has not revealed any direct relationship between minority populations and disproportionate environmental risk (Anderton et al. 1994). In addition to the changing place of demographics within research, another contributor toward the varying results was the wide array of study areas used in environmental justice research. In the following paragraphs we will discuss the issue of an appropriate scale of analysis.

The area of analysis has varied widely within environmental justice research, with much of the research focusing on a city-wide analysis (Mohai and Bryant 1992, Buzzelli, et al. 2003, Pulido 2000). Some research in the environmental justice field has been designed to model environmental risk at the county, state, and even national level (Margai 2001, Pastor et al. 2001). As can be expected, the findings within environmental justice research have therefore not only been highly variable, but contradictory as well. Early research regarding environmental justice often focused on a larger area such as the zip code, in order to examine distribution of risk (United Church of Christ 1987). Later investigations revealed the phenomenon of ecological fallacy, in which the heterogeneity of a particular area of study is often missed due to the area being too large (Anderton et al. 1994). In Anderton, et al. (1994), researchers utilized the census tract as the area of analysis in an attempt to capture the heterogeneity present within the study area. In an attempt to reduce the risk of ecological fallacy for this project we used a smaller area of analysis, the U.S. Census block group, which was the smallest area at which we could still obtain critical Census data. The challenge with selection of area of analysis for this project as with all environmental justice research is the development of a model which efficiently and effectively characterizes any disproportionate amount of environmental risk endured by any particular segment(s) of the population.

11.3 GIScience: GIS, RS & GWR

GIS has established itself as a tool well-suited for spatial analysis of environmental quality investigations, such as assessing questions of environmental justice. As environmental justice examines the geographical distribution of both status and risk, the benefits of using GIS are apparent. With state and federal government agencies realizing the importance of geographical data, a wealth of information has become available, including the locations of various facilities or sites which have been subjects of government enforcement. Such location data has proven useful when assessing questions of residential proximity to environmental risk or the siting of various facilities known or perceived to create environmental risk. Remote sensing technologies have also proven their effectiveness at revealing relationships perhaps otherwise not seen, such as that of quality of life and vegetation (Gatrell and Jensen 2002). In general, data gathered using remote sensing software can be combined with GIS data for effective modeling of environmental issues (Longley 2002). The combination of these technologies is what has been used here in order to compare data gathered by both GIS and remote sensing technologies.

In addition to the combination of GIS and remote sensing, another aspect of this research, which is discussed in greater detail later in this chapter, is the challenge of effectively representing the statistical interactions of risk and status across the study area. To address this issue we used, in part, the statistical technique known as geographically weighted regression (GWR). Whereas standard regression provides global statistics implying uniformity across space, GWR effectively calculates local statistics at regression points across a study area, which aids in visualization of phenomena (Fotheringham, et al. 2002). Given the variability in environmental justice research results, with all sizes of study areas considered, GWR is an important tool in examining the variation of environmental risk across and throughout the U.S. Census Block Groups. Indeed, GWR may unlock heretofore unseen relationships and/or problematize existing assumptions. The following pages will provide further insight into the uses of GIS, remote sensing, and GWR to investigate issues which lay at the intersection of humans and their environment.

11.4 Data and Methods

The primary objective of the investigation discussed here was to assess the relative efficacy of both environmental quality data and a normalized difference vegetation index (NDVI) as metrics of socioeconomic conditions. To follow is a discussion of the methods used, including the environmental quality data, socioeconomic variables, and statistical techniques, as well as a discussion of the creation of the NDVI for the study area, Vigo County, Indiana.

11.4.1 Environmental Data Sets

The United States Environmental Protection Agency (U.S. EPA) has required reporting of certain information under the guidance of environmental regulations for several decades. This information has provided extensive data sets for research relating to environmental quality. The first data set used for this investigation was the EPA's Toxics Release Inventory Program (TRI), which includes information regarding reported releases from regulated facilities throughout the United States. In particular, releases to air, soil, and surface water were used by first asking whether there has been a release, answered with a yes or no, and then adding the amounts released (air, soil, and water) to make one reported number or quantity. In this way, there was no differentiation between routes of release. Rather, the total amount of released contaminants from each facility or site is used. By not parceling out the release information by medium, we avoided an investigative slippery slope regarding the route of release, which leads to a consideration of the medium, meteorological, and hydrological factors.

Treatment, storage, and disposal facilities (TSDFs) databases were the second environmental quality data source used. TSDFs are regulated under the EPA's Resource Conservation and Recovery Act (RCRA), which in part, was designed to monitor the flow of hazardous waste from generation through to the time of disposal, a process commonly referred to as "cradle to grave".

The third environmental quality data set included the locations of Superfund sites within Vigo County, Indiana. This data consists of sites that are currently on the U.S. EPA's National Priorities List (NPL). Sites are placed on the NPL after regulatory officials investigate each site by following the Superfund cleanup process, beginning with notification to EPA of possible releases of hazardous substances. After each site is investigated it is either designated as needed no further remedial action or it is proposed for placement on the NPL.

TRI, TSDF, and Superfund data are location-based in their application to environmental justice research. The proximity of such facilities to particular communities or segments thereof is interpreted by many researchers as an indication of environmental risk, usually disproportionately distributed among the study area population. The TRI data was acquired from the U.S. EPA via its online data download library. The information is provided in the form of ESRI shapefiles and associated files, which was imported into ESRI's ArcMap software for display and analysis. TSDF data and Superfund site data were acquired from the Indiana Department of Environmental Management via the Indiana Geological Survey's online GIS data download library.

The fourth data set used provides levels of the metal lead (Pb) found in the blood of children within Vigo County, provided in the form of number of children within each zip code whose blood-lead levels were above a previously set criteria level. This data set could not only be a potential indicator of environmental quality, but it also represents actual human exposure to environmental contamination, as opposed to the other environmental quality data sets used here, which reflect a potential for environmental risk. Blood-lead level data were acquired from the Vigo County Health Department in the form of a hard-copy spreadsheet, with the data then being entered into a computer-based database and imported into ESRI's ArcMap software. The data set provides each zip code, as opposed to census block group, where an elevated level was revealed. In order to use this data at the block group level, the mean blood-lead level of results for those zip codes affected was assigned for each block group within that zip code.

11.4.2 NDVI

A NDVI map was created with the use of the remote sensing software ERDAS Imagine 8.7 (ERDAS) and a satellite-produced image of Vigo County. The satellite image was produced by Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) using the Terra satellite, yielding a spatial resolution of 15-meters in the near-infrared and visible portions of the electromagnetic spectrum. Within ERDAS a NDVI was calculated by incorporating the near-infrared and red channels into the following formula:

$$\frac{\text{Near-infrared} - \text{red}}{\text{Near-infrared} + \text{red}}$$

NDVI is based upon the principle that the red (visible) portion of the electromagnetic spectrum is highly absorbed by chlorophyll present within plants or vegetation, while the near-infrared energy is reflected at high levels by a plant's mesophyll leaf structure (Tucker 1979). The calculated vegetation index then indicates the relative strength or reflectance of vegetation throughout the satellite image of the study area. A higher NDVI value indicates a more robust presence of vegetation. NDVI is unique in that it normalizes the various reflectance values by converting them to a value between –1 and 1 for each pixel in the image, with –1 representing no vegetation and 1 indicating robust vegetation. This investigation examines a NDVI of Vigo County, Indiana to determine its efficacy as a metric for socioeconomic status, and then compares the resulting capacity to that of the environmental quality data. Specifically, NDVI variables used in this analysis were the following:

Standard deviation of NDVI values within a block group;
Minimum NDVI value observed within a block group;
Maximum NDVI value observed within a block group;
Range of NDVI values observed within a block group;
Interaction of NDVI with population density; and,
Mean NDVI value

These values were assigned to each of the census block group polygons within the spatial database.

11.4.3 Socioeconomic/Demographic Characteristics

Specific socioeconomic and demographic variables selected for use in this investigation have been applied in much of the earlier research regarding environmental justice. Such variables are often used as indicators of socioeconomic status. The following socioeconomic and demographic variables were acquired from the U.S. Census Bureau's online data library for the year 2000 and integrated into this analysis as indicators of socioeconomic status: (1) Median Household Income; and (2) Median Household Value.

11.5 Methods

This investigation uses three approaches: correlation, weighted least squares regression, and geographically weighted regression. Correlation—Pearson's R—was used to explore the relationships between variables and the significance of these variables. Using the Pearson's R results as a guide, weighted least squares regression models were tested using both enter and step-wise approaches. The weighted least squares regression was performed using population density as the weighting variable.

Geographically weighted regression was used as standard regression statistical techniques often treat phenomena as occurring equally across a study area. As Fotheringham, et al. (2002) discussed, spatial data often exhibit what has been termed spatial nonstationarity, or the nonuniform distribution of spatial information. The benefit of GWR in geographical research is that it accounts for unique characteristics of spatial data by calculating the necessary statistical measures at each point in the study area, which provides individual level or point-unique statistical information, allowing a researcher to identify disparities in the spatial distribution of various phenomena. GWR served this research well given that previous research has demonstrated the spatial nonstationarity of disproportionate environmental risk (Mennis and Jordan 2005).

The model for this investigation included the following variables analyzed through OLS regression and GWR, as well as analysis using Pearson's Correlation between the Socioeconomic and environmental metrics.

The variables used were:

Median Household Income
Median Household Value
U.S. EPA Toxics Release Inventory
RCRA TSDFs
Superfund Sites
Child blood-Lead levels (BLL) by Zip code
NDVI Minimum
NDVI Maximum
NDVI Mean
NDVI Range
NDVI Standard Deviation
Population Density

11.5.1 Interaction terms

In addition to the variables listed above, an interaction term was created using the expansion method (Casetti 1972, Gatrell and Bierly 2002, Jensen, et al. 2005). The expansion method developed by Emilio Casetti was an early challenge to the existing statistical paradigm that assumed spatial relationships are constant across a study area (Gatrell, J., Chapter 5 of Jensen, et al. 2005). Casetti (1972) also attributed the nonstationarity of spatial phenomena to the interaction of terms across space. We relied upon this interaction of terms as we created a model containing a multiplied interaction of NDVI data with observed population density for each block group within the study area. Population density was used as it has been shown to be effective when modeling environmental parameters in an urban environment.

11.6 The Models

Below the models are presented. The models presented were subjected to OLS, stepwise, and GWR. The study models are:

$$Y = \beta_0 + \beta TR(u,v) \underset{\cdot}{+} \beta TF(u,v) \underset{\cdot}{+} \beta B(u,v) + \beta S(u,v) \underset{\cdot}{+} \beta Sd\,(u,v) + \beta Min(u,v) \underset{\cdot}{+} \beta Mx(u,v) + \beta A(u,v) + \beta R(u,v) + \beta I(u,v) + \varepsilon(u,v)$$

where:

Y is the dependent variable (socioeconomic status), in this

case, either median household income or median household value;
β_0 is the constant;
TR is the U.S. EPA's Toxics Release Inventory data for Vigo County;
TF is IDEM Treatment, Storage, and Disposal Facilities;
B is the blood-lead level in children for Vigo County during 2000-2005;
S is the Superfund facility data;
Sd is the standard deviation of the NDVI values;
Min represents the minimum NDVI value;
Mx is the maximum NDVI value;
A is the mean NDVI value ;
R is the range of NDVI values;
I refers to interaction terms using population density and A; and,
ε refers to the statistical noise assumed to be present in the calculation;

The formula for GWR is the following:

$$y_i = \beta_0(u_i,v_i) + \Sigma_k\beta_k(u_i,v_i)x_{ik} + \varepsilon_i$$

where:

y_i is the dependent variable at location I;
β_0 is an independent variable;
(u_i,v_i) is the coordinate location for the ith point;
$\beta_k(u_i,v_i)$ is the function continuously measuring parameter values at each point I; and,
ε_i is the noise associated with each point i
(Fotheringham, et al. 2002).

11.7 Results

A challenge to effective statistical analysis of geographical relationships is spatial nonstationarity, or the discontinuity of relationships among and between geographical cases or phenomena throughout a study area (Fotheringham et al. (2002). While the process of WLS does capture variability across space as driven by varying population densities, "global" WLS does so based on discrete points rather than across a continuous surface (Fotheringham, et al. 2002). For this reason, GWR calculates local statistics, specifically local r-square values, to determine the model performance in "place" and across "space". In this paper, we use local r-square values derived from GWR to visualize, or map, the spatial dynamics and model performance across the study area. WLS and GWR 3.0 yielded both global and local coefficients of determination. WLS regression was performed on the data, using population density as the weighting variable, in order to evaluate the Pearson's correlation values. We first examined the distribution of the relationships between socioeconomic conditions and environmental quality data using WLS regression. WLS indicated a very weak relationship between both median household income and median household value and environmental disamenities. WLS was able to discern variability in that relationship across space within Vigo County, but the overall relationships were quite weak. Local r-square values generated within the GWR software were mapped to provide a visual reflection of the data (Fig. 2 and 3). GWR was used to determine whether there was spatial nonstationarity among the relationship(s) between socioeconomic conditions and environmental disamenities.

When examining median household income using WLS, all four of the environmental quality variables received correlation values of .05 or lower, with two of the four having negative values. The most closely correlated variables to median household income were the NDVI maximum value (.412) and the NDVI mean value (.440). When regressing the median household value data against the independent variables, the standard deviation of the NDVI as well as the NDVI mean value displayed the strongest correlation to household value at .320 and .283, respectively (See Tables 1 and 2). With such a drastic disparity between the roles of environmental quality variables and NDVI variables, it is apparent that within Vigo County, the geographical distribution of traditional environmental quality indicators do not statistically account for the observed variation in and/or spatial arrangement of median household income and property valu**es.**

Table 1. Diagnostics (Enter Method)

	Household Income	Household Value
Constant	-38900.584	-40424.255
	(-2.166)**	(-.870)
TR	-.002	-.003
	(.718)	(-.418)
TF	-12504.292	-105772.454
	(-.366)	(-1.198)
B	475.390	1225.025
	(1.424)	(1.419)
S	324.068	66288.976
	(.023)	(1.802)*
Min	33476.614	---
	(1.346)	---
Mx	115678.429	30372.046
	(1.130)	(.105)
Sd	28544.582	818233.782
	(.163)	(1.805)*
A	133434.982	372417.923
	(1.495)	(1.613)
I	1.001	-30.097
	(.035)	(-.410)
R	---	-93881.855
	---	(-1.460)
R-Square	**.234**	**.195**
F-Statistic	**3.662**	**2.911**

* Indicates the variable is significant at the .10 level.
** Indicates the variable is significant at the .05 level.

The step-wise approach provided insight as to which model may be considered more "elegant". In this case, it was the mean NDVI value which created a more elegant model when regressed against median household income. Regarding median household value, the standard deviation NDVI value contributed most to the performance of the model.

Table 2. Diagnostics (Stepwise Method)

	Household Income	Household Value
Constant	-10715.333	21478.996
	(-1.154)	(1.249)
Sd	---	783105.761
	---	(3.633)***
A	230062.271	---
	(5.274)***	---
R-Square	**.193**	**.102**
F-Statistic	**27.811**	**13.201**

*** Indicates the variable is significant at the .01 level.

As illustrated in Table 1, median household income was revealed as having the stronger model, with an r-square value of .234, as compared to that of median household value (.195). In this sense, income is more strongly related to the independent variables. Among the independent variables used here, NDVI variables possessed a stronger relationship to both income and household values, with a significantly weaker showing for the environmental quality variables. These figures provide for a discussion of the driving forces behind the phenomenon known as environmental justice. The environmental quality variables (TRI, TSDF, Superfund, and BLL) do not appear to be significantly related to the distribution of such indicators of socioeconomic status as income and household value. However, what this may indicate is less a case of environmental risk seeking out poor populations than wealthier populations seeking amenities, such as vegetation or “greenness”.

When the local r-square values for median household income, generated by GWR 3.0, were mapped (Fig. 2), there was a clear relationship within block groups of the urban core of Terre Haute between median household income and the independent variables used. This was not surprising as the urban center of Terre Haute contains block groups with some of the lowest income population. In addition, the model performed strongly in the northwest and southeast block groups, which are predominantly rural areas.

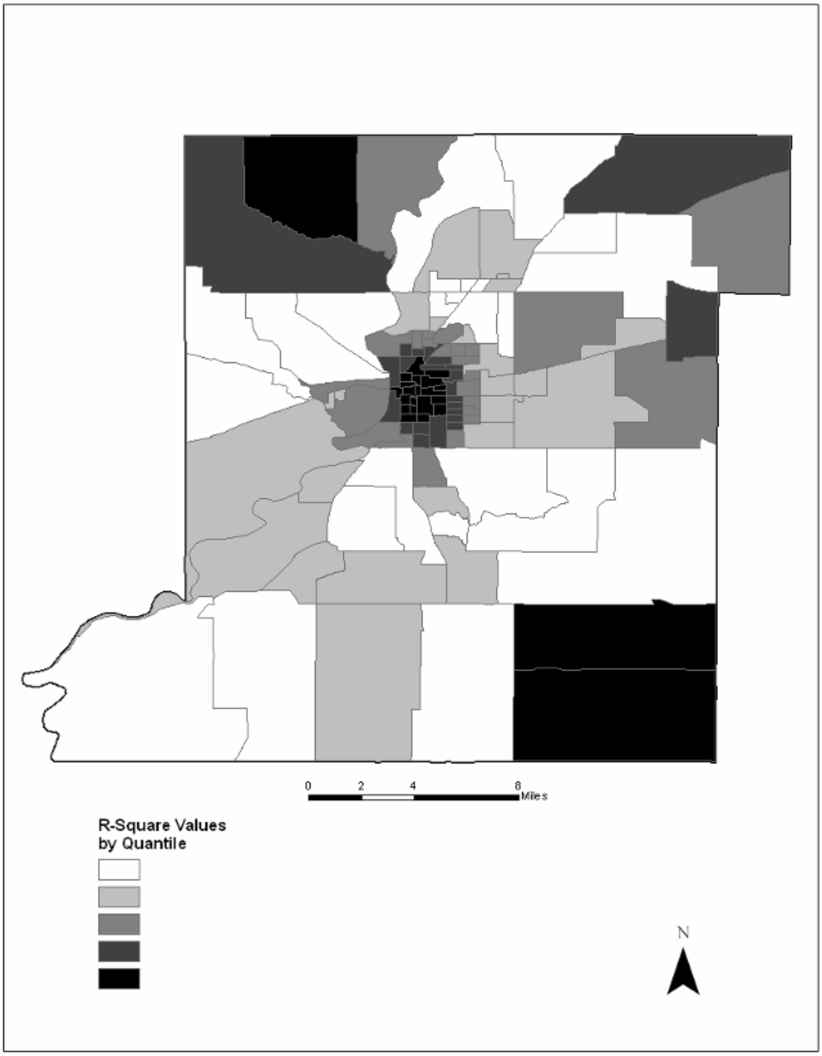

Fig. 2. Household income local r-square values by quantile.

A somewhat similar pattern emerged within the central (urban) and north central/northwest block groups when examining the r-square values for median household value. A strong relationship was observed in the central and northwest areas, but the difference between the mapped results of the two variables was the lack of any significant local r-square values for household value in the southeast corner of the study area (Fig. 3).

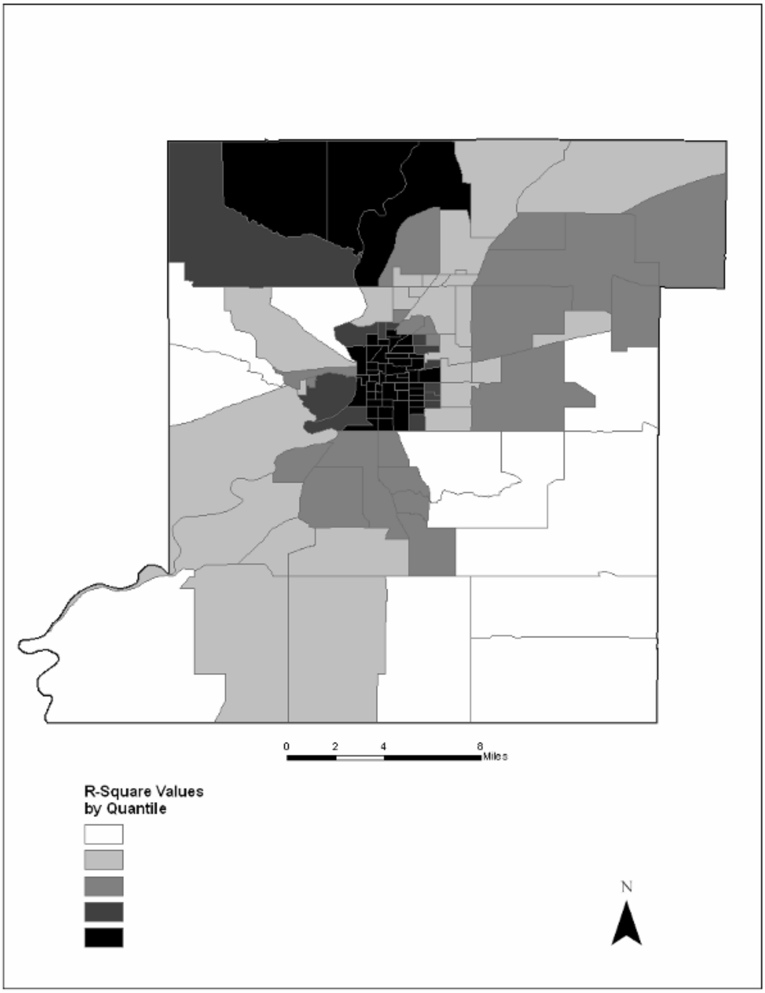

Fig. 3. Household value local r-square values by quantile.

11.8 Discussion

The results—while consistent with the earlier greenness research of Gatrell and Jensen and Jensen et al (2004, 2005)—suggest the environmental justice literature's focus on environmental disamenities may capture only part of the complex interactions that occur within and between social and natural systems in urban environments. That is to say, the basic assertion that the co-location of marginalized groups and environmental disamenities, represents only part of the complete picture. Rather, as this study suggests, the geography of environmental disamenities and socio-economic variables does little to explain the implied relationships between negative externalities and class and race. Instead, the urban environmental geography of class—and perhaps race, too—may be better understood within the context of access to environmental amenities as determined by key proxy variables, such as NDVI. Moreover, the unique geography of Vigo County suggests the distribution of negative externalities are only co-incident—and not necessarily correlated—as "risky sites" occur in a wide range of socio-economic contexts. To that end, the study suggests further research is needed to understand the observed spatial disconnect between the urban geography of amenities and disamenities.

References

Anderton, Douglas L., A. Anderson, J. Oakes, and M. Fraser. 1994. "Environmental Equity: The Demographics of Dumping." *Demography* 31: 229-48.

Buzzelli, Michael, M. Jerrett, R. Burnett, and N. Finklestein. 2003. Spatiotemporal perspectives on air pollution and environmental justice in Hamilton, Canada, 1985-1996. *Annals of the AAG* 93(3): 557-73.

Casetti, E. 1972. Generating models by the expansion method: applications to geographic research. *Geographical Analysis* 4:81-91.

Cutter, S., M. E. Hodgson, and K. Dow, 2001. Subsidized inequities: The spatial patterning of environmental risks and federally assisted housing. *Urban Geography* 22 (1): 29-53.

Fotheringham, A. Stewart, C. Brunsdon, and M. Charlton, 2002. "Geographically Weighted Regression: The Analysis of Spatially Varying Relationships." John Wiley & Sons, Ltd. West Sussex.

Gatrell, J.D. and Bierly, G.D. 2002. Weather and voter turnout: Kentucky primary and general elections, 1990-2000. *Southeastern Geographer* Vol. XXXXII (1): 1-21.

Gatrell, J. D. and R. R. Jensen. 2002. Growth through greening: developing and assessing alternative economic development programs. *Applied Geography* 22:331-350.

Jensen, R., J. Gatrell, J. Boulton, and B. Harper. 2004. Using Remote Sensing and Geographic Information Systems to Study Urban Quality of Life and Urban Forest Amenities. Ecology and Society 9(5): 5. http://www.ecologyandsociety.org/vol9/iss5/art5/

Jensen, R., J. Gatrell, and D. McLean (eds) 2005. "Geo-Spatial Technologies in Urban Environments." Heidelberg, Germany: Springer-Verlag.

Margai, Florence L., 2001. Health Risks and environmental inequity: a geographical analysis of accidental releases of hazardous materials. *The Professional Geographer* 53(3): 422-34.

Mennis, J. 2002. Using geographic information systems to create and analyze statistical surfaces of population and risk for environmental justice analysis. *Social Science Quarterly* 84: 281-297.

Mennis, J. and Lisa Jordan. 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), 2005, pp. 249-268.

Mohai, Paul, and Bunyan Bryant, eds., *Race and the Incidence of Environmental Hazards: A Time for Discourse*. Boulder, Colo.; Westview.

Pastor, M., J. Sadd, and J. Hipp. 2001. Which came first? Toxic facilities, minority move-in, and environmental justice. *Journal of Urban Affairs* 23: 1-21.

Pulido, L. 2000. Rethinking environmental racism: white privilege and urban development in southern California. *Annals of the Association of American Geographers* 90: 12-40.

Tucker, C.J. 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of the Environment* 8: 127-150.

United Church of Christ Commission for Racial Justice. 1987. *Toxic Wastes and Race in the United States: A National Report on the Racial and Socioeconomic Characteristics of Communities with Hazardous Waste Sites.* New York: Public Data Access.

www.epa.gov/tri/ U.S. EPA Toxics Release Inventory Program web page. 2006.

www.epa.gov/compliance/basics/ej.html U.S. EPA Office of Environmental Justice. 2006.