Modeling Predictors of Violent Crime in North Minneapolis, Minnesota USA using GIS-Based Regression Analysis

Edward Anderson
Department of Resource Analysis, Saint Mary's University of Minnesota, Minneapolis, MN 55404

Keywords: Crime, Spatial Analysis, Regression Analysis, Geographic Information Systems (GIS), Geographically Weighted Regression (GWR), Ordinary Least Squares (OLS), Minneapolis

Abstract

Despite the recent nationwide decline in violent crime rates, crime continues to be a significant problem in American cities. In recent years, Geographic Information Systems (GIS) has become an increasingly popular and effective tool for analyzing crime. The use of GIS-based regression analysis, specifically Ordinary Least Squares (OLS) and Geographically Weighted Regression (GWR) can help researchers identify underlying factors that correlate with crime. Using crime data from the City of Minneapolis, Minnesota USA as the dependent variable and independent variables consisting of demographic and socioeconomic data from the American Community Survey (ACS) and land use data from the Metropolitan Council, OLS and GWR models were constructed that explained 68% and 77%, respectively, of violent crime in North Minneapolis. Both models employed five significant variables: population density, percentage of population living in poverty, heterogeneity index, percentage of vacant housing units, and the percentage of land use classified commercial. This analysis gives criminologists, policymakers, and the general public a greater understanding of the underlying factors of violent crime in North Minneapolis.

Introduction

It is common knowledge violent crime is a problem in American cities. As a result, there is continued need to study crime in the hopes of effectively combating it. With the recent availability of ever-cheaper and faster computer processing power and the prevalence of GIS software such as Esri ArcMap, crime analysis with spatial statistics has become a viable and accessible option (Townsley, 2009). Regression analysis, for example, is often used to identify underlying variables that describe and explain instances of crime. The goal of identifying these variables is so that criminologists may understand the underlying factors that correlate with instances of crime. By subsequently informing government policymakers of these explanatory variables, measures can be taken to address them, leading, it is hoped, to fewer instances of violent crime (Chainey and Ratcliffe, 2005).

The widespread implementation and popularity of regression analysis has led to the research and development of more statistically-powerful regression models, such as Ordinary Least Squares (OLS) and Geographically Weighted Regression (GWR). It has also led to more interest and debate in crime analysis

Anderson, Edward. 2017. Modeling Predictors of Violent Crime in North Minneapolis, Minnesota USA using GIS-Based Regression Analysis. Volume 20, Papers in Resource Analysis. 12 pp. Saint Mary's University of Minnesota University Central Services Press. Winona, MN.

methodology. For example, a meta-analysis of crime studies by He, Paez, Liu, and Jiang (2015) indicate past analyses using regression analysis with data at geographic scales larger than the census block group may have been statistically flawed. A review of current literature indicates there are limited published violent crime analyses using OLS and GWR at the census block group level for North Minneapolis. In addition, although crime regression analyses have traditionally used decennial census data, few published crime analyses have utilized American Community Survey (ACS) 5-year estimates. The ACS 5-year estimates are the sole source of population, housing, economic, and social data published by the U.S. Census Bureau at the block group level (U.S. Census Bureau, 2009).

Methods

Study Area Considerations

In recent decades, North Minneapolis has been plagued by pernicious poverty and crime (Metropolitan Council, 2014). Historically an area of strong self-identification, North Minneapolis has well-defined borders: the Mississippi River to the east, the downtown and Interstate 394 to the south, Theodore Wirth Park and Victory Memorial Drive along the west, and 53rd Ave to the north. For these reasons it was chosen as the study area for this research.

GIS-based regression analysis requires the dependent and independent variables be divided into a spatial unit, such as census block groups, tracts or neighborhoods (Chainey and Ratcliffe, 2005). He *et al.* (2015) determined analysis units larger than the census block group should be avoided when conducting violent crime analysis as there is potential for geographic variation of independent variables to be masked. Cahill and Mulligan (2007) also advocated using the census block group for maximum statistical power. As a result, a shapefile of census block groups was downloaded from Hennepin County's GIS open data site and the 68 block groups that make up North Minneapolis were exported.

Dependent Variable

According to He *et al.* (2015), an important consideration is the necessity to use crime data concurrent with the demographic and economic data as it would be illogical to predict crime with independent variables from a different time period. Since the ACS 5-year estimate of 2010-2014 was to be used, datasets of geocoded crime incidents from 2010-2014 were obtained from the City of Minneapolis' open data portal. Violent crimes were selected from these datasets by use of the FBI's Uniform Crime Reporting (UCR) code, which classifies a crime as violent if it consists of murder or non-negligent manslaughter, forcible rape, robbery, or aggravated assault (He *et al.*, 2015). These crimes were then aggregated to the block group shapefile using a spatial join and averaged for a five-year period.

A common practice in crime analysis is to calculate a crime rate by dividing the number of violent crimes by the total population of the spatial unit of analysis (He *et al.*, 2015). However, at large geographic scales like the census block group, this is theoretically irrational as crimes do not necessarily take place where perpetrators or victims live. Zhang and Peterson (2007) state for crime analyses at the block group level, it is therefore best to calculate a crime density rather than a crime rate. For that reason, the five-year violent crime averages were divided by the total area of land per block group (Figure 1). The resulting densities were log transformed, a common data analysis practice to make the distribution of the data more normal. These violent crime density logs became the dependent variable for the Exploratory, OLS, and GWR regression models.

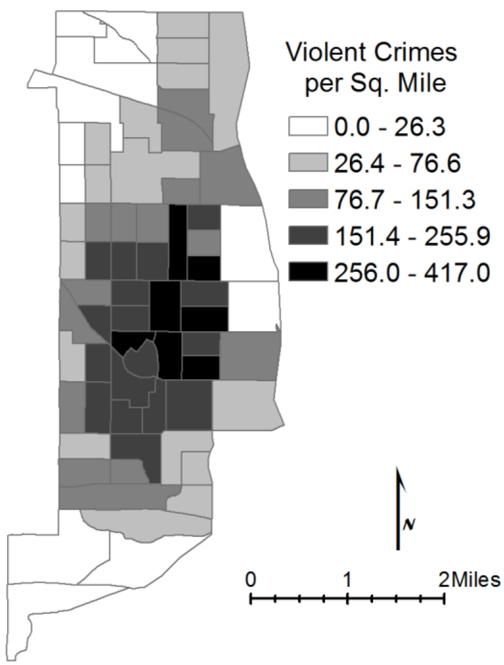

Figure 1. 2010-2014 average violent crime densities per census block group in North Minneapolis, MN.

Independent Variable Considerations

The theory of social disorganization has been highly influential in the field of criminology since its development at the Chicago School in the 1940s (Thompson and Gartner, 2014). It posits low economic status, residential mobility, and ethnic heterogeneity lead to social disorganization within a community, and this disorganization leads to greater levels of crime (He *et al.*, 2015). Past research has found strong correlations between crime and variables associated with these factors, as well as with variables representing family disruption, mixed-use land use, commercial land use, and population density (Cahill and Mulligan, 2007; Stucky and Ottensmann, 2009).

Low economic status was captured with three variables: percentage of total population living below poverty, percentage unemployed, and per capita income. It was hypothesized the percentage living below poverty and percentage unemployed would have positive effects on crime densities while per capita income would have a negative effect.

Residential mobility has a negative effect on the health of a community because it leads to weaker social bonds between residents (Sampson and Groves, 1989). Mobility was expressed with the following: the percentage of rental housing units, percentage of vacant housing units, and percentage of households in a different housing unit than they were a year ago. It was hypothesized all three would have positive associations with crime.

Racial and ethnic heterogeneity was captured by calculating a heterogeneity index for each block group (Figure 2). Past research has found a positive correlation between increasing racial and ethnic heterogeneity and crime. It is hypothesized that communities consisting of many self-identifying racial and ethnic groups have less social cohesion than homogenous populations and are, therefore, less able to prevent crime (He *et al.*, 2015). The heterogeneity equation takes into consideration both the number and size of different ethnic and racial groups within each block group (Sampson and Groves, 1989). This research used seven census-defined racial and ethnic groups: White, Black or African American, American Indian, Asian, Pacific Islander, Hispanic, and other. In theory, the index has a range of 0 to 1 with a completely homogenous population expressing a value of 0. For this research, the heterogeneity index ranged from 0.125 to 0.841.

$$1 - \Sigma p_i^2$$

Figure 2. Heterogeneity index equation. p_i is the proportion of the total block group population for each racial group. The proportions are squared, ^{2}, added together, Σ, and subtracted from 1.

Family disruption has also been shown to lead to increased levels of crime (He *et al.*, 2015). It was represented using the following: percentage of population unmarried, percentage of households living alone, and percentage of households consisting of nonrelatives. All three variables were expected to be positive in their relation to crime.

Past research has found positive correlations between both mixed-use and commercial land uses and crime, so these two variables were also included (Cahill and Mulligan, 2007). A shapefile of generalized land use from 2010, created by the Metropolitan Council, was obtained from the Minnesota Geospatial Commons. Land use class percentages were reclassified and calculated for each block group using Python scripting and ArcMap 10.3.

Population density was also included in this research, though its effect on crime in past research is mixed. Criminology theories such as routine activity theory hypothesize greater population densities mean more potential perpetrators and targets and thus more crime, but it can also equate to a greater number of capable guardians, which can have a negative effect on crime (Cahill and Mulligan, 2007). It was calculated by dividing the block group population by the area (in square miles).

At the time of this analysis, the ACS 5-year estimate for 2010-2014 were the most recent data available. Demographic data were downloaded in tabular format and percentages were calculated where necessary. The heterogeneity index was calculated using a Python script and ArcMap. The tabular data was joined to the block group shapefile via a common field, the block group spatial IDs.

Regression Analysis Considerations

OLS is considered global in scope as it computes one set of equation parameters for the entire study area (Figure 3). OLS is a form of linear regression, meaning correlations between the dependent and independent variables will be either positive or negative (Leung, Mei, and Zhang, 2000).

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 \ldots \beta_n X_n + E$$

Figure 3. OLS regression equation. Y signifies the dependent variable, β the coefficients, X the value of the independent variable, and E the residuals, i.e. the under and over predictions.

The benefit of GWR is its ability to account for shifting influence between the dependent and independent variables across a study area, often referred to as nonstationarity. In contrast to the global nature of OLS, which assumes a consistent relationship between the independent and dependent variables no matter their location, GWR is local, meaning it computes an equation with potentially different parameter values for each block group (Rosenshein, Scott, and Pratt, 2011). The block groups themselves are not considered wholly independent, however, as GWR uses a weighting matrix to assign influence from the values of neighboring block groups into its calculation (Leung *et al.,* 2000). From a theoretical standpoint, GWR supports Tobler's First Law of Geography, which posits nearby features have a greater impact than distant features. Thus, models which take adjacency into account, like GWR, are more rationally defensible, especially at small analysis units such as block groups (Townsley, 2009). In the past, studies using GWR have explained dependent variables with more statistical significance than global regression methods like OLS (Cahill and Mulligan, 2007).

Results

The thirteen independent variables were tested for significance using the Exploratory Regression tool, which used OLS to evaluate the candidate variables to determine which ones made viable OLS models. After 2,379 trials, Exploratory Regression identified a model of five variables which successfully passed the checks of model performance, significance, stationarity, redundancy, model bias and spatial autocorrelation. In order of probability significance at predicting crime, these variables were: population density, percentage of population living in poverty, heterogeneity index, percentage of vacant housing units, and the percentage of land use classified commercial. These five variables were then used to build OLS and GWR regression models.

OLS regression modeling found all five variables had probabilities that were statistically significant at the 95 percent confidence level, meaning they all fell under the 0.05 error threshold. The coefficients were also found to be positive, which was expected (Table 1). This means as the value of the variables increase, so does crime. The Adjusted R-Squared value for the OLS model was 0.678, meaning nearly 68% of violent crime density values were predicted using the five variables. Adjusted R-Squared, however, is only one indication of how well a model describes the dependent variable. There were other diagnostics that had to be successfully passed before the model was considered properly specified (Rosenshein *et al.,* 2011).

Table 1. P-values and coefficients for OLS variables. The smaller the p-value, the greater the significance of the variable at predicting crime.

Variable	**P-Value**	**Coefficient**
Population Density	0.0000	0.0002
% Impoverished	0.0002	1.9241
Heterogeneity Index	0.0002	2.2084
% Vacant	0.0014	3.4819
% Commercial	0.0490	4.8819

The Joint F-Statistic, an indication of model significance, was 29.19 with a p-value of 0.000000. Any value less than 0.05 indicates a model that is statistically significant. The Joint Wald Statistic, an alternate indication of model significance, also had a statistically significant p-value of 0.000000, well under the threshold of 0.05.

The degree of multicollinearity, or redundancy in the independent variables, is indicated by the variance inflation factor (VIF). Calculated for each explanatory variable, values larger than 7.5 indicate redundancy. All five variables had VIF values under 1.4 (Table 2).

Table 2. Variance Inflation Factor (VIF) values for independent variables. No variable exceed 7.5, which would be an indication of variable redundancy.

Variable	**VIF**
Heterogeneity Index	1.118
Population Density	1.166
% Impoverished	1.207
% Vacant	1.281
% Commercial	1.365

The Jarque-Bera Statistic was 0.4353 with a p-value of 0.8044. It was not significant, which was an indication residuals were normally distributed and not biased. Bias would indicate missing explanatory variables in the model.

The Koenker (BP) Statistic, employed to quantify the amount of variable nonstationarity, was 8.3342, with a p-value of 0.1388. It was not significant in the OLS model, meaning the variables explained crime densities consistently across the study area.

An additional diagnostic of model performance is spatial autocorrelation, or clustering of the OLS residuals. The presence of residual clustering would be an indication of a misspecified regression model. The Spatial Autocorrelation (Global Moran's I) tool returned a z-score of -0.124, which indicated the residuals were spatially distributed in a random pattern.

A map of the OLS residual standard deviations reveals most over and under predictions fall within -1.5 and 1.5 standard deviations (Figure 4). Blues are block groups where the model over performed (i.e. the actual crime density is lower than the model predicted) and reds are block groups where the model underperformed (i.e. the crime density is actually higher than predicted). No block group residual exceeded -2.18 standard deviations, but one under prediction in the northwest exceeded 2.74 standard deviations.

The five independent variables employed in the OLS model were next used to construct a GWR model. The OLS model's non-significant Koenker (BP) Statistic suggested GWR would not produce a model with improved performance, since the relationship between the dependent and independent variables was found to be consistent.

Nevertheless, the Adjusted R-Squared value of the GWR regression model was 0.766, an increase of 0.088 over OLS. Thus, over 76% of the crime density values are predicted by GWR.

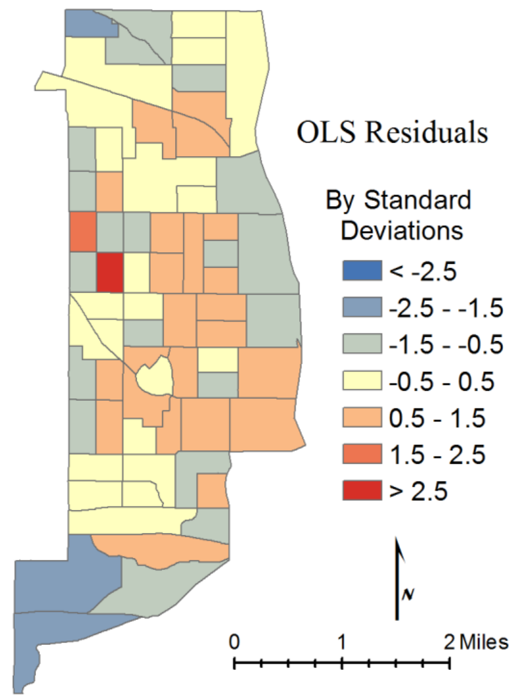

Figure 4. OLS residual standard deviations. Blues indicate census block groups where crime is lower than the model predicted.

The Akaike's Information Criterion (AICc), useful for comparing performance between models with the same dependent variable, also decreased by 12.642 in the GWR model, from 147.067 to 134.425. Any decrease over 3.0 is considered significant. The Spatial Autocorrelation (Global Moran's I) tool, ran on the GWR model residuals, returned a z-score of -0.4397, an indication of a random spatial pattern. Similar to the OLS regression residuals, 94% of the GWR residuals fell within -1.5 and 1.5 standard deviations of the mean (Figure 5). The largest over prediction was -2.14 standard deviations while the largest under prediction was 2.77 standard deviations. This block group was also under predicted to a similar degree by the OLS model.

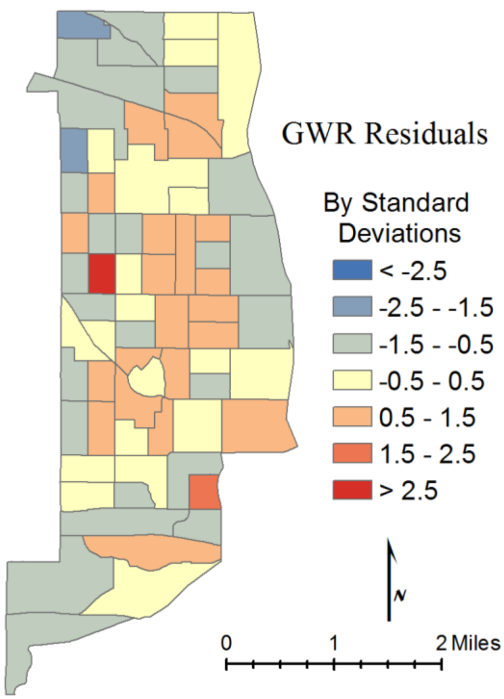

Figure 5. GWR residual standard deviations. Reds indicate census block groups where the regression model underperformed.

Discussion

Findings

Although appearing to be 100% significant in the summary table of Exploratory Regression significance (Table 3), per capita income did not meet the minimum spatial autocorrelation p-value of 0.10, meaning it was clustered to such a degree that it caused regression models to fail if it was included. In contrast, population density was just as significant as per capita income, with the benefit of not being as spatially clustered.

The unpredictable effect population density has on crime was reinforced by it being the most significant independent variable in the OLS regression model, with a p-value of 0.0002. Following routine activity theory, this is an indication that in North Minneapolis, the number of perpetrators and targets is greater than the number of capable guardians.

One of the central tenets of social disorganization theory is that low economic status has a positive impact on crime. Ultimately, this was represented in the final OLS and GWR models by the percentage of the total population living below poverty variable. Although it was found to be significant in only 42% of trials, it had a strong p-value of 0.0002 in the final OLS model.

Table 3. Summary of Exploratory Regression significance for all independent variables after 2379 trials.

Variable	**% Significant**
Per Capita Income	100.00
Population Density	100.00
% Vacant	96.60
% Nonfamily	72.67
Heterogeneity Index	70.78
% Rental	58.31
% Commercial	51.89
Different House %	50.13
% Impoverished	42.32
% Unmarried	41.56
% Unemployed	33.00
% Living Alone	32.32
% Mixed Use	8.06

The heterogeneity index was significant in nearly 71% of Exploratory Regression trials, and had the third most significant p-value of the independent variables. This is another confirmation of the previously-supported hypothesis that communities with a high degree of racial and ethnic heterogeneity (ultimately determined by self-identification on census forms) have greater crime than homogenous populations.

The percentage of vacant housing units, a variable representing the residential mobility factor of social disorganization theory, was significant in over 96% of Exploratory Regression trials.

It was found to be more significant than either the percentage of rental housing units or the percentage of people who were living in a different housing unit than they were a year ago.

Despite having a theoretical effect on crime, no variable representing family disruption was found to be consistently significant in the Exploratory Regression trials to warrant its inclusion in the OLS and GWR regression models. The percentage of households consisting of nonrelatives was found to be significant in 73% of the trials, but the percentage of the population unmarried and the percentage of households living alone were not significant in more than 42% and 32%, respectively, of trials. Though none of the variables representing family disruption were found to be significant, this does not negate or diminish the damaging effects family disruption has on communities. It should be understood that only the variables chosen to represent this theme in this research were found to be not significant.

The percentage of land use classified as commercial was significant in slightly over half of the Exploratory Regression trials. It also had the least significant p-value of the OLS variables at 0.0490, just under the 0.05 threshold. As a result, it has a significant effect on explaining crime density, but less so than the other variables. The percentage of land use dedicated to mixed use, though significant in past research, was found to be significant in only 8% of trials and was not included in the final OLS or GWR regression models. It should also be noted that this research did not assign a qualitative measure to commercial or mixed use land use. As a result, whether or not a parcel is an alcohol establishment or a convenience store, both of which have been shown to lead to increased levels of crime, was outside of this project's scope.

The significant increase in the Adjusted R-Squared value from OLS to GWR and the decrease in the AICc indicate though the independent variables had consistent enough relationships with crime as to not warrant a significant Koenker (BP) Statistic, their relationships were not completely consistent. This is substantiated by mapping the GWR variable coefficients. For example, a map of population density coefficients indicates it was a strong predictor of crime in the southern part of the study area, but underperformed in the northwest (Figure 6).

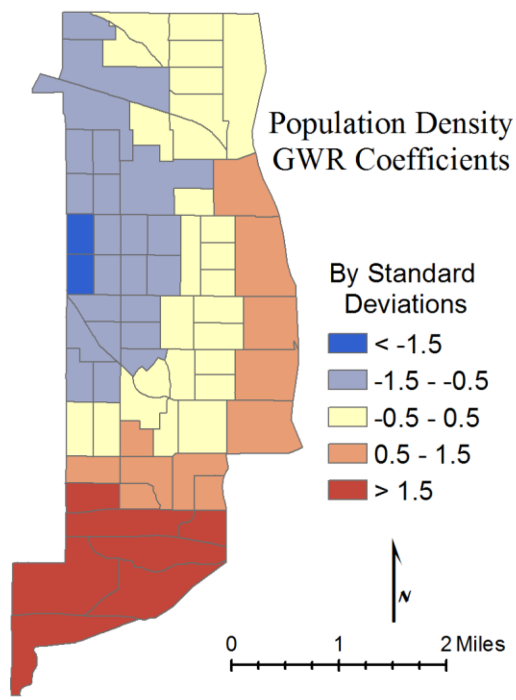

Figure 6. GWR Coefficients for Population Density. Larger coefficients, in red, indicate areas where population density is a strong predictor of crime.

The block group with the largest under prediction of its violent crime density is in this northwest area. Both OLS and GWR models had under predictions for this block group which were more than 2.7 standard deviations from the mean. A brief analysis of this block group offers insight into this under prediction. The area is bounded by the following avenues: Penn to the east, Lowry along the south, Sheridan on the west, and 35th along the north. The block group had 72% of its area classified as residential, and consisted predominantly of single-family houses. An elementary school, park, and large post office building occupied the majority of the remaining land. This means the block group had a low population density in comparison to nearby block groups. It also had a low percentage of its population living in poverty (7.1%), had a low percentage of vacant housing units (7.3%), a very small percentage of commercial land use (0.8%), and an average heterogeneity index (0.67). As a result, both models predicted this block group to have low violent crime density when in fact the crime density was relatively high, at 180 crimes per square mile. A reason for the high crime rate may be because one of the businesses in this block group is a liquor store, which is located adjacent to a multi-story apartment complex and 150 feet from the intersection of Penn Avenue and Lowry Avenue. Both roads are major transportation routes in North Minneapolis. Indeed, further analysis revealed 38% of the block group's instances of violent crime from 2010 to 2014 took place within 250 feet of these two buildings.

An additional parameter estimate calculated by the GWR regression model was crime density predictions (Figure 7). Most of the block groups with high predictions were located along major avenues such as Penn and Broadway. The three block groups along Emerson Avenue which actually had high violent crime densities in 2010-2014 (Figure 1) were not predicted by GWR to have as high a level of crime density as they in fact did. The overall similarity between the two maps, with elevated crime densities in the central region of the study area, is a further validation of the GWR regression model's performance.

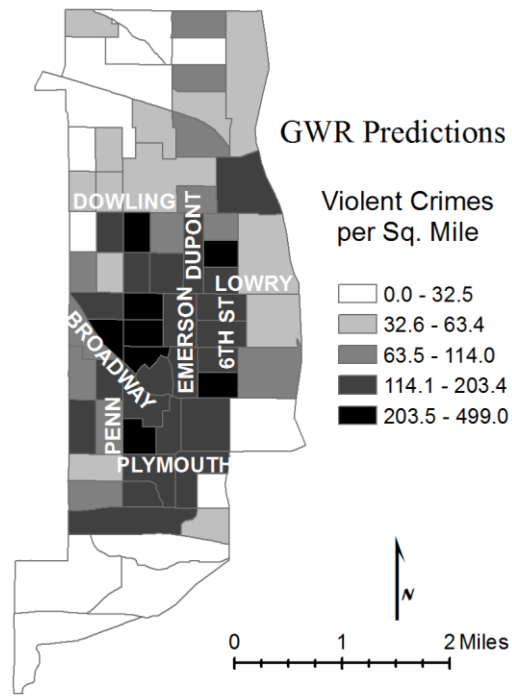

Figure 7. 2010-2014 violent crime densities as predicted by the GWR regression model. Compare to the actual crime densities in Figure 1.

Potential Sources of Error

Crime is a complex subject, and even though GWR was able to account for nearly 77% of violent crime densities, the model could theoretically be improved. When conducting regression analysis, there is always the possibility a potentially significant explanatory variable may have been overlooked or not considered in the planning stages of the analysis. This can lead to models that do not accurately predict the dependent variable (Thompson and Gartner, 2014). It is also possible that independent variables determined to be not significant on their own, such as those associated with family disruption, would become significant if they were combined into a composite variable or index. Cahill and Mulligan (2007) and Stucky and Ottensmann (2009) found composite variables to be statistically significant in their regression models.

Though North Minneapolis was chosen because it is a well-defined area with an established sense of community, it is not geographically isolated. Its proximity to the rest of Minneapolis to the east and south, and the suburbs of Brooklyn Center, Robbinsdale, and Golden Valley to the north and west, means there is potential influence on North Minneapolis from these areas that were not measured in this study.

Suggestions for Future Research

This research validates the feasibility of using GIS-based regression analysis to model crime densities in North Minneapolis. A logical next step would be to expand the geographic scope of analysis to encompass the entire city of Minneapolis. With the understanding that Minneapolis is itself not a geographically isolated city, an analysis including the greater Minneapolis-St. Paul metropolitan area would be of benefit to regional planners.

With the U.S. Census Bureau publishing ACS 5-year estimates on an annual basis, it would be beneficial to conduct a longitudinal study to compare which independent variables remained significant from year to year, and how their significance changed.

Though the OLS and GWR models account for 68% and 77%, respectively, of crime densities, undoubtedly there are additional variables which may add additional understanding. Finding additional significant variables could lead to greater dependent variable explanation in the regression models. It would also be advantageous to assign a qualitative measure to some variables, such as land use. Doing so would allow researchers to account for alcohol establishments, for example, or high density residential development.

Conclusion

Both OLS and GWR regression models offer explorations of violent crime densities in North Minneapolis as both have Adjusted R-Squared values well above the 50% threshold. For OLS, the wealth of diagnostic output further reinforces the fact that it is a properly-specified model. The well-established checks for variable and model significance were successfully passed, and the independent variables were determined to be unbiased with no redundancy. Both models had residuals free of spatial autocorrelation.

GWR did a statistically better job than OLS at modeling crime in North Minneapolis. Its Adjusted R-Squared value indicates it explained 77% percent of violent crime densities while OLS explained 68%. The robust performance of both models reinforce central tenets of social disorganization theory, as low economic status, ethnic and racial heterogeneity, and residential mobility were found to have significant and positive effects on crime. However, the significance of both population density and the percentage of commercial land use also indicate crime densities cannot be completely explained by social disorganization theory.

Though results of this research are static and do not take into account the dynamic and unpredictable nature of crime, it can nevertheless give policymakers valuable guidance. Due to the independent variables' positive correlations with crime, this research indicates programs aimed at reducing poverty and housing vacancies would lead to lower violent crime densities. In addition, this research suggests programs focused on community development and engagement as well as racial and ethnic inclusiveness would also lead to reduced levels of violent crime. These measures could offset the positive correlations of crime and population density, especially since population density is often viewed as desirable in contemporary urban planning. With regards to commercial land use, by itself not necessarily an indication of crime, measures could be taken to encourage positive community-reinforcing development. Such measures may have prevented the construction of a liquor store next to a multi-story housing complex, as witnessed in the block group with the large under prediction.

This research reinforces the notion that crime is a complex phenomenon with numerous variables at play, not easily described by tidy theories. Nonetheless, GIS-based regression analysis such as OLS and GWR can be employed to quantitatively assess ecological crime theories, giving policymakers and the general public an additional tool in the effort to make communities safer.

Acknowledgements

I would like to thank Mr. John Ebert, Mrs. Greta Poser, and Dr. David McConville for their instruction and guidance throughout my time at Saint Mary's University of Minnesota graduate GIS program.

In addition, I would like to thank my family and friends for their support throughout the pursuit of my graduate degree.

References

Cahill, M., and Mulligan, G. 2007. Using Geographically Weighted Regression to Explore Local Crime Patterns. *Social Sciences Computer Rreview, 25,* 174-193.

Chainey, S., and Ratcliffe, J. 2005. *GIS and Crime Mapping*. Chichester, West Sussex, England ; Hoboken, NJ : Wiley.

He, L., Paez, A., Liu, D., and Jiang, S. 2015. Temporal Stability of Model Parameters in Crime Rate Analysis: An Empirical Examination. *Applied Geography, 58,* 141-152.

Leung, Y., Mei, C., and Zhang, W. 2000. Statistical Tests for Spatial Nonstationarity Based on the Geographically Weighted Regression Model. *Environment and Planning, 32,* 9-32.

Metropolitan Council. 2014. *Choice, Place and Opportunity: An Equity Assessment of the Twin Cities Region,* (5). Retrieved February 1, 2017 from metrocouncil.org/METC/files/35/35358ee4-7976-42e6-999d-9e54790d45fe.pdf.

Rosenshein, L., Scott, L., and Pratt, M. 2011. Finding a Meaningful Model. Retrieved February 2, 2017 from www.esri.com/news/arcuser/0111/files/findmodel.pdf.

Sampson, R., and Groves, W. 1989. Community Structure and Crime: Testing Social-Disorganization Theory. *American Journal of Sociology, 94*(4), 774-802.

Stucky, T. D., and Ottensmann, J. R. 2009. Land Use and Violent Crime. *Criminology: An Interdisciplinary Journal, 47*(4), 1223-1264.

Thompson, S. K., and Gartner, R. 2014. The Spatial Distribution and Social Context of Homicide in Toronto's Neighborhoods. *Journal of Research in Crime and Delinquency, 51,* 88-118.

Townsley, M. 2009. Spatial Autocorrelation and Impacts on Criminology. *Geographical Analysis*, *41*(4), 452-461.

U.S. Census Bureau. 2009. A Compass for Understanding and Using American Community Survey Data: What Researchers Need to Know. Retrieved August 8, 2016 from www.census.gov/content/dam/Census/library/publications/2009/acs/ACSResearch.pdf.

Zhang, H., and Peterson, M. P. 2007. A Spatial Analysis of Neighborhood Crime in Omaha, Nebraska Using Alternative Measures of Crime Rates. *Internet Journal of Criminology, 1*.