Landscape and Urban Planning 106 (2012) 262–270

Contents lists available at SciVerse ScienceDirect

Landscape and Urban Planning

journal homepage: www.elsevier.com/locate/landurbplan

The relationship between tree canopy and crime rates across an urban–rural gradient in the greater Baltimore region

Austin Troy[a,*], J. Morgan Grove[b], Jarlath O'Neil-Dunne[a]

[a] University of Vermont, Rubenstein School of Environment and Natural Resources and Spatial Analysis Lab, Aiken Center, 81 Carrigan Drive, Burlington, VT 05405, United States
[b] USDA Forest Service Northern Research Station, United States

ARTICLE INFO

Article history:
Received 6 October 2011
Received in revised form 9 March 2012
Accepted 27 March 2012
Available online 27 April 2012

Keywords:
Urban tree canopy
Crime
Urban vegetation
Public safety
Geographically weighted regression

ABSTRACT

The extent to which urban tree cover influences crime is in debate in the literature. This research took advantage of geocoded crime point data and high resolution tree canopy data to address this question in Baltimore City and County, MD, an area that includes a significant urban–rural gradient. Using ordinary least squares and spatially adjusted regression and controlling for numerous potential confounders, we found that there is a strong inverse relationship between tree canopy and our index of robbery, burglary, theft and shooting. The more conservative spatially adjusted model indicated that a 10% increase in tree canopy was associated with a roughly 12% decrease in crime. When we broke down tree cover by public and private ownership for the spatial model, we found that the inverse relationship continued in both contexts, but the magnitude was 40% greater for public than for private lands. We also used geographically weighted regression to identify spatial non-stationarity in this relationship, which we found for trees in general and trees on private land, but not for trees on public land. Geographic plots of pseudo-*t* statistics indicated that while there was a negative relationship between crime and trees in the vast majority of block groups of the study area, there were a few patches where the opposite relationship was true, particularly in a part of Baltimore City where there is an extensive interface between industrial and residential properties. It is possible that in this area a significant proportion of trees is growing in abandoned lands between these two land uses.

1. Introduction

A considerable literature has addressed the relationship between urban vegetation and crime but it is divided as to the direction of this relationship. Several studies have suggested that low, dense vegetation is positively associated with actual or perceived crime risk because it affords criminals a place to hide (Fisher & Nasar, 1992; Nasar, Fisher, & Grannis, 1993). Michael, Hull, and Zahm (2001) discuss how park police have indicated that dense vegetation is regularly used by criminals and how automobile thieves say they use dense vegetation to shield many of their activities, including target selection, examination of stolen goods, and disposal of unwanted goods. Stoks (1983) found that dense vegetation was a common characteristic of rape sites. In their guide to park design, Forsyth, Musacchio, and Fitzgerald (2005) discuss the importance of eliminating concealing undergrowth in parks to make users feel safer.

On the other hand, many other studies have found that vegetation is associated with decreased crime. Several potential reasons have been proposed for this effect. One relates to Jacobs' (1961) contention that places with more "eyes on the street" have more checks on dangerous behavior. Kuo (2003) suggests that well-designed green space might actually decrease crime by attracting people to spend time outdoors. The presence of more people in public places means that it is harder for criminals to go unnoticed. It also can result in an informal system of surveillance (Kuo & Sullivan, 2001), furthered by the fact that increased outdoor encounters foster social networks and relationships (Yancey, 1971). Having stronger social networks also means less likelihood of crime from within the community, for instance in the case of a public housing development (Sullivan & Kuo, 1996).

On the other hand, paved areas with no vegetation are often seen as "no-man's lands" which discourage residential interaction and reduce "eyes on the street," thereby making it easier for criminals to go unnoticed. This result is consistent with the findings of Kuo, Sullivan, Coley, and Brunson (1998), who found that residents disliked and avoided barren common spaces, typical of many unmaintained inner city parks, but that they liked photo-simulations of the same spaces showing the addition of grass and trees. It is also consistent with the results of Coley, Kuo, and Sullivan

* Corresponding author. Tel.: +1 802 656 8336.
E-mail addresses: atroy@uvm.edu (A. Troy), mgrove@fs.fed.us (J. Morgan Grove), Jarlath.ONeil-Dunne@uvm.edu (J. O'Neil-Dunne).

0169-2046/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.landurbplan.2012.03.010

(1997), who found that the amount of time residents spend in common outdoor neighborhood spaces is associated with the presence of trees and that the closer trees are to residential buildings, the more people spent time outside near them.

Another related reason why well maintained vegetation might reduce crime is that it can be seen as a "territorial marker" or a "cue to care," signifying to criminals that the residents actively care about and are involved with their surroundings (Brown & Bentley, 1993), even if they see no residents on the street. The presumption is that when looking for a place to commit crime, a perpetrator would move on to a neighborhood where cues suggest a weaker social organization and lesser neighborhood involvement. This is consistent with the "broken window theory," which posits that neighborhoods displaying visual cues of neglect or poor maintenance experience higher crime because these cues suggest to criminals a lack of effective law enforcement, while maintained neighborhoods send the opposite cue (Wilson & Kelling, 1982).

Perhaps one reason for the discrepancy in the literature has to do with differences in the type of vegetation being analyzed. Those studies finding a positive correlation between vegetation and crime appear to focus more on low, dense vegetation. Most of the studies finding an inverse relationship between trees and crime are not explicit about the type of vegetation being studied (e.g. are trees tall or low and immature?), although it can be inferred that they probably mostly deal with more open trees and grass. One study that is explicit about this is Kuo and Sullivan (2001), which looked specifically at grass and widely spaced, high-canopy trees, finding that their presence decreased crime around Chicago public housing apartment buildings (although that predictor only explained about 8% of the variance). The authors point out that the vegetation being studied in this case was not the type that would afford concealment; therefore the vegetation's crime-fighting characteristics outweigh its crime-inducing effects. In another study that controlled for vegetation type, Donovan and Prestemon (2012) found that low trees that decreased views from first floor windows on private lots in Portland, OR were associated with increased crime occurrence, while taller trees on private lots were associated with decreased crime. Street trees were generally associated with decreased crime.

Most of the research conducted to date on this topic has focused on relatively restricted geographic areas (e.g. a housing development or neighborhood) using relatively qualitative methods. The main exception is Donovan and Prestemon (2012), which sampled over 2800 individual housing units across Portland, OR and used a combination of aerial and ground photo interpretation. The study presented here represents the first known attempt to continuously analyze the relationship between crime and vegetation using fine resolution data across such a large extent—in this case multiple counties. It is also the first study to include such a wide range of land use conditions, from dense inner city areas to rural agricultural zones. It is able to look at such a large area at such fine detail because of the availability of accurate, high resolution data on tree canopy, in addition to the availability of geocoded crime data.

2. Research questions

There are three goals for this research. The first is to assess whether there is a statistical association between indictors of certain crime types and measures of tree cover in Baltimore City and Baltimore County, MD, when controlling for potentially confounding factors. The second goal is to determine whether this relationship may differ depending on whether trees are located in public or private land. To gain further insight on the different types of public trees, we look at how results might differ when street trees were included or excluded from this analysis. The third goal is to understand how the relationship between trees and crime might vary across space in a way that cannot adequately be controlled for in a linear regression model. By looking at the spatial pattern in this relationship, we hope to identify key omitted variables that could potentially be quantified in subsequent research. Answering these questions will also help policy makers determine where to strategically target urban forestry investments.

3. Methods

3.1. Study area

The study area includes Baltimore City and Baltimore County in Maryland. This area contains a wide range of neighborhood crime rates, from among the highest in the nation, to near-zero. While the average total crime rate by block group for the 1200 block groups in these two counties is approximately 2.5 times the national average, the range is immense, from areas with nearly non-existent crime to areas with seven times the national average. For robbery alone, the range is even greater, from near-zero to seventeen times the national average. Baltimore City (which is also its own county, but distinct from Baltimore County) has the highest concentration of crime. The mean total crime rate there by block group is 3.5 times the national average, while the mean robbery rate is nearly eight times the national average. Nonetheless, statistics are improving. Violent crimes, which were at nearly 22,000 (3% of Baltimore's population) in 1993, were down to around ten thousand (1.6% of population) by 2009.

Land use and land cover are also extremely variable across the study region, ranging from dense urban environments near downtown Baltimore to high agricultural or forested rural areas in northern Baltimore County. Tree canopy percentage by block group ranges from 0% in the central city to 87% in some of the more distant exurbs. Population densities range from as high as 250 people per acre in the densest urban core to as low as 75 people per square mile in the rural fringe.

3.2. Data

Crime data for Baltimore City and County came from Spotcrime (http://spotcrime.com), a service that aggregates and address geocodes crime data from public record police reports, augmented by news stories and user input. Spotcrime is the most widely distributed Internet source of crime mapping in the United States. The crime database went from the middle of 2007 to the end of 2010 and gave the crime type, a brief description and the geographic coordinates. We developed a crime index that was used as a dependent variable. The crime index consists of the density of combined robbery, burglary, theft, and shooting crimes. Robbery includes the taking or attempted taking of goods from a person by force or threats, and it commonly includes holdups. Burglary includes the unlawful entry into a structure, such as a house or store, to commit theft or some other felony. This includes "breaking and entering" and "forcible entry." Theft includes various forms of stealing not included in the previous categories. It is often synonymous with "larceny." One of the largest categories it includes is theft in or of motor vehicles. Shooting includes murder and attempted murder. Together, these account for over 70% of all crimes in the region. We chose this combination of variables not only because it represents such a large percentage of all crimes, but also because these are all crimes that can potentially benefit from concealment and be deterred by "eyes on the street." For instance, robberies or car thefts are generally outdoors and the chance that they will be attempted in a given place very much depends on the absence of bystanders and opportunities for concealment. Burglaries also are far more likely when there are few "eyes on the street" and better places to hide. The only other frequent crime types in the database were assault and vandalism. Assault was not chosen because the descriptor column indicated that a large percentage of incidents under this category were domestic assaults—an event that very often happens indoors. This category includes rape, which is an important crime indicator, but separating out rapes from domestic assaults was not feasible with the dataset. While domestic assault may potentially be lessened by the presence of vegetation (Sullivan & Kuo, 1996), we chose not to include it as a category because the mechanisms by which vegetation may affect its prevalence are both different from those affecting the other crime types and less understood. Vandalism was not included because it was attributed in the database for the County but not for the City.

Tree canopy data came from a 2007 high resolution (1 m) land cover layer for Baltimore City and County. The tree canopy layer was derived from 2007 1-m resolution color infrared imagery sourced from the National Agricultural Imagery Program (NAIP) along with surface models generated from light detection and ranging (LiDAR) data. The use of LiDAR was particularly valuable in that it allowed for the detection of trees within areas obscured by building shadow and the differentiation of canopy trees versus low woody vegetation, like shrubbery. The imagery and LiDAR were integrated into an object-based image analysis (OBIA) system in Definiens eCognition software. A series of segmentation, classification, and morphology routines were used to extract the tree canopy based on the spectral and spatial information contained within the data (Zhou & Troy, 2008, 2009). The smallest patch of tree canopy that could be detected using this approach was 9 m^2. The entire dataset was then manually reviewed at a scale of 1:2500 and all identifiable errors were corrected. To assess the accuracy of the tree canopy data following manual corrections we performed a stratified random sample following Congalton and Green (2009). 150 points were sampled for areas classified as tree canopy, and another 150 sampled for areas not classified as tree canopy. The source imagery and LiDAR were used as the reference data. The user's and producer's accuracy for the tree canopy class were 95% and 92% respectively. The overall accuracy was 93%.

Tree canopy data and crime data were summarized by Census block groups, the smallest geography for which the needed socio-economic control variables from the Summary File (SF)-3 long form dataset were available. Block groups vary in size but have similar numbers of households and roughly correspond with small-scale neighborhoods. The lack of the long form in the 2010 Census (it has largely been replaced with the American Community Survey, which has much smaller samples and larger margins of error, making it problematic at fine geographic scales), meant that 2000 Census data were used even though the crime and land cover data analyzed are closer in time to 2010. There were 1208 block groups in the study area, ranging in size depending on population density.

Crime data were summarized by taking a simple point density of crime points meeting the crime type criteria. The search radius, or area over which density is calculated for a given pixel, was 500 m. The choice of a 500 m radius was based on a study by Peters and Efflers (2010), which found that barriers to "crime trips" do exist but that they are most significant when less than 500 m from the criminal's origin location. Treating trees as a "psychological barrier" to crime, the density interpolation for crime at a given pixel is assumed then to be a function of actual crimes as a far as 500 m away for the purposes of our study. A 1000 m search radius was also attempted and resulted in only slightly different results, not presented here. Densities were calculated in units of points per square kilometer. Densities were then summarized by block group using the ArcGIS Zonal Statistics function, which returns the average of a raster surface for a polygon. Census variables were joined to this table. In order to break down tree coverage by public and private ownership, a parcel map with ownership codes was combined (using the Union function) with block group boundaries. This served to subdivide block groups by ownership type. They were then re-aggregated to give percent canopy cover for both private and public lands by block group. This was done in two different versions: in the first version tree canopy percentage was given for all public and private lands; in the second, public rights of way (street polygons) were excluded from the public lands calculations. This was done so that we could assess how crime is associated with street trees versus other types of public trees, like those in parks. An obvious limitation of this approach is that non-street trees often overhang the public right of way and street trees often overhand non-right of way land. However, we believe this limitation has relatively little impact on our conclusions.

Table 1
Variable names and summary statistics (all variables are by US Census block group).

Variable name	Variable description	Mean	STDV
PTREE	Percent of land area covered by tree canopy from 2007	0.29	0.17
PTREE.PUB1	Percent of land area covered by tree canopy in public ownership including rights of way	0.06	0.09
PTREE.PUB2	Percent of land area covered by tree canopy in public ownership excluding rights of way	0.09	0.08
PTREE.PRIV	Percent of land area covered by tree canopy in private ownership	0.20	0.14
POP00.SQMI	Population density by square mile from 2000 Census	11,122.24	9290.55
MED.HH.INC	Median household income from 2000 Census	41,189.85	21,808.98
P.SFDH	Percent of housing that is single family, detached from 2000 Census	0.31	0.33
P.HH.RUR	Percent of population classified as "rural" from 2000 Census	0.031	0.16
MED.YR.ALL	Median year of construction from 2000 Census	1954.387	13.92
P.WH	Percent white population from 2000 Census	0.4935	0.3924
P.PROTLAND	Percent of land that is protected open space	0.0682	0.1324
P.AGPR	Percent of land that is in agricultural preserve designation	0.0037	0.0263
CRIME	Density of crime points from late 2007 to late 2010, including robbery, shooting, burglary and theft (per square kilometer)	476.94	479.36
POWNOCC	Percent of housing that is owner occupied	0.061	0.268

3.3. Statistical methods

Once data were processed, a number of statistical models were run to test the relationship between crime and tree canopy. Ordinary least squares regression was first used with crime index as a dependent variable, percent tree canopy cover by block group as an independent variable, and a number of control variables at the block group level. All variables are described in Table 1. Socio-economic status (income, race), housing type and tenure, and the natural environment were all expected to be significant confounders based on past studies that found that these factors are correlated with the presence of urban vegetation (Grove et al., 2006; Troy, Grove, O'Neil-Dunne, Pickett, & Cadenasso, 2007). Several dozen variables were found to be significant within these different categories. These include: median year of housing construction, percent of households classified as "rural," median household income, percent white, percent single family detached homes, population density, percent owner occupied housing, percent of land in agricultural trust designations, and percent of land protected as open space. A Box–Cox transformation was used to help determine the optimal transformation of the dependent variable. This approach, first described by Box and Cox (1964) and later elaborated on by others (Bender, Gronberg, & Hwang, 1980; Halverson & Pollakowski, 1979; Spitzer, 1982), should yield the transformation that allows the model to best meet many of the assumptions of linear regression, such as constant variance or errors. It is described by the following power transformation:

$$y^{(\ )} = \begin{cases} \dfrac{(y\ \ -1)}{\ } & \text{for} \quad \neq 0 \\ \ln y & \text{for} \quad = 0 \end{cases} \tag{1}$$

The parameter is estimated through maximum likelihood to find the optimal transformation of a variable. This can be done for both dependent and independent variables, but in this case we only do this for the dependent. Once that parameter is estimated, the dependent variable can be transformed according to (1), although there are three "special cases" where $=1$, $=0$, and $=-1$, corresponding to linear and natural log and reciprocal models respectively. Although many use the Box–Cox transformation as an actual variable transformation, it can also be considered as a diagnostic to determine if a model fits one of those three special cases. For our model, we found the optimal to be 0.15, which is extremely close to a log transformation. Given this closeness and the difficulty in interpreting the output of a non-special case Box–Cox-transformed dependent variable, we felt justified in using a log transformation. Doing so increased the *R*-squared from 0.60 to near 0.85. Furthermore, it dramatically reduced heteroskedacticity of the error term, as evidence by changes in the residual plot, and resulted in a more normally distributed error term, as evidenced by a Shapiro–Wilke test.

Three models were run: one with aggregate percent tree cover, PTREE (model 1); one with tree cover distinguished by public and private ownership, PTREE.PUB1 and PTREE.PRIV (model 2); and one with tree cover separated by public and private but with public rights of way excluded, PTREE.PUB2 and PTREE.PRIV (model 3).

We also ran a spatially adjusted regression in order to determine if our results were robust to potential spatial autocorrelation (Cliff & Ord, 1981). When autocorrelation exists, either in the dependent variable or the error term, it can lead to a number of statistical problems. One of those is pseudo-replication: since observations are not truly independent, they are effectively overcounted when calculating degrees of freedom, which in turn leads to type 1 errors (Fortin & Dale, 2005). When the error term is autocorrelated, it violates the regression assumption of independent error terms. When the dependent variable, *y*, is inherently autocorrelated, this violates the assumptions of independent observations and can lead to biased regression estimates because *y* in one location is a function of *y* in neighboring locations, yet the traditional regression model does not account for this.

Spatially adjusted regressions serve to address these statistical assumption violations. The two most common types of spatial regression are the spatial error and spatial lag models. The former assumes that the error term is subject to spatial autocorrelation (largely expected to be a result of omitted variables), while the latter assumes generally (there are several variants) that the response variable (and potentially the error term too) is subject to spatial autocorrelation. The fact that a Moran's I test revealed that both the residual term from model 1 and the dependent variable were high autocorrelated meant that either spatial lag or spatial error models could have been appropriate. Both types were compared using a number of heuristics including pseudo *R*-squared, Akaike's Information Criterion, Log-likelihood, Lagrange Multiplier tests, and the Breusch–Pagan Test for heteroskedasticity. All clearly pointed to the superiority of the spatial lag model. Most importantly, though, theory supported the use of this model, for it has been well established in the literature that crime begets crime—that is, crime in one neighborhood can easily spill over into adjacent neighborhoods.

This left the question of how to parameterize the neighbor weight matrix, which in turn defines what is considered a neighbor for the purposes of adjusting for spatial autocorrelation. We chose to use a fixed distance band for defining neighbors rather than a fixed number of neighbors because the size of polygons (and hence the spacing between centroids) is extremely variable throughout the study area. Using a fixed number would mean that in the rural northern part of the study area, where block groups are many times the size of their urban counterparts, objects would be considered "neighbors" at extremely long distances. This is inappropriate because it would result in too many rural block groups being considered neighbors and too few dense urban block groups (where the vast majority of the sample is) being counted as such. Also because such block groups are large and internally heterogeneous, spatial autocorrelation is less critical to capture in these areas. We instead used a fixed distance band, which meant that all neighbors within the specified distance of a given observation would be considered as neighbors. Using this method ensures that an adequate number of small urban block groups are counted in neighborhood calculations. To determine the appropriate distance band we first ran semivariograms (a graph showing how variance between point pairs varies as a function of distance of the points) of the dependent variable. This indicated a terrace-type pattern with an initial sill (the point at which autocorrelation of point pairs levels off) at a range of approximately 2 km. We compared some fixed distance bands for the neighbor matrix both slightly above and below this value and found that 2 km resulted in the best *R*-squared and hence decided to use this value in building the neighbor matrix. We also found that the best fitting models used an inverse distance weighting function, rather than a simple binary measure of neighbor connections.

Spatial lag regressions were then run for aggregate tree cover (model 4), tree cover broken down by public and private ownership (model 5), and tree cover separated by public and private but with public rights of way excluded (model 6).

Finally, a geographically weighted regression (Brunsdon, Fotheringham, & Charlton, 1998; Fotheringham, Brunsdon, & Charlton, 2002; Fotheringham, Charlton, & Brunsdon, 1998) was used to determine whether there was significant spatial non-stationarity in the relationship between tree cover and crime. Spatial non-stationarity means that the relationship between a dependent and independent variable is non-constant over space. GWR, as it is known, is a form of moving window regression in which parameter estimates are deterministic functions of spatial location, as described by the following:

$$Y(x) = \alpha(u_i, v_i) + \sum_k \beta_k(u_i, v_i) x_k + e_i \tag{2}$$

where *u* and *v* describe spatial coordinates. In it, a separate regression is run for each observation, centered on that observation and, in our case, using an adaptive spatial kernel that subsets and weights nearby observation based on distance from that central point. The size of the kernel adapts in response to the density of observations. GWR is commonly used as a diagnostic tool to help identify potential omitted effects from models describing

Table 2
Coefficients and significance levels for non-spatial regression models.

	Model 1	Model 2	Model 3
(Intercept)	47.4040***	48.7649***	46.6514***
PTREE	−2.064130***		
PTREE.PUB1		−1.75262***	
PTREE.PUB2			−2.410941***
PTREE.PRIV		−1.8143***	−1.745036***
MED.YR.ALL	−0.020784***	−0.02151***	−0.020397***
P.HH.RUR	−2.553094***	−2.58181***	−2.593853***
P.Protland	−0.619685***	−0.76996***	−0.364267**
MED.HH.INC	−0.000004***	−0.000004***	−0.000005***
P.Agpr	−5.201022***	−4.92173***	−5.382584***
P.WH	−0.163993***	−0.15511***	−0.183161***
P.SFDH	−0.722312***	−0.77379***	−0.712791***
POP00.SQMI	0.000020***	0.000021***	0.000020***
POWNOCC	−0.432310***	−0.40784***	−0.409406***
R-squared	0.842	0.836	0.841

*Significant at 90% level.
**Significant at 95% level.
*** Significant at 99% level.

Table 3
Coefficients and significance levels for spatially adjusted models.

	Model 4	Model 5	Model 6
(Intercept)	7.2459***	7.3262***	6.6505***
PTREE	−1.1828***		
PTREE.PUB1		−1.2097***	
PTREE.PUB2			−1.5676***
PTREE.PRIV		−0.88917***	−0.9092***
MED.YR.ALL	−0.00149	−0.00158	−0.0012
P.HH.RUR	−1.74154***	−1.73641***	−1.7518***
P.Protland	−0.72094***	−0.69965***	−0.4633***
MED.HH.INC	−0.000003***	−0.0000035***	−0.0000038***
P.Agpr	−5.08048***	−5.02147***	−5.2554***
P.WH	0.30029***	0.313898***	−0.291***
P.SFDH	−0.35722***	−0.39392***	−0.3609***
POP00.SQMI	0.000006	0.000007***	0.000006***
POWNOCC	−0.27705***	−0.2483***	−0.2565***
Pseudo *R*-squared	0.902	0.90	0.902
Rho	0.084***	0.086***	0.085***
Log likelihood	−701.177	−713.432	−702.142

*Significant at 90% level.
**Significant at 95% level.
*** Significant at 99% level.

complex phenomena. It is particularly useful for situations in which there are large number of potential influences and interactions that cannot be reasonably or parsimoniously controlled for. Given the vast number of potential drivers and interactive effects influencing crime, we deemed it worthwhile to run GWR to see if the resulting visual patterns might provide any additional insight about the mechanism behind the tree–crime relationship. One of the key tests of GWR is the Monte Carlo significance test for spatial variability of parameters. This indicates which of the parameters, if any, exhibit spatial nonstationarity. It does so by testing whether parameter values for a particular variable differ significantly from those in a theoretical distribution generated with a Monte Carlo simulation technique (Hope, 1968). Once nonstationarity is established, parameter estimates or test statistics can be mapped.

The first GWR model (model 7) was analogous to model 1, the second (model 8) to model 2, and the third (model 8) to model 3 in terms of the variables used.

4. Results

All three ordinary least squares (OLS) regression models had *R*-squared values around 0.84 and all variables were significant at the 99% confidence level with the expected sign (Table 2). Year of construction, percent rural, percent protected land, median income, percent agricultural preserve, percent white, percent single family detached home and percent owner occupied housing were all negative. All of these factors were expected to be associated with reduced crime. Population density was associated with higher crime, also as expected. In model 1, the tree variable was negative and significant at the 99% level, with a coefficient of −2.06. Given the log transformation of the dependent variable, this can be interpreted as a roughly 20% decrease in crime for a 10% increase in tree canopy at the block group level. Model 2 shows that when this term was broken down into trees on public and private land, both variables are negative and significant at the 99% level, with similar magnitudes, both slightly lower than the tree variable in model 1. However, model 3 shows that when public lands do not include rights of way (PTREE.PUB2), the magnitude of the public lands coefficient grows to be roughly 40% higher than the private lands coefficient.

Under models 4, 5, and 6 (Table 3), the tree variables were still significant with the same sign, but with somewhat lower magnitudes. The coefficient on the total tree cover variable in model 4 is much lower than in model 1 (at −1.18 vs. −2.06). In model 5, the magnitudes of coefficients are lesser than in the corresponding OLS model (model 2) for both public trees (−1.209 vs. −1.753) and private trees (−0.889 vs. −1.814). In model 6, the magnitudes of coefficients are lesser than in model 3 for both public trees outside of rights of way (−1.568 vs. −2.41) and private trees (−0.909 vs. −1.74). In other words, the more conservative spatial regression method mostly yields smaller marginal effects. The significant values for the spatial lag coefficient, rho, for models 4–6 indicate there is strong spatial dependence in the sample data and that it was warranted to use this type of modeling approach. As for control variables, all remain significant at the 99% confidence level except the housing age variable, which dropped out at even at the 90% confidence level. Pseudo *R*-squared values are all near 0.9 for the three models.

The most important output from GWR is the Monte Carlo significance test for non-stationarity. It essentially shows which independent variables have a spatially non-constant relationship with the dependent variable. Then, based on that, the variation in parameter values can be mapped out geographically to look for patterns. The GWR model with a single variable for tree cover (model 7) indicated that five parameters were spatially nonstationary at the 99% confidence level, including percent trees. Other nonstationary variables included population density, percent single family detached homes, median year of construction and percent white. In other words, the relationship between all these variables and crime was found to vary across space. For model 8, the private tree variable was significantly non-stationary at the 99% confidence level, but not the public tree variable. The control variables that were significantly non-stationary the same as in model 7. In model 9, the private trees variable was significantly non-stationary at the 99% level, and the significant control variables were the same as in models 7 and 8 except that the variable for single family detached homes was only non-stationary at the 95% level.

Spatially varying significance test measures on the tree canopy variable, known as “pseudo *t*-statistics,” were then plotted out geographically. This was done in favor of plotting out actual parameter values because doing so fails to account for which of those values are statistically significant or not (that is, significant in terms of slope, not non-stationarity). Using −1.65 and positive 1.65 as the upper and lower thresholds for significance, pseudo *t*-statistics were then plotted out to show block groups where the relationship between crime and trees was significant and positive, significant and negative, and non-significant at the 95% confidence level. The resulting map for model 7 is given in Fig. 1. Pseudo *t*-statistics on the private tree variable were then plotted out from model 8. The

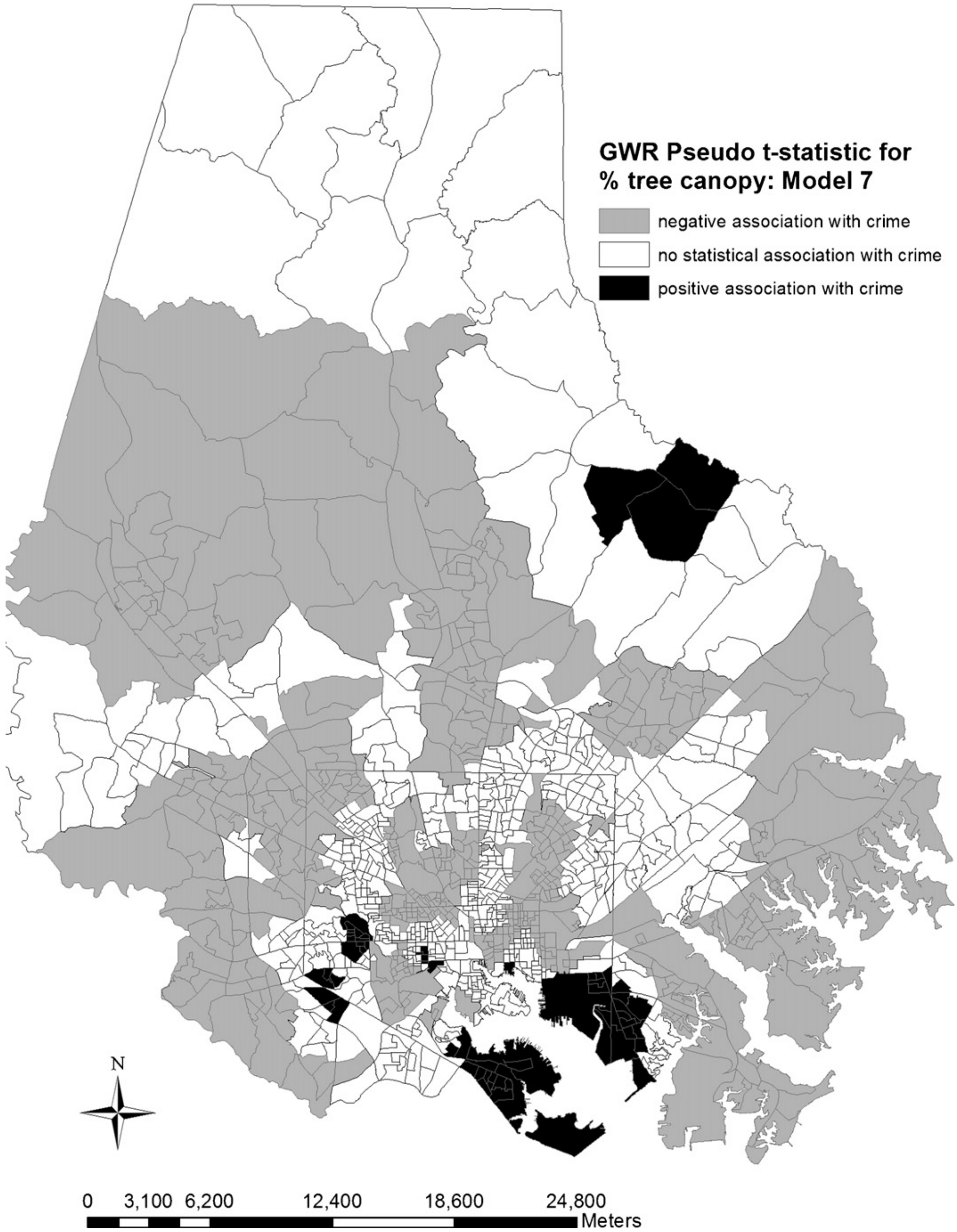

Fig. 1. Plot of pseudo *t*-statistic for percent tree canopy variable from GWR for model 7. The map shows Baltimore City and Baltimore County.

resulting map is given in Fig. 2. Because private trees are characterized the same way in both models 8 and 9 (only public trees are different and the variable on public trees was stationary), no map was made for model 9.

5. Discussion

The results of this study indicate that crime has a strong negative association with tree cover, even after controlling for socio-economic variables such as income, housing age, ruralness, race, housing type, housing tenure, population density, and amount of protected or agricultural land. However, the exact magnitude of this impact varies depending on the model. The more conservative—and probably more accurate—spatially adjusted model indicates a coefficient of lower magnitude on tree cover of −1.18. This suggests that a 10% increase in tree cover would be associated with an 11.8% decrease in crime rate, all else equal. It seems unlikely that the entire magnitude of this effect is purely causal. Rather, it is probably at least partially accounted for by omitted variables. Still, the fact that *R*-squared values are as high as they are suggests that there is some genuine relationship between trees and crime. The results of the spatial regression indicate that model results are not an artifact of spatial autocorrelation.

The relationship between trees and crime appears to vary somewhat between public and private land. When tree cover is broken up into these categories, the magnitude of the effect goes down slightly for each, and those effects are of roughly the same magnitude (model 2). However, when the more conservative spatial model is used (model 5), a big difference appears between the public and private tree effects, with the former being nearly 40% larger.

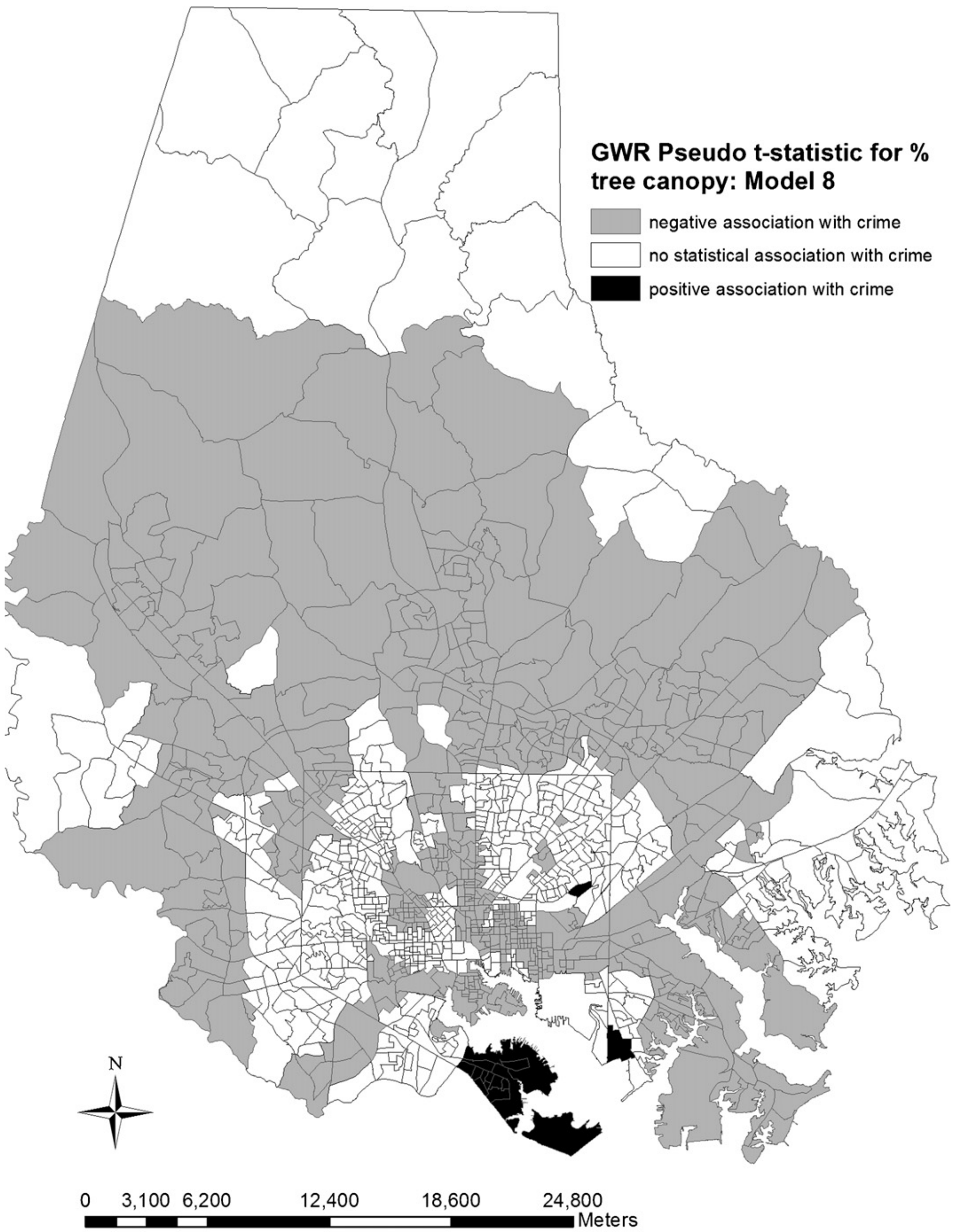

Fig. 2. Plot of pseudo *t*-statistic for percent tree canopy on private land from GWR for model 8. The map shows Baltimore City and Baltimore County.

This would suggest that planting trees on public lands might yield somewhat higher crime-reduction benefits than planting on private. This would suggest that the private land trees result in model 2 is confounded by spatial autocorrelation. When public rights of way are not included in the analysis of public trees a differential between public and private trees becomes evident in the non-spatial model (model 3), while in the spatial model (model 6) the gap between public and private increases relative to model 5, as does the magnitude of both coefficients. In this case, the magnitude on public trees is nearly 50% greater than that of private trees. This would suggest that trees in non-right of way public lands are the most effective components in terms of reducing crime. These types of public lands might include parks, other protected open space, and major government buildings or facilities.

Finally, GWR analysis indicates that total tree percentage and private tree percentage are both spatially nonstationary while public tree canopy percentage is not. This suggests that our model adequately explains the global relationship between public trees and crime, while it omits some unknown interaction effect that conditions the relationship between private trees and crime. The pattern displayed in the GWR outputs in Figs. 1 and 2 indicate that there is a patchy relationship between private trees and crime. In the vast majority of block groups, the relationship is either significant and negative or non-significant. The big exceptions are in Baltimore City's Brooklyn Park, Wagners Point, and Dundalk neighborhoods located ringing the outer harbor in the far south of the map, where crime appears to be positively associated with tree cover—both in aggregate and specifically for private land. What makes these places different? A possible explanation is that there is a considerable amount of lower, early successional, and apparently unmanaged stands of trees both on small residential lots and on larger private institutional/industrial parcels in some of these neighborhoods. With major harbor facilities nearby, the neighborhoods in question have a significant interface zone between industrial and residential land that could be considered a "no man's land." It is quite possible that the small patches of unmanaged trees that are often found adjacent to warehouses, truck yards, factories, etc., provide an excellent hiding place for criminals looking to prey upon residents going to and from nearby homes. A preliminary photographic analysis of the area shows that there are a considerable number of such low, overgrown patches. But whether they are found in greater preponderance here than in other neighborhoods is uncertain, for there are certainly other such patches of unmanaged vegetation elsewhere, particularly on Baltimore's large number of vacant lots.

The nonstationarity of the parameter on trees is consistent with the literature's finding of opposing effects of trees on crime. It supports the contention that trees can be both an asset for criminals by providing concealment and a deterrent to them by increasing "eyes on the street" (Jacobs, 1961; Kuo, 2003) or by giving a territorial marker cue that residents actively care about their neighborhood (Brown & Bentley, 1993; Kuo et al., 1998; Wilson & Kelling, 1982). We did not control for the potential concealment value of trees (i.e. whether they are short and scrubby or canopy trees) or for their level of management, which would be expected to correlate with the territorial marker value of trees. Given the thousands of abandoned lots and the large number of industrial–residential interface zones in Baltimore, it would appear that unmanaged trees in these areas might explain why crime and trees are positively correlated in some areas. Testing this would be an auspicious direction for future research. One possible approach to this would be to use LiDAR data to account for the height of trees and to include location in and around vacant lots or industrial interface zones as covariates. Street-level photographic analysis could also be used.

Another potential explanation for the non-stationary effects is that neighborhood crime levels actually influence the perception of trees in certain contexts. It may be that heavy criminal activity—combined with other location-specific socio-economic or design characteristics—causes residents to perceive vegetation as more threatening. This is consistent with the findings of Troy and Grove (2008), who found that an increase in crime rate in and around parks in Baltimore caused the effect of parks on property values to go from being positive to being negative. Where this effect is present, it could cause people to further avoid vegetated areas, in turn reducing the positive "eyes on the street" effect and eventually creating a downward spiral in which vegetated areas become increasingly unmanaged as they become less desirable. In this sense, non-stationarity could be caused by legacy effects related to "tipping points" or thresholds crossed in the past. This would be another fruitful area for future research.

Our finding that public land trees have a bigger negative association with crime suggests that fear of public trees as crime facilitators may be misguided. More importantly, the fact that private land is the culprit in areas where crime is positively associated with trees suggests that private land trees are more likely to act as facilitators to crime than public trees. This would then suggest that a strategy to use trees to combat crime should have two prongs: first, increase canopy in areas with few trees in order to maximize the "eyes on the street" and "cues to care" effects; and second, encourage better management of potential crime-facilitating vegetation on private lands so that it cannot serve as a screen for criminals, particularly on abandoned lots or in interface zones around industrial properties. This could be accomplished not only by appropriately spacing trees and choosing the right species, but also by frequent pruning.

These findings fit within a larger context and growing trend in urban forest management and urban sustainability. Specifically, there is increasing recognition that urban trees are great "multitaskers," providing diverse benefits such as esthetics (Acharya & Bennett, 2001; Morancho, 2003; Tajima, 2003) moderation of temperature (Akbari, 2002; Akbari, Pomerantz, & Taha, 2001; Shashua-Bar & Hoffman, 2000), and stormwater processing. These benefits connect to programmatic and regulatory concerns that extend beyond traditional urban forest management agencies such as Departments of Recreation and Parks or Public Works. For instance, a esthetics connects to departments of Real Estate and Economic Development, moderation of temperature to Public Health, and stormwater regulation to Public Works.

Our findings recruit an additional agency partner with an interest in promoting urban forestry and management: Police Departments. Our findings extend existing theories of community policing in that we find an environmental component to the "broken window" theory, one that includes empty tree pits and poorly maintained trees. This suggests that there could be potential benefits from incorporating public safety criteria into the prioritization of areas for tree planting (Locke et al., 2010) and from involving Police Departments in the process.

6. Conclusion

Our findings add to the literature on the relationship between crime and vegetation in a number of ways. First, we find a strong inverse association between crime rates and tree canopy cover in the Baltimore region, adjusting for many confounding factors. Second, this result holds for both public and private land, but it is stronger for public land. Third, when spatial autocorrelation is adjusted for, the overall result still holds, but the magnitude is not as great. Finally, it appears there is some slight geographic variability in the relationships between crime and trees and that a few isolated areas see a positive relationship between crime and trees. Results in these anomalous areas may relate to the fact that they contain relatively large interface zones between residential and industrial uses where vegetation tends to be more unmanaged, such that the concealment value of the vegetation outweighs its deterrent effect. These results do not establish causality, but suggest a strong need for further research to determine the role of vegetation in mediating crime.

Acknowledgements

This research was made possible by the generous support of the U.S. Forest Service's Northern Research Station and Northeastern Area State & Private Forestry Program (USDA03-CA-11244225-531), the National Science Foundation Long-Term Ecological Research program (NSF DEB-0423476) and the National Science Foundation Human and Social Dynamics Program (award #0624159), which in turn supported the Baltimore Ecosystem Study. Thanks also go out to R. Holli Howard for her expert GIS help and to Colin Drane of Spotcrime for his help with interpreting crime data.

References

Acharya, G., & Bennett, L. L. (2001). Valuing open space and land-use patterns in urban watersheds. *Journal of Real Estate Finance and Economics*, *22*(2–3), 221–237.

Akbari, H. (2002). Shade trees reduce building energy use and CO_2 emissions from power plants. *Environmental Pollution*, *116*, S119–S126.

Akbari, H., Pomerantz, M., & Taha, H. (2001). Cool surfaces and shade trees to reduce energy use and improve air quality in urban areas. *Solar Energy*, *70*(3), 295–310.

Bender, B., Gronberg, T., & Hwang, H. S. (1980). Choice of functional form and the demand for air quality. *Review of Economics and Statistics*, *62*, 638–643.

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society*, *26*(2), 211–252.

Brown, B., & Bentley, D. (1993). Residential burglars judge risk – The role of territoriality. *Journal of Environmental Psychology*, *13*(1), 51–61.

Brunsdon, C., Fotheringham, S., & Charlton, M. (1998). Geographically weighted regression – Modelling spatial non-stationarity. *Journal of the Royal Statistical Society Series D-The Statistician*, *47*, 431–443.

Cliff, A. D., & Ord, J. K. (1981). *Spatial processes: Models & applications*. London: Pion.

Coley, R., Kuo, F., & Sullivan, W. (1997). Where does community grow? The social context created by nature in urban public housing. *Environment and Behavior*, *29*(4), 468–494.

Congalton, R., & Green, K. (2009). *Assessing the accuracy of remotely sensed data: Principles and practices* (2nd ed.). Boca Raton: CRC Press/Taylor & Francis.

Donovan, G. H., & Prestemon, J. P. (2012). The effect of trees on crime in Portland, OR. *Environment and Behavior*, *44*(1), 3–30.

Fisher, B. S., & Nasar, J. L. (1992). Fear of crime in relation to 3 exterior site features prospect, refuge, and escape. *Environment and Behavior*, *24*(1), 35–65.

Forsyth, A., Musacchio, L., & Fitzgerald, F. (2005). *Designing small parks: A manual addressing social and ecological concerns*. Hoboken, NJ: J. Wiley.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2002). *Geographically weighted regression: The analysis of spatially varying relationships*. Chichester, England: Wiley.

Fotheringham, A. S., Charlton, M. E., & Brunsdon, C. (1998). Geographically weighted regression: A natural evolution of the expansion method for spatial data analysis. *Environment and Planning A*, *30*(11), 1905–1927.

Fortin, M., & Dale, M. (2005). *Spatial analysis: A guide for ecologists*. Cambridge. England: Cambridge University Press.

Grove, J. M., Troy, A. R., O'Neil-Dunne, J. P. M., Burch, W. R., Cadenasso, M. L., & Pickett, S. T. A. (2006). Characterization of households and its implications for the vegetation of urban ecosystems. *Ecosystems*, *9*(4), 578–597.

Halverson, R., & Pollakowski, H. (1979). Choice of functional form for hedonic price equations. *Journal of Urban Economics*, *10*, 37–49.

Hope, A. C. A. (1968). A simplified Monte Carlo significance test procedure. *Journal of the Royal Statistical Society Series B (Methodological)*, *30*(3), 582–598.

Jacobs, J. (1961). *The death and life of great American cities*. New York: Random House.

Kuo, F., & Sullivan, W. (2001). Environment and crime in the inner city – Does vegetation reduce crime? *Environment and Behavior*, *33*(3), 343–367.

Kuo, F., Sullivan, W., Coley, R., & Brunson, L. (1998). Fertile ground for community: Inner-city neighborhood common spaces. *American Journal of Community Psychology*, *26*(6), 823–851.

Kuo, F. E. (2003). The role of arboriculture in a healthy social ecology. *Journal of Arboriculture*, *29*(Part 3), 148L 155.

Locke, D., Grove, J. M., Lu, J. W., Troy, A., O'Neil-Dunne, J., & Beck, B. (2010). Prioritizing preferable locations for increasing urban tree canopy. *Cities and the Environment*, *3*(1). Article 4

Michael, S., Hull, R., & Zahm, D. (2001). Environmental factors influencing auto burglary – A case study. *Environment and Behavior*, *33*(3), 368–388.

Morancho, A. B. (2003). A hedonic valuation of urban green areas. *Landscape and Urban Planning*, *66*(1), 35–41.

Nasar, J., Fisher, B., & Grannis, M. (1993). Proximate physical cues to fear of crime. *Landscape and Urban Planning*, *26*(1–4), 161–178.

Peters, M. P., & Efflers, H. (2010). Do physical barriers affect urban crime trips? *Crime Patterns and Analysis*, *3*, 38–49.

Shashua-Bar, L., & Hoffman, M. E. (2000). Vegetation as a climatic component in the design of an urban street – An empirical model for predicting the cooling effect of urban green areas with trees. *Energy and Buildings*, *31*(3), 221–235.

Spitzer, J. J. (1982). A primer on Box–Cox estimation. *Review of Economics and Statistics*, *64*(2), 307–313.

Stoks, F. G. (1983). Assessing urban public space environments for danger of violent crime – Especially rape. In D. Joiner, G. Brimikombe, J. Daish, J. Gray, & D. Kernohan (Eds.), *Conference on people and physical environment research* (pp. 331–343). Wellington, NZ: Ministry of Works and Development.

Sullivan, W., & Kuo, F. E. (1996). In U. a. C. F. USDA Forest Service Southern Region (Ed.), *Do trees strengthen urban communities, reduce domestic violence?* USDA Forest Service Southern Region, Urban and Community Forestry.

Tajima, K. (2003). New estimates of the demand for urban green space: Implications for valuing the environmental benefits of Boston's big dig project. *Journal of Urban Affairs*, *25*(5), 641–655.

Troy, A., & Grove, J. M. (2008). Property values, parks, and crime: A hedonic analysis in Baltimore, MD. *Landscape and Urban Planning*, *87*(3), 233–245.

Troy, A. R., Grove, J. M., O'Neil-Dunne, J. P. M., Pickett, S. T. A., & Cadenasso, M. L. (2007). Predicting opportunities for greening and patterns of vegetation on private urban lands. *Environmental Management*, *40*(3), 394–412.

Wilson, J. Q., & Kelling, G. L. (1982). The police and neighborhood safety: Broken windows. *Atlantic Monthly*, *249*, 29–38.

Yancey, W. L. (1971). Architecture, interaction, and social contreol: The case of a large-scale public housing project. *Environment and Behavior*, *3*, 3–21.

Zhou, W., & Troy, A. (2008). An object-oriented approach for analysing and characterizing urban landscape at the parcel level. *International Journal of Remote Sensing*, *29*(11), 3119–3135.

Zhou, W., & Troy, A. (2009). Development of an object-based framework for classifying and inventorying human-dominated forest ecosystems. *International Journal of Remote Sensing*, *30*(23), 6343–6360.