Master Thesis in Geographical Information Science nr 96

A Spatial Analysis of Homicide Crime's Distribution and Association with Deprivation in Stockholm Between 2010-2017

Dennis Lindell Schettini

2019
Department of
Physical Geography and Ecosystem Science
Centre for Geographical Information Systems
Lund University
Sölvegatan 12
S-223 62 Lund
Sweden

Dennis Lindell Schettini (2019). A Spatial Analysis of Homicide Crime's Distribution and Association with Deprivation in Stockholm Between 2010-2017.
Master degree thesis, 30- credits in Master in Geographical Information Science
Department of Physical Geography and Ecosystem Science,
Lund University

A Spatial Analysis of Homicide Crime’s Distribution and Association with Deprivation in Stockholm Between 2010-2017

Dennis Lindell Schettini

Master thesis, 30 credits, in Geographical Information Sciences

Supervisor: Mahdi Farnaghi

Department of Physical Geography and Ecosystems Science, Lund University

Abstract

This dissertation explores homicide crimes in a spatial context in Stockholm, Sweden between the years 2010-2017. The locations of homicide crimes were obtained from various sources and confirmed in an extensive verification process. The first part of the thesis utilizes GIS-tools to locate homicide hot spots through Nearest Neighbour Hierarchical Clustering (NNH). The second part aims to further address the spatial distribution of homicide crimes by investigating its relationship with material deprivation on a district level through Poisson based regression. Material deprivation was defined by a deprivation index (NDI) consisting of four variables: income; education; unemployment; and social welfare. The combined results show several significant hot spots, and a positive correlation between economic disadvantage and homicide rates. The relationship between homicide crimes and social welfare was particularly substantial.

Table of contents

List of Figures

List of Tables

1. Introduction

Interest in the geographic dimension to explain spatial patterns and causal dynamics of crimes can be traced back to the very beginnings of criminology as a scientific discipline. Crime theorists generally agree that inquiry of the spatial component of criminal behaviour is essential in order to yield valuable insight and knowledge (Messner and Anselin, 2004).

Homicide research includes various disciplines such as criminology, sociology, public health, and geography (Ye and Wu, 2010). Both western societies and developing countries usually list violent crimes as one of its leading public health and social problems (Cole and Gramajo, 2009). The relationship between material deprivation and homicide rate has mainly been studied in the United States and reveals positive correlation between poverty and national homicide rates (Rogers and Pridemore, 2013). The use of spatial analysis is substantively important for detecting different predictors in varying geographic areas and for the diffusion process of homicide crimes (Ye and Wu, 2010).

1.1. Problem statement and goals

Homicide crimes are complex in its nature. They are rare events depending on numerous factors and tend to be distributed differently depending on geographic location. At present, no research has been published that tackles the nature of homicide crimes in a spatial context in Sweden. The first part of the dissertation will strive to fill a gap in the literature, where spatial clustering of homicide crimes in Stockholm will be investigated between the years 2010-2017. The second part will aim to further explore the spatial distribution on an aggregated district level in relationship to a deprivation index (NDI).

1.2. Research questions

The following research questions were composed in order to address the problem statement:

1. Is there significant spatial clustering of homicide crimes in Stockholm municipality? Are the patterns different depending on the inclusion of homicide crimes of adjacent municipalities?

2. Is there a statistical relationship between homicide crimes and material deprivation for Stockholm municipality on a district level?

1.3. Thesis structure

The first part of the thesis addresses spatial homicide patterns through a Nearest Neighbour Hierarchical Clustering analysis (NNH). The second part of the thesis includes an exploration of the relationship between the homicide rates on a district aggregated level and socio-economic factors using Poisson based regression with a Neighbouring Deprivation Index (NDI).

2. Literature review

Homicide crimes have a serious impact on the perceived liveability of a community. They can be used as an accurate indicator of community violence as well as a generalized predictor of general community well-being (Ye and Wu, 2011). Recent research reveals a strong association between poverty and homicide rates (Rogers and Pridemore, 2012). A variety of studies on both micro- and macro levels (Wang and Arnold, 2008; Ye and Wu, 2011; Messner and Zimmerman, 2012; Thompson and Gartner, 2014) have intended to explain the nature of homicide crimes in a spatial context; both where they occur; and which variables that would be most suitable to explain and predict the patterns.

2.1. Homicide research in a spatial context

The empirical research from the early years of the 20th century embraced a spatial context. The common thread in the sociological study of crime was the premise that criminal behaviour was not randomly distributed. This resulted in an understanding that inquiry into the spatial pattering could yield invaluable insight into their causal dynamics (Messner and Anselin, 2004).

Further on, the middle years of the 20th century was characterized with a decrease in spatial interest. This was due to major innovations in survey methodology, where the field of sociology fostered an “empirical behaviourism” (Coleman, 1986). For sociological studies of crime this led to the empirical research changing its analytic focus from social groups and territorial based social aggregates, for instance cities and neighbourhoods, to individuals (Bursik and Grasmick, 1993).

The early 21th century passed an extensive change with an increasingly implementation of formal tools of spatial analysis to describe and explain variations in levels of homicide and other crimes. The rediscovering of the importance of geographic information in crime studies was driven primarily by recognized and influential place-based theories such as crime pattern theory where geographic space plays a central role in the flourishing research on crime hot spots and the diffusion of violence and neighbourhood collective efficacy. Both the increased availability of georeferenced information on crime events and the proliferation of spatial analysis to the empirical research in criminology have had a crucial role in the reestablishment of a geographic interest, resulting in that geographic space returned to the forefront of criminological inquiry (Messner and Anselin, 2004).

2.2. Criminal theories and previous homicide studies

A vast amount of leading criminology theories imply at least two common denominators; the positive relationship between concentrated disadvantage or income inequality and crime; and that more densely populated areas increases the occurring of crimes. Spatial studies of criminal behaviour are often based on three major theories: (i) routine activity theories, (ii) strain theories, and (iii) social disorganisation theories.

(i) Routine activity theories argue for a crime generating scheme based on the fact that perpetrators makes conscious risk assessments before committing crimes. The scheme consists of three essential elements: motivated perpetrators; potential targets; and the lack of capable guards. Economic inequality increases the potential targets which encourages more motivated offenders, resulting in more crimes. Removing one of these three elements from the equation would mean that crimes could be prevented (Wang and Arnold, 2008).

(ii) Strain theories proposes that disadvantage and inequality tend to create pressure when the less fortunate individuals compare themselves to the more prosperous. Higher strain and frustration due to the interpreted injustice from these individuals would then lead to more acts of violence (Wong, 2012).

(iii) Social disorganisation theories argue that high rates of crime in urban areas can be predicted by residential instability, poverty, and ethnic heterogeneity (Wong, 2012). The residential instability may weaken the communities' social networks and strain resources to deal with the settlement of new members (Sampson and Raudenbush, 1997). More impoverished communities lack the ability to satisfy their inhabitants, where especially the youth activities become harder to monitor and control (Bursik and Grasmick, 1993). Social networks in the communities may struggle to be developed either due to ethnic heterogeneity or to a high proportion of ethnic minority populations, resulting in higher crime rates (Sampson and Groves, 1989). Social disorganisation theories and routine activities theories are especially represented in homicide studies, where multivariate statistical techniques have become an important area of quantitative criminological research (Rourke-McBride, 2014).

Various schools of theories argue for certain angles as essential in the explanation of the spatial relationships between crime and deprivation (Wang and Arnold, 2008). Shifts in crime rate trends, changes in social and economic trends, and theoretical developments has lead researchers to develop and introduce specific structural indicators as independent predictors to crime rates. The most successful of these studies involves the examination of variables such as racial composition, social bonds, concentrated poverty and labour market conditions (McCall et al., 2011).

Hipp (2007) suggested in a large study which included nineteen cities that there is strong evidence of the importance of ethnic heterogeneity for crime rates, and that material inequality within the ethnic group is strongly correlated with an increase in crime rates. Lafree (1999) argued in his extensive research over United States' crime rates that economic stress, political legitimacy, and family disorganisation are the major contributors to homicide crimes. He continued to claim that homicide research in general lies in the complex nature of finding full and consistent variables for different groups on different geographic levels.

The youth as a group has been target for research. Wong (2012) have studied youth crime in Canadian municipalities where he showed that the effect of single parenthood was positive whereas that of divorce was unexpectedly negative. McCall et al. (2008) suggest that the relative size of the youth population is connected to homicide rates; where a higher percentage of youths in a population is positively associated with higher homicide rates.

No significant differences have been recognized regarding neighbouring effects on violent crimes and homicide crimes. The findings of national wide studies in the United States (Kirk and Laub, 2010; Peterson and Krivo, 2010; Messner and Zimmerman, 2012) are remarkably consistent across regions of the country, cities of different sizes, and levels of homicide. Urban neighbourhoods characterized by high levels of economic disadvantage, single parent families, and racial isolation and inequality with high levels of violent crimes also have higher homicide rates (Thompson and Gartner, 2014).

2.3. Methods of previous homicide studies

As mentioned earlier, previous research has focused on finding the most suitable predicting variables for homicide rates on different geographic levels. Multivariate statistical techniques are the most common methods to use in order to investigate statistical significance. Linear regression models are commonly used techniques, where especially Ordinary Least Square regression (OLS) has been widely adopted in studies (Altindag, 2012; Thompson and Gartner, 2014; Rourke-McBride 2014). Studies find that homicide can be contagious, i.e. communities adjacent to high crime areas, independent of their social and structural development, can by their geographic location experience higher crime rates. Wang and Arnold (2008) argue for that the spatial dependence, i.e. the correlation between homicide rate at one location and homicide rates at nearby observations, will result in inconsistent and inefficient estimators by

OLS. The misrepresented spatial dependence can be mitigated by the implementation of spatial regression models.

The relationship between homicide rate and socioeconomic factors at community area level in Chicago from 1960 to 1995 was studied by Ye and Wu (2011). Exploratory spatial data analysis (ESDA) was applied to investigate dynamic spatial patterns. These methods provide measures of global and local spatial autocorrelation, which has latter been estimated in their regression analysis. The study identified various hot spots over time, and found that neither OLS nor spatial regression could generate consistent findings for all the time points and addressed this by using spatial panel regression.

Wang and Arnold (2008) aimed to explain homicide rates in Chicago by using social disorganisation theory. Localized income inequality index was implemented in GIS to measure economic inequality. The study illuminated the problem with spatial autocorrelation in homicide research and uses OLS regression as a reference to a spatial lag model and a spatial error model.

The social ecology of homicide in Toronto between the years 1988-2003 was examined by Thompson and Gartner (2014). OLS was used to analyse the structural correlates of 965 homicides occurring in 140 neighbourhoods. Negative binomial models were used to analyze the rare events of homicides instead of basic Poisson distribution models due to overdispersion in the data.

Osgood (2000) suggested the use of Poisson regression models when analyzing crime rates. The study illuminated the problems with the usage of OLS for small populations and low-base rates, in this case for juvenile arrest rates as a respondent variable. The result suggested that OLS regression can be ill-suited for smaller numbers of offenses due to inappropriate distributions and insufficient accuracy. Regression models based on Poisson distribution would instead be preferable for the purpose because they are built on assumptions about error distributions that are consistent with the nature of event counts.

2.4. A cross-national context and research in Stockholm

Thompson and Gartner (2014) argued that the city of Chicago is the birthplace of theory and research on the social ecology of crime and has been subject of more studies of how neighbourhood context shapes homicides and other crimes than any other city in the world (e.g., Zimmerman and Messner, 2011; Browning et al., 2004; Bursik and Grasmick, 1993;

Shaw, 1929; Sampson, 2012; Block and Block, 1992). The research regarding "neighbourhood effects" has after the turn of the century extended to other American cities (e.g., Hannon, 2005; Kubrin, 2003; Martinez et al., 2008; Lee et al., 2001; Peterson et al., 2000). The findings suggested significant consistence across the regions of the country, cities of different sizes, and level of homicide rates (Thompson and Gartner, 2014).

Since most of the research derives from the United States, questions have been raised on how well the findings can be generalized to resolve foreign cities. Sampson and Wikström (2008) have stated that we know surprisingly little about interpersonal violence in a cross-national, comparative context. Violence tend to correlate with structural factors such as neighbourhood poverty. However, research lack to resolve whether concentrated poverty, defined similarly, can explain cross national differences in violence, and how distributions of inequality directly compare at the neighbourhood level.

Sampson and Wikström (2008) aimed to address the neighbourhood social order of interpersonal violence in the cities of Chicago and Stockholm in order to link comparative measures of structural inequality with community-level mechanisms to predict violence despite the different national settings. Their findings suggested distinct similarities in the social structural characteristics of neighbourhoods associated with violence.

The comparative study by Sampson and Wikström (2008) examines however non-lethal violence and not homicide. The difference between the United States and other industrialized countries is not significant when it comes to violent crimes. However, the difference is substantial when it comes to homicide crimes. The nature of homicide crimes in the United States tend to be more male-dominated, more likely to be committed with firearms, more highly concentrated in urban areas, and less likely to involve family members or intimate partners. Because of these differences, the nature of homicide explaining neighbouring factors cannot be assumed to parallel those in other nations, even those with similar economic and political structures (Thompson and Gartner, 2014).

The measures of neighbourhood characteristics are not in general comparable between less-developed countries and the western world. However, independent of location and economic status of a country, economic disadvantage on a local level is consistently and positively associated with higher homicide rates (Thompson and Gartner, 2014).

Studies regarding the relationship between neighbourhood and homicide studies in the west is sparse. This scarcity is supported by Thompson and Gartner (2014) which could only find one

English-language study conducted in a western industrialized country, a Dutch study by Nieuwbeerta et al. (2008). During the research for this dissertation, no Swedish study regarding the relationship between homicide crimes and material deprivation was found.

Apart from neighbourhood correlates with homicide rates, one study could be found regarding spatial clustering of crimes in Stockholm. Uittenbogaard and Ceccato (2012) produced space-time clusters of crimes in Stockholm between the years 2006-2009 using Kulldorff's scan test. Their findings suggested distinct patterns regarding concentration of violence and property crimes over time and space. Violent crimes tend to happen more often during the night and is heavily concentrated in large parts of the city centre, with some extension to socially disorganized areas in the west and south Stockholm.

2.5. Study area

Stockholm municipality has the largest population out of all the 290 municipalities in Sweden, and is also the most densely populated. It is divided into fourteen district areas based on geographic location, and these district areas are further divided into smaller communities. The district areas are of interest in this dissertation and will be referred to as "districts". The districts' department boards are responsible and influential of issues regarding their own administrative areas, deciding how funds shall be allocated to satisfy local needs (Stockholm Stad, 2018). Stockholm municipality was selected as study area due to its high population density, prominent status and administrative division of districts, making it perhaps the ideal municipality in Sweden to study in terms of homicide crimes in a spatial context.

The municipality itself divides the districts in two fractions: the "inner-town"; and the "outer-town". Due to major differences in the neighbouring context between these two fractions, the regression analysis was solely made on the "outer-town". The "inner-town" might be better analysed separately, where important variables such as tourism and night life activities could be integrated in a regression model. The "inner-town" constitutes of Östermalm, Norrmalm, Kungsholmen, and Södermalm. The "outer-town" included in the regression analysis consists of: Rinkeby-Kista; Spånga-Tensta; Enskede-Årsta-Vantör; Skarpnäck; Skärholmen; Hässelby-Vällingby; Hägersten-Liljeholmen; Farsta; Bromma; and Älvsjö. Figure 1 includes the whole municipality, while Figure 2 shows the districts included in the regression analysis.

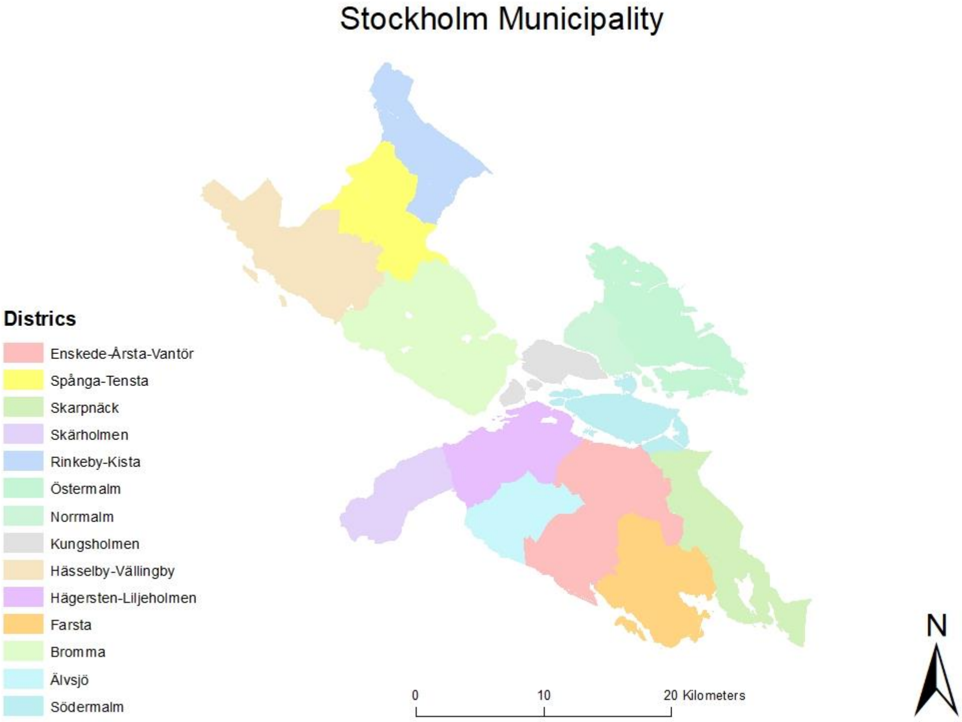

Figure 1. Stockholm municipality with containing districts.

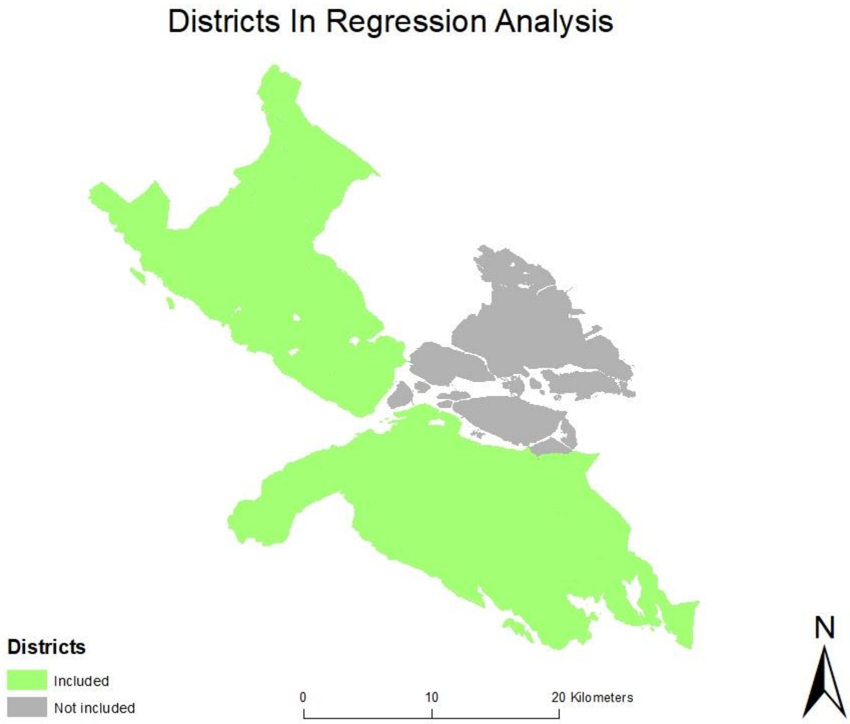

Figure 2. Stockholm municipality with outer-town shown in green and inner-town shown in grey. Only outer-town was used in the regression analysis.

2.6. Homicides in study area

The Swedish National Council for Crime Prevention (Brå) is responsible for the official national crime statistics, which includes producing, following, analysing, and reporting on criminality and the criminal justice system's responses to crime. Brå generates statistics which are based on large-scale surveys and other special data collections (Bra.se, 2018).

This dissertation includes homicide crime identification on a smaller scale than Brå, which means that comparisons can be made on a district level rather than a region level. When comparing homicide rates with the national rate, several districts in Stockholm municipality are subject to particular high values of homicide rates. Figure 3 shows the aggregated homicide rates within the study area. Figure 4 illustrates the rates in comparison to the national homicide rate.

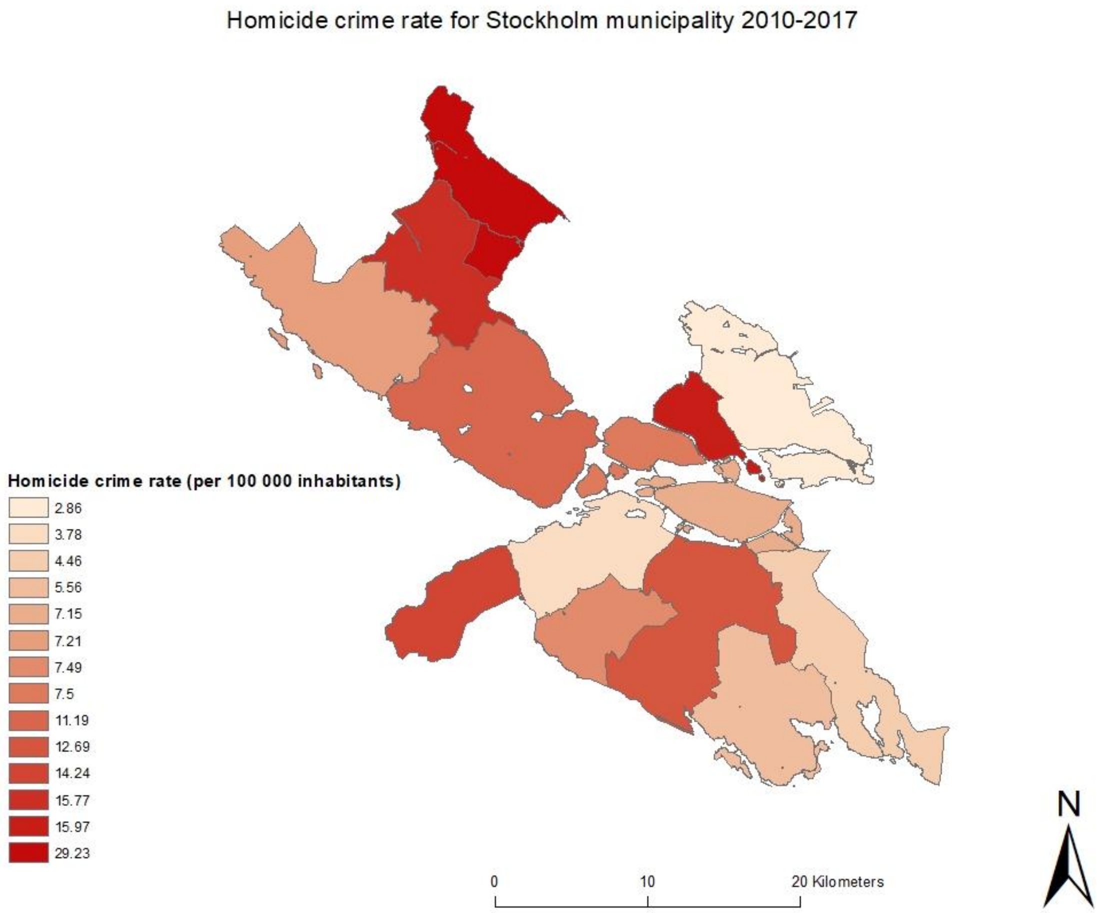

Figure 3. Aggregated homicide crime rate for Stockholm municipality between the years 2010-2017 (per 100 000 inhabitants).

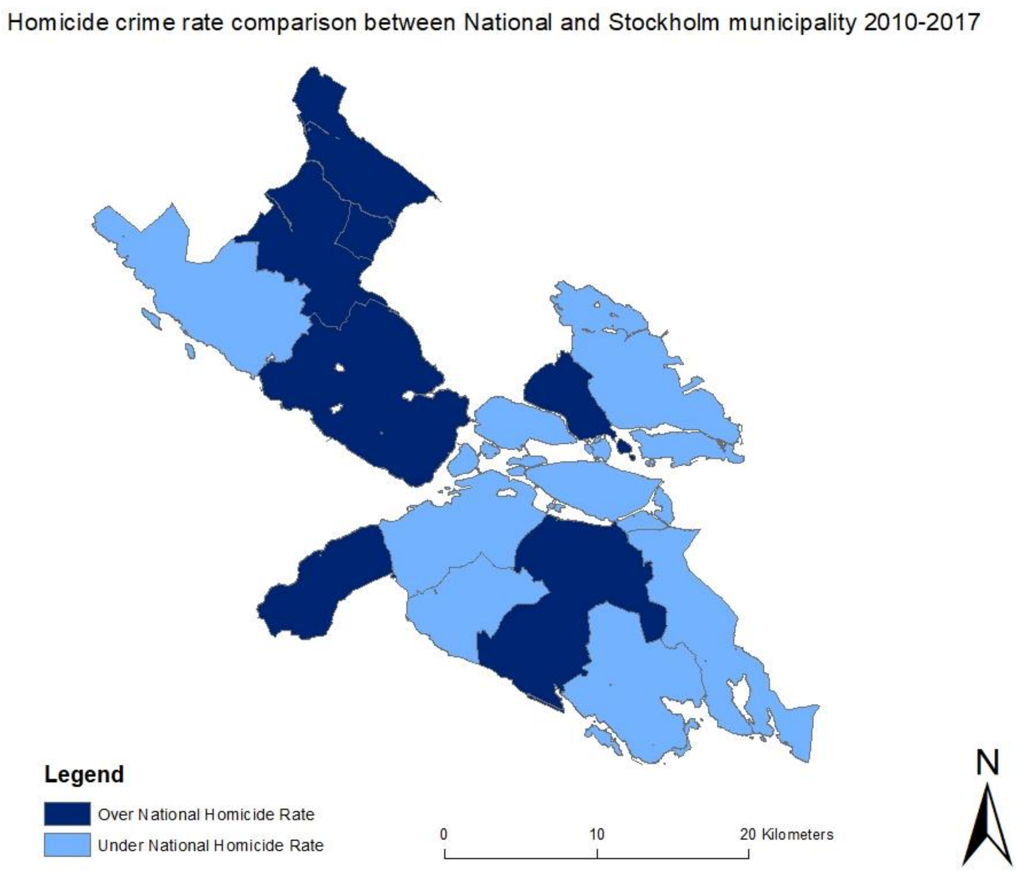

Figure 4. Districts in Stockholm municipality which lies above or below the national homicide rate between the years 2010-2017 (per 100 000 inhabitants).

2.7. Summary

Modern research acknowledge a strong relationship between poverty and homicide rates (Rogers and Pridemore, 2012). Studies on micro and macro levels tend to break down the term poverty into specific constituting variables in order to find the variables suitable to explain and predict homicide crime patterns.

Major criminology theories include routine activity theories, strain theories, and social disorganization theories. These theories suggests divergent explanations on the spatial relationships between crime and poverty. However, they agree upon the fact that there is a positive relationship between concentrated disadvantage and crime, and that more densely populated areas increases the occurring of crimes.

Research on homicide studies in a spatial context are mainly based on studies in the United States. This is problematic due to the fact that the nature of homicide explaining neighbouring factors in the United States cannot be assumed to parallel those in other nations, even those with similar political and economic structures.

3. Data and Methodology

The dissertation's methodological approach was divided into three steps. The first step consisted of obtaining homicide crime data and deprivation data and digitize it into a GIS-database. This step included an extensively verification process on homicide crimes coordinates. The second step addressed the first research question through spatial clustering data analysis. Lastly, the third step included a deeper understanding in the spatial distribution of homicide crimes by assessing economic disadvantage as an explanatory variable in generalized linear regression models.

3.1. Homicide data verification process

The criterions for the homicide crimes used in the dissertation follows the exact definition used by Brå. To be considered as homicide crimes, they were required to meet any of the following juridical conditions: murder; manslaughter; involuntary manslaughter as a result of assault and battery; or infanticide. Cases which involved killing in self-defence were not considered as homicide.

The verification process followed the same structure as Brå's own verification process, which is defined as:

"*The information has been verified on different ways depending on how long the matter has come in the criminal trial procedure. If the matter has led to a judgment, the information is verified firstly through the judgment and secondly by withdrawals of courts' information on legally binding judgments or from superior courts in the database of Brå. If the matter is still under investigation it is verified through oral or written information from the police or prosecutor which has handled the case or is requisitor*" (Brottsförebyggande rådet, 2017).

The complex nature of obtaining the homicide data gave rise to an extensive verification process in order to establish reliability. All data samples went through the same process, which consisted of several control segments. The verification process was developed as a necessity to interconnect and compare data from various sources, where the sources could be broken down into a spectrum of reliability.

Sources of higher reliability were original electronic- and paper documents of verdicts from involved courts; preliminary investigation protocols from the Swedish Police authorities; electronic event reporting from the police; mail exchange with local police authorities; and statistics from governmental institutions. As a general rule, data deriving from these sources would automatically be considered confirmed. Sources of lower reliability would include media records; footage of crimes scenes; quotations from officers in charge or other individuals working for the Police published in newspapers. For these lower reliable sources, investigations and multiple comparisons had to be made in order to guarantee reliability.

The verification process model was constructed to manage three types of homicide crime data. The first type contained homicide crimes where perpetrators had been sentenced through a district court, a court of appeal, or a supreme court. The second type involved cases where the prosecutor brought a prosecution towards the accused, but the judgment has yet not taken legal effect due to appeals. Issues arose for these second types when the cases involved questions of guilt – i.e. whether or not the defendants actually had committed crimes. The third type consisted of cases where the Police had classified a crime according to the thesis' criterions for a homicide crime in the preliminary investigation, but no perpetrators had been identified, meaning that large amounts of information had been classified and thereby not obtainable. This type was the most complex type to account for and one of three main reasons why the thesis would not stretch further back in time than the year of 2010.

The verification process consisted of several steps. The order of these steps differed from homicide crimes committed between the years 2012-2017, and the years 2010-2012. In general, convictions from 2012-2017 were available as electronic documents which had to be purchased after a certain amount. Due to the Swedish Data Protection Act (2015:728), the majority of the courts decided to interpret the law as very restrictive in text search. Searches could not be designed to handle the specific search words of interest (for instance homicide). Instead only the parent category "crime against person" could be distinguished in the text search, which would have to be inspected manually by the purchaser for its relevance.

The majority of the convictions before 2012 were not available for digital searching, hence only accessible in paper form except if the exact date of a conviction was known. These convictions were stored in large volumes containing criminal cases and civil cases, mixed and without table of contents. As a reference point, two years' worth of convictions for a single court would be held in twenty-five volumes, each volume containing approximately 1500 pages. This extensive search for convictions older than 2012 was the second main reason to the demarcation in the thesis' time-interval stretch.

Investigating the reliability followed five steps in the verification process. The first step involved the analyzing of conviction documents from the relevant courts. The courts investigated were: Stockholm District Court; Nacka District Court; Solna District Court; Södertälje District Court; Attunda District Court; and Norrtälje District Court. A major part of the data samples could be verified through these courts' documents. The crime scenes would not always be mentioned with specific coordinates. Instead further analysis had to be made in the documents to establish the meaning of terms such as "*residence*" and "*courtyard*". The uniqueness of every conviction document made the procedure of establishing crime-scene specific coordinates mutable.

A second step in the verification process was to order non-classified preliminary investigation protocols from the Police of cases where the first step was not sufficient. A vast majority of these homicide crimes were non-finished cases and thereby classified.

The third step included to establish an aggregated judgment based on statements made by the Police authorities on their event-reporting sources; interviews and quotations from the Police through various media sources; and mail exchange with local police departments. This step was particular essential for unsolved homicide crimes. If the last announced classification of a crime by the Police was stated to be a homicide crime, it was used as data.

The fourth step involved searching through recognized media sources for crime-scene photos and reports. Information from media sources could involve trial dates which could then be used for text searching at the courts for convictions between the years 2010 - 2012. Photos of crime-scenes were especially useful as a compliment to all cases and compared between multiple media sources, and analysed together with Google Satellite and Google Street View for visual confirmation.

The last step comprised of comparing all homicide crimes obtained with the total amount of homicide crimes for Stockholm region, which was provided by Brå. Brå began to provide the public with statistics of regional homicide crime statistics in 2010. In order to compare this study's homicide data for Stockholm with the total regional statistics, homicide data for all homicide crimes of the whole region (Region Stockholm) were obtained. Total regional statistics were not obtainable before the year of 2010. This was the third main reason why the time-interval was not stretched further back in time than 2010.

Classified material together with the cumbersome process of obtaining especially older data made the data sampling extensive in time. Media and newspapers were in general accurate on crime events and their approximate locations. A study based only on media reports would probably have yielded faster information and the ability to go further back in time. However, due to its unreliable nature this was not a preferable data collecting routine suitable for this dissertation, where instead reliability and preciseness of the data were prioritized.

3.1.1. Data limitations

This dissertation was mainly limited by the time-consuming process and restrictiveness regarding homicide crime data. The extensive search of convictions older than 2012 was a major contributing factor to the demarcation in the thesis' time interval.

The sample size can be considered to have had an effect on the overall reliability of the regression and clustering conclusions. Because of the small sample size, extreme events such as the terrorist attack in Norrmalm 2017 with several casualties, did affect the clustering pattern to a great extent. Accounting for a larger time period would have fine grained the calculated rates and offered less sensitivity to extremes.

3.2. Hot spot analysis

This dissertation uses homicide crime data in Stockholm to assess spatial clustering. Hot spot analysis is an umbrella term for analysis techniques which focuses on identifying statistically significant hot spots on a given surface. Nearest Neighbour Hierarchical Clustering (NNH) was chosen as hot spot method and done in the software CrimeStat III (see appendix 8.2).

3.2.1. NNH

NNH identifies groups of incidents that are spatially close. This method is a hierarchical clustering routine where points are clustered together on the basis of a criteria. The clustering is repeated until either all points are grouped into a single cluster or else the clustering criteria fails (Levine, 2004).

Hierarchical clustering methods are among the oldest cluster routines (King, 1967; Everitt, 1974). Various forms of spatial clustering methods have been adopted in spatial analysis studies, including the centroid method, median clusters, group averages, nearest neighbour method, and minimum error method (Levine, 2004). Research suggests that crime incidents often are generated from micro-environments defined by neighbourhoods (Wachs and Shirazi, 1986; Gordon and Friedman, 1989). The hierarchical approach of NNH offers identification of small geographical environments of concentrated incidents.

3.2.2. NNH routine

The NNH routine is based on two criteria: threshold distance; and minimum number of points. The threshold distance was determined using the default random nearest neighbour algorithm. It is defined by a method that compares the threshold to the distances for all pairs of points. The method compares pairs of points to a distance expected in a random distribution of points in the jurisdiction's area, clustering groups of pairs that are unusually close together (Levine, 2004)

The second criteria is the minimum number of points that are required for each cluster. It is used to reduce the number of very small clusters as well as reduce the likelihood that clusters could be found by chance (Levine, 2004).

With these criteria, first order clusters are constructed of the points. The analysis is then conducted again on the first-order clusters where clusters that are unusually close together will be circled in second-order clusters, thus providing a hierarchical approach (Levine, 2004).

Due to the user's ability to accustom the analysis' parameters, including the minimum number of points and threshold distance, it is possible to identify and measure the level of clustering on smaller sample sizes (Levine, 2004). The linkage between smaller clusters can be seen through higher level clusters, displaying hot spots that are adjacent to other hot spots.

3.2.3. Statistical significance and limitations of NNH

Testing the statistical significance of a cluster analysis is a complex task. The NNH routine needs to cluster pairs of points and cluster as many points as possible that fall within the threshold distance. There is also an additional requirement of the minimum number of points defined by the user. The probability distribution is not known for the given situation and is resorted with a Monte Carlo simulation of randomness (Levine, 2004).

As with other clustering techniques, there are some technical and theoretical limitations to hierarchical clustering. When the confidence interval around the mean random distance is used as the threshold distance criteria, the size of the grouping area would be dependent of the sample size, meaning that crime distributions with many incidents would have smaller threshold distances than distributions with fewer incidents. Also, establishing the minimum points per cluster leaves room for arbitrariness as the user is obligated to define a meaningful cluster size. The size of a hot spot may be interpreted differently by two users, thus leaving room for subjectivity as a premise for the result. Almost all clustering techniques allows users to adjust the parameters and by that being a subject to manipulation and this involvement of subjectivity constitutes a statistical weakness (Levine, 2004).

3.2.4. Parameters used

Random nearest neighbour was used for the threshold distance where the p-value was selected as $p<0.5$. This p-value groups pairs of points that have less than 50% probability of being randomly allocated, so there is a 50% chance that the points are actually clustered and not a statistical fluke (Levine, 2004).

For this study, the size of a neighbourhood could be considered a significant cluster. With this in consideration, the p-value was set to $p<0.5$. A smaller threshold value would simply resulted in fewer clusters. Both the threshold distance of $p<0.5$ and the number of simulations made (1000 times) follows the recommendations of Levine (2004) in order to yield a meaningful analysis. The choice of minimum number of points per cluster was set to three. A higher value in this parameter greatly reduces the likelihood that CrimeStat will identify “false positives” in its NNH analysis (Levine, 2004). Considering the quantity of data points for the given time interval, the minimum value of three points per cluster was considered representative on a district scale. The parameters used is shown in table 1.

Type of analysis: Nearest Neighbour Hierarchical (NNH)
Threshold radius: Random Nearest Neighbour ($p>.0.5$)
Minimum points per cluster: 3
Standard deviation ellipses: 1.5
Simulation runs: 1000
Distance measurement: Indirect (Manhattan)

Table 1. Parameters used in CrimeStat.

The visualisation parameters involved the geometric shape of the hot spots and the specification of numbers of standard deviations for the shape. Ellipses were used because of the improved readability they offer. The value of standard deviation was set with regards to the recommended value of 1.5. A lower setting on standard deviations can be hard to view at a small scale, in comparison to a higher value which tend to exaggerate the size of the hot spot (Levine, 2004).

3.3. Regression analysis

Regression analysis studies the complete process of the casual relationship between a dependent variable and a set of independent, explanatory variables (Rogerson, 2001). The statistical relationship between homicide crimes and socio-economic factors were analysed using Poisson based regression, which is a generalized linear model form of regression.

Problems occur when analysing aggregate crime rates. When the population size of an aggregate unit is small relative to the offense rate, crime rates must be computed from a small number of offenses. Poisson based regression was chosen as method since it is more appropriate to use than traditional regression when analysing rare events due to the method's assumptions about error distributions that are consistent with the nature of event counts (Osgood, 2000).

Homicide crime rate was set as the dependent variable, and a deprivation index containing four variables (NDI) was used as explanatory variables. An average value for every explanatory variable from the time interval of 2010-2017 was derived and used as a representative for every single district.

The Poisson distribution is a probability density function that models the number of outcomes obtained in a given interval of time and space, where its variance is equal to its mean. It is used as an approximation of the binomial distribution, assuming that the timing of the events is random and independent (Gooch, 2011). Due to the nature of its variance, the probability of observing specific number of events declines as the mean count increases. This would result in a broader range of values having relevant probabilities of being observed (Osgood, 2000).

Overdispersion is ubiquitous in analyses of crime data (Osgood, 2000). A negative binomial regression model that allows for overdispersion was used as a reference to the basic-Poisson model in the analysis and compared in statistical fit. The statistical fit was measured using Akaike information criterion (AIC). By comparing the models' AIC, it can be read which of the used models that produce a probability distribution with the smallest discrepancy from the true distribution. A lower AIC value would indicate a better fit (Busemeyer and Diederich, 2014)

In order for the explanatory variables to have the same influence on the regression analysis, the census data was standardised with the use of z-scores (Equation 1). The homicide crime rates for the districts were calculated per 100,000 people (Equation 2).

$$Zx = \frac{Xi - \bar{x}}{Sx}$$

Equation 1. Calculation z-score using x variable, sample mean, and sample standard deviation.

$$Homicide\ crime\ rate = \frac{Homicide\ crime\ count}{District\ Population/100{,}000}$$

Equation 2. Calculation of districts' homicide crime rate.

3.3.1. Spatial autocorrelation

Literature suggest that homicide crimes can be contagious, i.e. higher crime rates for districts can affect adjacent districts. In regression analyses, omission of spatial autocorrelation might bias parameter estimates and yield incorrect standard error estimates (Mohebbi et al., 2011). It is implausible that a spatial dataset will have a spatially stationary distribution. In order to explain these relationships, it is possible that a locally model need to be implemented in order to account for the variations (Rourke-McBride, 2013). Homicide events for the study area were tested with Global Moran's I. The method evaluates whether the pattern expressed is clustered, dispersed or random. In order to identify spatial outliers, the districts' homicide rates were also tested with Local Moran's I. This method measures if features' values are surrounded by similar or dissimilar neighbouring features' values (Fu et al., 2014).

3.4. Deprivation index

A vast majority of the studies have been done in American cities, where material deprivation not necessarily follows the same standards as material deprivation in Sweden. In regards to this, recognized Swedish deprivation variables were used for the study.

The characterization of districts' socio-economic environments was represented by a Neighbourhood Deprivation Index (NDI), containing four variables; low educational status; low income; unemployment; and social welfare (see appendix 8.4). A higher NDI-score indicates higher rate of material deprivation. The index has been used by Winkleby et al. (2007) as a deprivation indicator in their Swedish study about coronary heart disease fatality.

The index involves general deprivation variables that can be calculated from census data. The index was applied because of its previous usage in a Swedish study as well as the good fit to available statistics on the district level. The relationships between the deprivation variables and how they may interact with each other were not investigated further in this dissertation and was neither so in the study by Winkleby et al. (2007).

The common thread in the literature is that the definition of deprivation as an explanatory variable for crime rates is no different than for other respondent features, for instance general community health. Major theoretical and empirical developments in the field of criminology during the past 50 years suggest that the same social environmental factors that predict geographic variation in crime rates may also be relevant for explaining community variations in health and wellbeing (Kawachi et al., 2001).

4. Result

Nearest Neighbour Hierarchical Clustering (NNH) was applied to two different study areas and evaluates the data from different perspectives. The data set containing Stockholm municipality will be referred to as D1, while the data set containing Stockholm municipality with adjacent municipalities will be referred to as D2.

4.1. NNH: Stockholm municipality (D1)

Seven clusters of first order are identified in the analysis within the range of three to eight points per cluster. Figure 5 illustrates the spatial patterns of the hot spots deriving from the municipalities' fourteen districts. Four of the clusters are identified in the inner city while three are identified in the outer city. The highest cluster containing eight points and also densest with a measure of 35.5 is located in Norrmalm. The second highest cluster with seven points is located in Rinkeby-Kista. Rinkeby-Kista is furthermore together with Södermalm the only districts which contains two hot spots. The largest cluster is located in Södermalm, with an area of 1.429 km^2 containing four homicide crimes. The hot spots are exclusively located in the inner city and the north western part of the municipality. Besides a small stripe, no hot spots are identified south of the inner city. No second order clusters were identified.

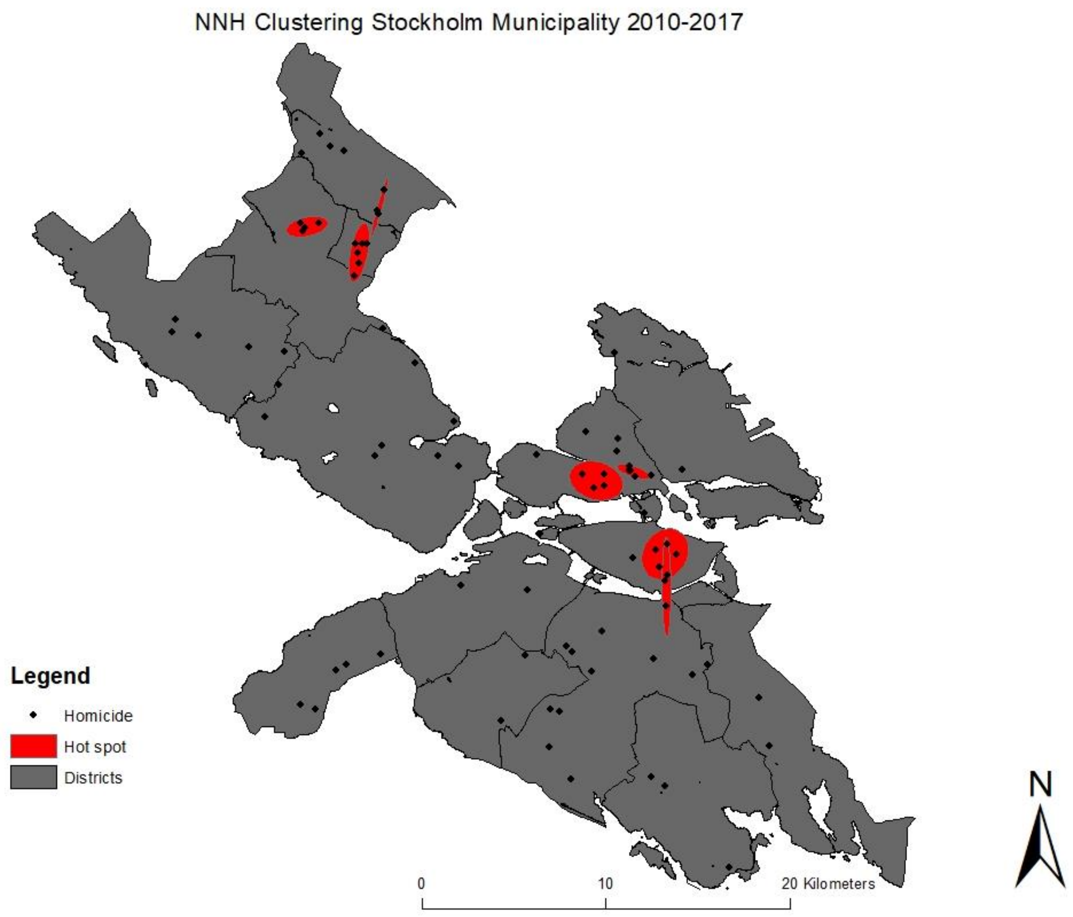

Figure 5. Hot spots over Stockholm municipality generated by Nearest Neighbour Hierarchical Clustering between the years 2010-2017. Clusters are shown in red ellipses and the homicide crimes' locations are shown in black dots. Several homicide events has occurred on the same or nearby locations as others, which means that the visual representation of the black dots not necessarily is representative as they are overlaid. See appendix 8.3 for exact numbers of homicide crimes per district.

4.2. NNH: Stockholm together with adjacent municipalities (D2)

With the addition of the areal extent and homicide crimes committed in adjacent parishes, twelve first order clusters and one second order cluster are identified. Figure 6 shows the spatial distribution of hot spots over the extended area. A larger version of previous highest hot spot in Norrmalm continues to be the highest, with fourteen points and less density. The second highest hot spot is located in the central part of Sundyberg. The most clustered hot spot is located in the southern part of Rinkeby-Kista, measuring nine in density. Several other hot spots are identified, with a majority of them located in the north western districts. Three clusters are identified in the southern parts, where one appeared in Enskede-Årsta-Vantör, which in the previous data set did not contain a hot spot. A second order cross-border cluster is identified containing five first order hot spots from four districts; Spånga-Tensta (1); Sollentuna (1); Sundbyberg (1); Rinkeby-Kista (2).

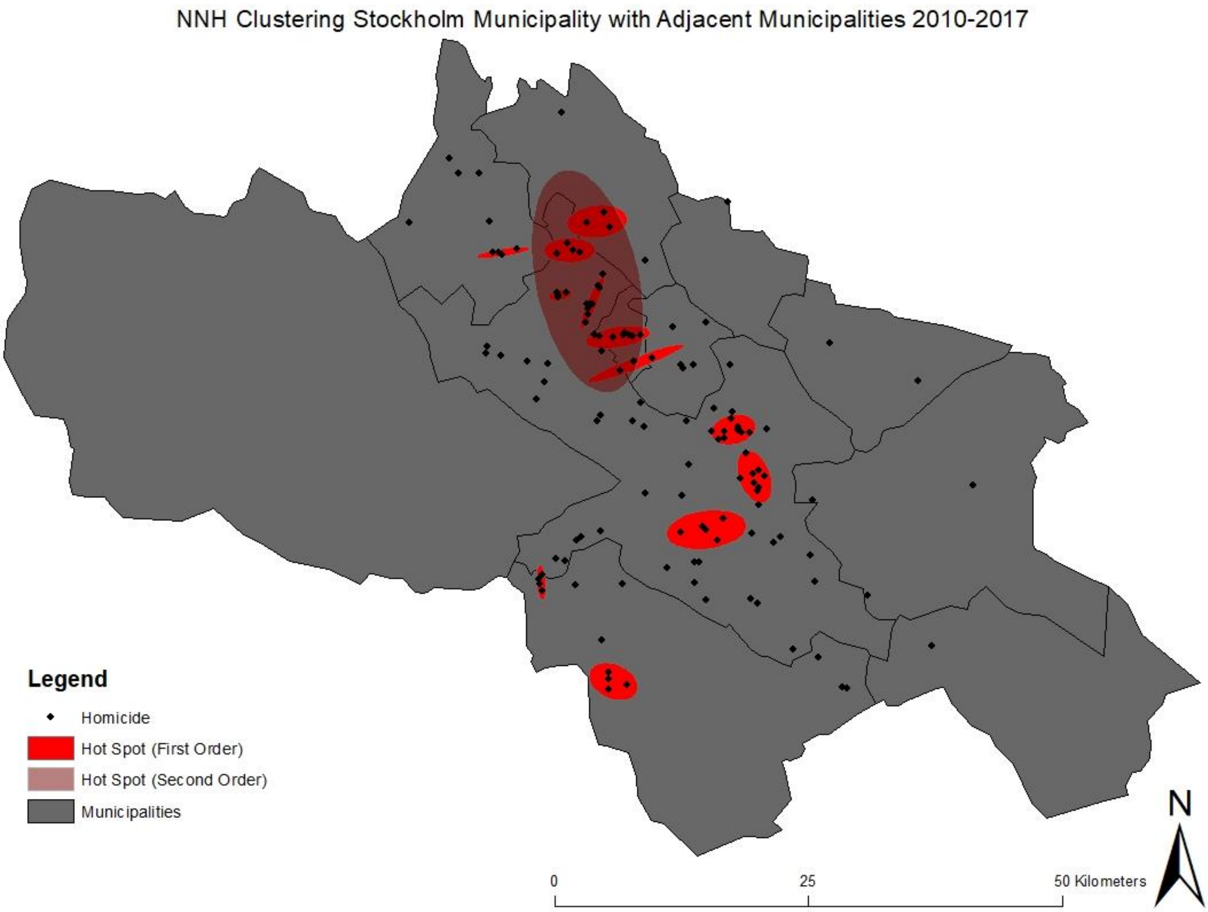

Figure 6. Hot spots over Stockholm municipality together with adjacent parishes generated by Nearest Neighbour Hierarchical Clustering between the years 2010-2017. Clusters are shown in red and the homicide crimes' locations are shown in black dots. Several homicide events has occurred on the same or nearby locations as others, which means that the visual representation of the black dots not necessarily is representative as they are overlaid. See appendix 8.3 for exact numbers of homicide crimes per district.

4.3. Statistical significance

The random nearest neighbour distance defines the probability that two points could be grouped together on the basis of chance (Levine, 2004). The probability level was set to

$p<0.5$, meaning that approximately 50% of all pairs of points would be found under a random distribution.

Besides the clustering of pairs of points, the routine needs to cluster as many points as possible that falls within the threshold distance with the additional requirement of a user specification of minimum number of points. The probability distribution for this situation is unknown, hence it was necessary to involve a Monte Carlo simulation of randomness for first order clusters to investigate the probability distribution under the conditions of the NNH analysis. The structure of second order clusters are dependent on first order clusters and therefore not simulated. The 95^{th} percentile was considered as a reference for testing statistical significance, meaning that we are willing to accept a one-tailed type l error of only 5% for finding out greater number of clusters than by chance. Table 1 shows the results for D1 while table 2 illustrates the outcome for D2.

D1 identified seven clusters in comparison to the 95^{th} percentile finding two. This means that it is unlikely that all seven would be due to chance. The 95^{th} percentile for D2 was three clusters, and with twelve identified clusters it is highly unlikely the identified clusters being due to chance. Both D1 and D2 can be said to be significant clustered. The routine does not include which of the clusters for D1 and D2 that could have been due to chance. Realistically, one can assume that the clusters with the lowest densities are less certain to be real (Levine, 2004).

The number of points per cluster found in the simulation for the 95^{th} percentile was five for both D1 and D2. This would mean that in terms of number of points per cluster, five clusters could have been due to chance for D1. In other words 29% of the clusters for D1 had more points than what would be expected on the basis of chance distribution. The percentage of clusters for D2 would on the same premise account for 33%.

Density is calculated as the number of points per km^2. The 95^{th} percentile of density is 17.5% per cluster for D1, meaning than six out of seven clusters could be due to chance. For D2, eight out of twelve could be due to chance for the same percentile.

The testing of significance of a cluster analysis is a complex task. The actual distribution could be evaluated according to several criterions. Levine (2004) argues that the number of points identified in the clusters should be the main focus rather than the area or density by themselves since the area has to be defined by a geometrical shape. The number of points would then be the relevant criteria since it is one of the criterions used for the clustering in the NNH algorithm.

Stockholm Municipality (D1)		**Order**	**Cluster**	**Points**	**Area (Km)**	**Density**
Total Clusters First Order	7	1	1	8	0.225	35.5
		1	2	7	0.638	11
		1	3	4	0.486	8.2
		1	4	4	1.429	2.8
		1	5	4	0.181	22
		1	6	4	1.242	3.2
		1	7	3	0.526	5.7
Monte Carlo 95th Percentile Estimation (%)						
Simulated Clusters	2			5	2.687	17.5

Table 2. Analysis' results from the Nearest Neighbour Hierarchical Clustering and Monte Carlo simulation for Stockholm municipality between the years 2010-2017.

Stockholm Municipality with Adjacent Municipalities (D2)		**Order**	**Cluster**	**Points**	**Area (Km)**	**Density**
Total Clusters First Order	12	1	1	14	2.683	5.2
Total Clusters Second Order	1	1	2	10	2.694	3.7
		1	3	10	1.113	9
		1	4	8	3.451	2.3
		1	5	5	6.215	0.8
		1	6	3	2.494	1.2
		1	7	4	0.657	6.1
		1	8	4	0.486	8.2
		1	9	4	2.504	1.6
		1	10	4	3.955	1
		1	11	4	0.876	4.6
		1	12	4	3.525	1.1
		2	1	5 (Clusters)	46.614	0.1
Monte Carlo 95th Percentile Estimation (%)						
Simulated Clusters	3			5	9.61	4.65

Table 3. Analysis' results from the Nearest Neighbour Hierarchical Clustering and Monte Carlo simulation for Stockholm municipality together with adjacent municipalities between the years 2010-2017.

4.4. Regression analysis

The result from the NNH analysis indicates significant clustering of homicide events in several areas in Stockholm between the years 2010-2017. This section of the dissertation aims to address the relationship between homicide rate and deprivation by assessing Poisson based regression. The insight of this relationship might then help to explain the spatial distribution of homicide crimes.

A negative binomial model was applied as a reference to the Poisson model due to indication of some overdispersion in the data and compared with the Poisson model with Akaike information criterion (AIC).

4.4.1. Index efficiency

Table 5 illustrates the use the deprivation index (NDI) as explanatory variable and homicide rate as respondent variable. When using Poisson distribution regression, the result shows a high level of statistically significance. The more restrictive negative binomial model is not showing a statistically significant relationship.

The exponentiated values of the coefficients (Exp(B)) for the Poisson regression is 1.140. In other words: for every integer increase in the standardised NDI, homicide rate would increase with 14%. The Poisson regression model's AIC is lower than the negative binomial model's, indicating that the statistically significant Poisson regression is a better model to use.

Model	**Sig.**	**Exp(B)**	**AIC**
Poisson loglinear	0.000	1.14	58.777
Negative binomial with log link	0.169	1.118	71.12

Table 4. The regression models' efficiencies when measuring NDI as explanatory variable and homicide rate as respondent variable.

Apart from testing statistical significance through regression models, a visual scanning might roughly indicate certain patterns. Figure 1 represents the Poisson regression included districts also referred to as "outer-town". From the figure it can be noticed that the three districts with highest homicide rate also suffers from the highest NDI scores.

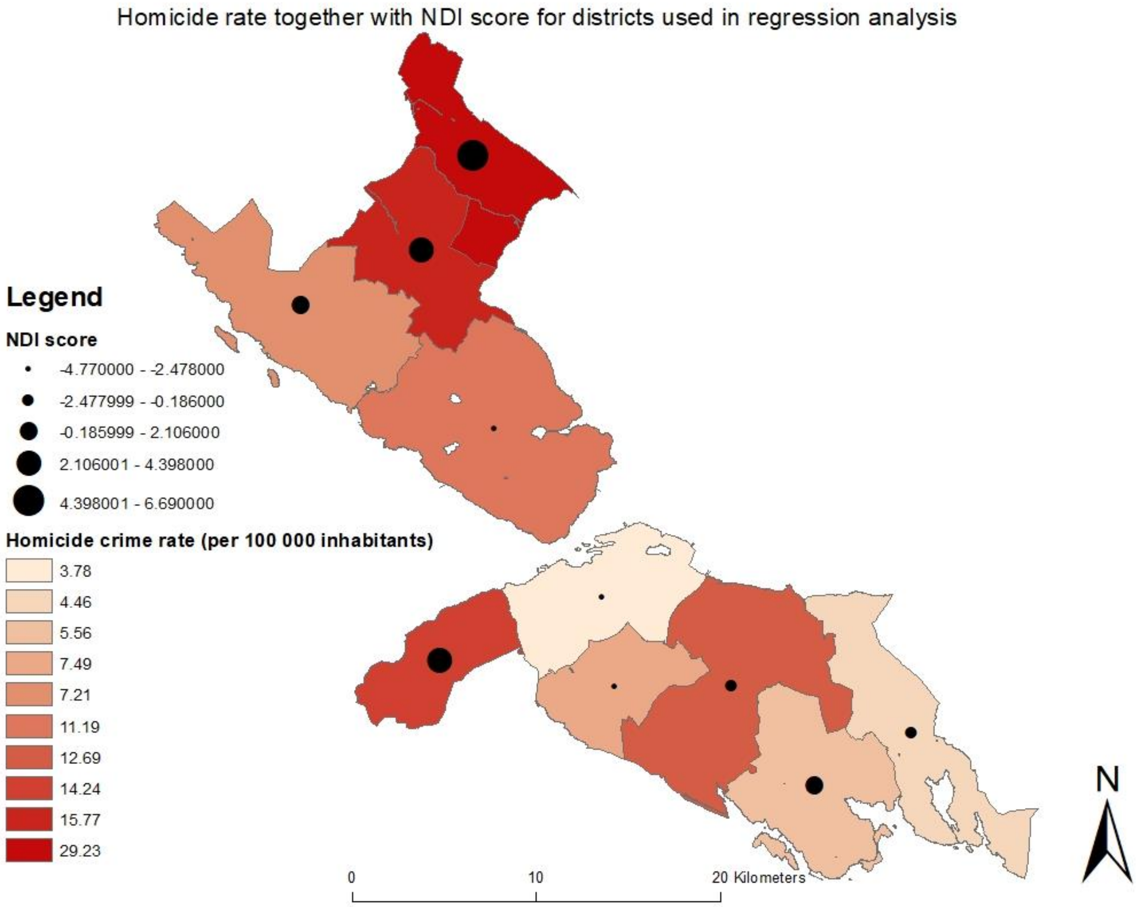

Figure 7. Visual representation of the relationship between districts' homicide rate and NDI-score. The homicide crime rate is shown in a red spectrum from light red (lowest rate) to dark red (highest rate). The NDI score is shown in black circles in five equal breaks. A larger circle represents a higher NDI-score.

4.4.2. Variable performance

The NDI is purely a construct used in order to try to address relative deprivation in Stockholm. To gain a deeper knowledge in the spatial dimension of homicide research it is necessary to investigate which variables that actually performed in the model.

Table 6 illustrates the Poisson model outcome of using the four variables separately as explanatory variables against homicide rate as a dependent variable. Unemployment and social welfare are statistically significant in the Poisson model. The result suggests a strong association between social welfare and homicide rate. None of the four variables show statistical significance in the negative binomial model.

Variable	Sig.	Exp(B)
Education	0.691	1.128
Income	0.318	1.668
Unemployment	0.035	0.262
Social Welfare	0.025	3.199

Table 5. The single variables' performances in the Poisson regression model

4.4.3. Spatial autocorrelation

Homicide rates were tested with Global Moran's I. The results show a non-significant positive clustering (p=0.198). A significant value of either dispersion or overdispersion would have meant that the regression model would have to be integrated with for instance a spatial lag or error model.

Local Moran's I was tested for feature individual homicide characteristics. Only one district (Enskede-Årsta-Vantör) was significant (p=0.05) as a high homicide district being surrounded by low homicide rate districts.

5. Discussion

The results indicate both spatial clustering of homicide crimes and positive correlation between deprivation and homicide rates. The discussion section tries to address the meaning of the results and other methods that were considered in the dissertation. The section continues with suggestion for further research and how utilization of spatial tools can be used for crime prevention.

5.1. Hot spot analysis

The results from D1 and D2 differs greatly in terms of number of hot identified spots. Changing the study area illuminates on how the clustering technique is empirical derivatives of a procedures, where its substance can be argued for. By expanding the study area including the homicide crimes committed in adjacent parishes, hot spots that were not identified in Stockholm municipality in D1 became so in D2.

Hot spots are perceptual constructs. They do not exist in reality and offers room for arbitrariness in its user defined parameters. The perceiving of a meaningful study area and cluster size will affect the outcome, and it is up to the user to interpret the results and decide whether or not the information is substantive. Clustering techniques are powerful tools to use in crime preventing analyses. However, the involvement of arbitrariness will invariably constitute a statistical weakness.

5.1.1. Other methods considered

A few other hot spot techniques were investigated as suitable methods. Spatial and Temporal Analysis of Crime (STAC) is a well-used point-based method. It counts the number of crimes that occur in overlapping circles spread evenly across the study area, where the largest number of crime incidents and the areas with the densest crime activity can be identified. STAC is primarily a portioning method, while NNH is primarily a risk-based technique (Levine, 2004). STAC was rejected as clustering method due to its less hierarchical approach to hot spots.

Kernel estimation is a common used hot spot technique. It is a procedure for smoothing point data to estimate density across a study area. The density at each location reflects the concertation of points in the surrounding area (Levine, 2004). Interpolated risk surfaces generated by Kernel estimation was not used in the project as a result of the discrete nature of homicide crimes and the length of the study interval.

5.2. Regression analysis

The second part of the dissertation investigates the topic further by attempting to assess the relationship between material deprivation and homicide rate for the municipality of Stockholm. A vast majority of homicide research studies are conducted in the United States, and explaining neighboring factors cannot be assumed to parallel those in other nations, even those with similar economic and political structures (Thompson and Gartner, 2014). It is therefore important that Swedish homicide research is not built blindly on assumptions on homicide explaining factors deriving from the United States.

The economic disadvantage was represented by a similar NDI as the one used in a Swedish study by Winkleby et al. (2007) as deprivation indicator. The Poisson regression model shows a positive relationship between the NDI and homicide rate. Assessing the variables separately, unemployment and social welfare showed statistical significance. The relationship between social welfare and homicide rate was shown to be particularly substantial.

Spatial autocorrelation was measured with Moran's I. Scientific studies suggests that homicide crimes can be spatial correlated and need to be accounted for in regression analysis. However, the results of this study did not involve a spatial component in the regression analysis due to non-significance in Global Moran's I for the given dataset. The lack of evidence of spatial effect can mainly be traced to the scarcity of data samples.

5.2.1. Why Poisson should be applied to homicide rates

Statistical research of violent crime rates often suffers from low offense counts, explaining why it has been uncommon to see analyses of homicide for populations less than several hundred thousand. A common solution has been to increase the level of aggregation, for instance using cities instead of districts. If the population sizes of the aggregate units are large relative to the average homicide rate, it would mean that the calculated rates would have been sufficiently fine-grained. The rates could then without harm be thought of as continuous and by that applying ordinary least-squared regression (OLS). However, this puts pressure on the predicting variables as they get coarser and have to be representative for larger areas. Material deprivation rates might vary considerable within a city and might be better analyzed through less aggregation (Osgood, 2000).

Several problems occur when applying OLS to low homicide counts. Homicide crimes are discrete values which cannot be ignored for populations under several hundred thousand. For smaller populations, a single homicide crime might correspond to a high crime rate. Smaller crime counts cannot be assumed to have normal or even symmetrical error distributions of crime rate (Osgood, 2000). The lowest possible crime count is zero, which means that the error distribution need to become increasingly skewed approaching this lower bound. Furthermore, violating the assumption of homogeneity of error variance is inescapable when the precision of the estimated crime rate depends on population size. Larger populations expects a smaller error of prediction for per capita crime rate than smaller populations (Osgood, 2000).

5.2.2. Problems with over dispersion in Poisson

The basic Poisson regression model is appropriate if the assumptions underlying the Poisson distribution are fully met by the data. One assumption is that the residual variance should be equal to the fitted values, which implies that the exploratory variables account for all the meaningful variation among the aggregate units (Osgood, 2000). There is no more reason to expect that a Poisson regression should explain all of the variation in the homicide rates than an OLS regression would explain all variance other than error of measurement.

The variance will also be greater than the mean if the assumption of independence among the individual crime events is inaccurate. Dependence can arise in several ways, for instance if offenders are not being incarcerated or if one homicide event triggered others. Applying the basic Poisson regression model to data that is overdispersed might generate misleading significance tests (Osgood, 2000).

The data indicated overdispersion and for this reason a negative binomial model, which accounts for overdispersion, was used as a reference. By comparing these two regression models it could be established that the Poisson model proved to be a better fit than the negative binomial model for analyzing homicide crimes in Stockholm for the given time interval.

5.2.3. The NDI and other deprivation indexes considered

Much effort was put in to find a suitable deprivation index which would represent districts in Stockholm municipality. The studying of explanatory deprivation variables and their relationship with homicide is central in the field of crime prevention. This study of homicide crimes in Stockholm is a first of its kind. It does not claim that the NDI index used is an optimised deprivation index to be studied together with homicide, but it was the only one available for the study. As with a majority of international deprivation indexes, the variables are likely to overlap each other. What actually defines relative neighbour deprivation needs to be study further in Sweden in order for crime preventing studies to reach higher accuracy.

Several other indexes and single variables were contemplated to use as measures. The measurements fit to Swedish deprivation standards, the homogeneity in data supply and their definition by the districts, and the actual availability of statistics on a district level came ultimately down as the deciding factors.

The Townsend index described by Peter Townsend (1988) incorporates four variables for describing material deprivation within a population; household overcrowding; non-car ownership; non-home ownership; unemployment. This index is applied to a much greater extent in the UK. Several of its variables were not considered a good fit as a Swedish metropolitan deprivation measurement and therefore rejected.

The Carstairs index is an index of deprivation developed for Scotland and used in spatial epidemiology to identify socio-economic confounding. It is similar to the Townsend index but without integrating households in the deprivation measures. It consists of four variables; overcrowding; low social class; lack of car ownership; and male unemployment (Cairstairs and Russel 1991). It was ultimately disregarded due to the same reasons as the Townsend index.

5.3. Future research

Homicide crimes are recognized among the most severe crimes. Homicides complicated nature can be problematic to statistically study on a quantified level, or at least draw meaningful conclusions from. The literature often uses less severe crimes and homicide crimes synonymously and tries to find common denominators as explaining variables. This dissertation has no intention in explaining why certain factors would be contributors to homicide crimes. Instead, its aim is to consider a spatial approach to homicide crimes for Stockholm and apply suitable statistical mathematics to investigate these relationships.

A common feature in the literature is the seeking of thresholds that would explain homicide crimes for specific metropolitan areas. Treating homicide crimes as purely constructs of deprivation variables and with that perception draw conclusions might not always be an appropriate approach. However, I believe that this understanding of homicide crimes and relationships with deprivation variables are crucial to further study in order to gain deeper knowledge about how to predict, and ultimately to prevent, homicide crimes.

Suggestions for further research would be to study homicide crimes on a smaller neighbourhood level, on a longer time period and compare with other Swedish cities. By doing this, national conclusions might be drawn and used as a compliment in governmental decision making and distribution of resources.

6. Conclusions

This dissertation utilizes spatial analysis techniques to investigate the spatial distribution of homicide crimes in Stockholm. Homicide crimes are very severe crimes and therefore argued as discrete in its spatial nature.

The NNH analysis indicates significant homicide clusters in Stockholm municipality. These are mainly concentrated in the city centre and the north western parts. The inclusion in the clustering analysis of adjacent parishes did not change the patterns extensively but allowed for more hot spots to pop up around the existing ones.

The statistical relationship between economic disadvantage and homicide rate were modelled though Poisson based regression models. The result shows a positive correlation between the NDI and homicide rate on a district level. The variables constituting the NDI were investigated for its single efficiencies. The result indicates statistical significance for unemployment and social welfare, with a particular substantial positive correlation between homicide rate and social welfare.

This thesis show that GIS methods can assist researchers to gain a deeper knowledge of the spatial distribution of homicide crimes. It contributes to the homicide research in Stockholm, and show how spatial tools can be utilized in the field. The thesis illuminates the importance of establishing a deeper understanding of the homicide nature in Sweden in order for better homicide crime control. In addition, the thesis' results can help develop the prevention of homicide crimes and assist researchers' further studies in the field.

7. References

Altindag, D. T. 2012. Crime and unemployment: Evidence from Europe. *International Review of Law and Economics*, 32: 145-157.

Block, C. R., and Block, R. L. 1992. Homicide Syndromes and Vulnerability: Violence in Chicago's Community Areas over 25 Years. *Studies on Crime and Crime Prevention* 1: 61–87.

Brottsförebyggande rådet, 2018. Kriminalstatistik 2017. Konstaterade fall av dödligt våld. Retrieved 30 March, 2018, from https://www.bra.se/publikationer/arkiv/publikationer/2018-03-27-konstaterade-fall-av-dodligt-vald.html. (In Swedish).

Brottsförebyggande rådet, 2018. Om Brottsförebyggande rådet. Retrieved 10 April, 2018, from https://www.bra.se/om-bra.html. (In Swedish).

Browning, C. R., Feinberg, S. L., and Dietz, R. D. 2004. The paradox of social organization: networks, collective efficacy, and violent crime in urban neighborhoods. *Social Forces* 83: 503–534.

Bursnik R., and Grasmick H. G. 1993. *Neighborhoods and Crime: The dimensions of Effective Community Control*. New York: Lexington Books.

Busemeyer, J. R., and Diederich, A. 2014. *Estimation and Testing of Computational Psychological Models*. In Neuroeconomics second edition, pp 560. New York: Elsevier.

Carstairs, V. M., and Russel. 1991. Which Deprivation? A Comparison of Selected Deprivation Indexes. *Journal of Public Health Medicine* 13: 318-326.

Ceccato, V. 2008. Expressive Crimes in Post-Socialist States of Estonia, Latvia and Lithuania1. *Journal of Scandinavian Studies in Criminology and Crime Prevention* 9: 2-30.

Cole, J. H., and Gramajo, A. M. 2009. Homicide Rates in a Cross-Section of Countries: Evidence and Interpretations. *Population and developments review* 35: 749-776.

Coleman, J.S. 1986. Social Theory, Social Research, and a Theory of Action. *The American Journal of Sociology* 91: 1309-1355.

Everett, B. 1974. *Cluster Analysis*. Heinemann Educational books, London.

Fu, W. J., Jiang, P. K., Zhou, G. M., Zhao, K. L. 2014. Using Moran's I and GIS to study the spatial pattern of forest litter carbon density in a subtropical region of southeastern China. *Biogeosciences* 11: 2401-2409.

Hannon, L. E. 2005. Extremely Poor Neighborhoods and Homicide. *Social Science Quarterly* 86: 1418–34.

Hipp, J. R. 2007. Income inequality, race, and place: Does the distribution of race and class within neighborhoods affect crime rates? *Criminology* 45: 665–697.

Kawachi, I., Kennedy, B. P., and Wilkinson R. G. 1999. Crime: social disorganization and relative deprivation. *Social Science & Medicine* 48: 719-731.

King, B. F. 1967. Stepwise clustering procedures. *Journal of the American Statistical Association* 62: 86-101.

Kirk, S. D., and Laub, H. J. (2010). Neighborhood Change and Crime in the Modern Metropolis. *Crime and Justice* 39: 441-502.

Krivo, L. J., and Peterson, R. D. 2010. *Divergent Social Worlds: Neighborhood Crime and the Racial-Spatial Divide*. New York: Russel Sage Foundation.

Kubrin, C. E. 2003. Structural Covariates of Homicide Rates: Does Type of Homicide Matter? *Journal of Research in Crime and Delinquency* 40: 139–70.

Lafree, G. 1999. Declining violent crime rates in the 1990s: Predicting crime booms and busts (English). *Le déclin du nombre de crimes liés à la violence au cours des années 1990: Prédire l'explosion et la chute du nombre des crimes* (French). *Annual Review of Sociology* 25: 145-168.

Land, K. C., McCall, P. L., and Cohen L. E. 1990. Structural covariates of homicide rates: Are there any invariances across time and space. *American Journal of Sociology* 96: 1441-1463.

Lee, M., Martinez, R., and Rosenfeld, R. 2001. Does Immigration Increase Homicide? Negative Evidence from Three Border Cities. *The Sociological Quarterly* 42: 559–80.

Levine, A., 2004. Hot Spot Analysis. In: *CrimeStat III*. 6.1-6.55.

Martinez, R., Stowell, J., and Cancino, J., 2008. A Tale of Two Border Cities: Community Context, Ethnicity and Homicide. *Social Science Quarterly 89*: 1–16.

Messner, S. F., and Anselin, L. 2004. *Spatial analyses of homicide with areal data*. In Spatially integrated social science, ed. Janell D. G. 480 pp, Oxford: Oxford University Press.

Messner, S. F., and Zimmerman, G. M. 2012. Community-Level Influences on Crime and Offending. *The Oxford Handbook of Crime Prevention*, 155–71, ed. Welsh, B. C., Farrington, D. P. New York: Oxford University Press.

Messner, S. F., Teske, R. H. C., Baller, R. D., and Thome, H. 2013. Structural Covariates of Violent Crime Rates in Germany: Exploratory Spatial Analyses of Kreise. *Justice Quarterly* 30: 1015-1041.

Nieuwbeerta, P., McCall, P. L., Elffers, H., and Wittebrood, K. 2008. Neighborhood Characteristics and Individual Homicide Risks: Effects of Social Cohesion, Confidence in the Police, and Socioeconomic Disadvantage. *Homicide Studies* 12: 90–116.

Osgood, D. W. 2000. Poisson-based regression analysis of aggregate crime rates. *Journal of Quantitative Criminology* 16: 21–43.

Peterson, R. D., Krivo, L. J., and Harris, M. A. 2000. Disadvantage and Neighborhood Violent Crime: Do Local Institutions Matter? *Journal of Research in Crime and Delinquency* 37: 31–63.

Rogers, M. L., and Pridemore, W. A. 2013. The effect of poverty and social protection on national homicide rates: Direct and moderating effects. *Social Science Research* 42: 584-95.

Rogerson, P. 2001. *Statistical Methods for Geography*. Wiltshire: The Cromwell Press Ltd.

Rourke-McBride, A. 2014. *Utilizing a Data Integration Approach to Analyse the Nature of Homicide Crimes in the Cities of Edinburgh and Glasgow Throughout the Past Century* (Unpublished master's thesis). Scotland: University of Edinburgh.

Sampson, R. J., and Groves, W. B. 1989. Community structure and crime: Testing social-disorganization theory. *American Journal of Sociology* 94: 774–802.

Sampson, R. J., and Raudenbush, S. W., & Earls, F. 1997. Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science* 277: 918–924.

Sampson, R. J., and Wikström, P-OH. 2008. The social order of violence in Chicago and Stockholm neighbourhoods: a comparative inquiry, 97-119, ed. Shapiro I, Kalyvas S, Masoud T, Cambridge: Cambridge University Press.

Shaw, C. R., and McKay, H. D. 1942. *Juvenile delinquency and urban areas*. Chicago: University of Chicago Press.

Stockholm Stad, 2018. Stadsdelsförvaltningar. Retrieved 5 April, 2018, from http://www.stockholm.se/OmStockholm/Forvaltningar-och-bolag/Stadsdelsforvaltningar/.

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.

Townsend, P., Phillimore, P., and Beattie, A. 1988. *Health and deprivation: inequality and the North*. London; New York: Croom Helm.

Uittenbogaard, A., and Ceccato, V. 2012. Space-time Clusters of Crime in Stockholm, Sweden. *Review of European Studies* 4: 148-156.

Wang, F. 2005. Job Access and Homicide Patterns in Chicago: An Analysis at Multiple Geographic Levels Based on Scale-Space Theory. *Journal of Quantitative Criminology* 21: 195-217.

Wang, F., and Arnold, M. T. 2008. Localized income inequality, concentrated disadvantage and homicide. *Applied Geography* 28: 259-270.

Winkleby, M., Sundquist, K., and Cubbin, C. 2007. Inequites in CHD incidence and case fatality by neighborhood deprivation. *American Journal of Preventive Medicine* 32: 97-106.

Wong, S. K. 2012. Youth crime and family disruption in Canadian municipalities: An adaption of Shaw and McKay's social disorganization theory. *International Journal of Law Crime and Justice* 40: 110-114.

Ye, X., and Wu, L. 2011. Analyzing the dynamics of homicide patterns in Chicago: ESDA and spatial panel approaches. *Applied Geography* 31: 800-807.

Zimmerman, G. M., and Messner S. F. 2011. Neighborhood context and nonlinear peer effects on adolescent violent crime. *Criminology* 49: 873–903.

8. Appendix

The appendix shows homicide census data; hot spot analysis specifications; and NDI calculations and statistics details.

8.1. Homicide data

Table 5 represents homicide count per year per district between the years 2010-2017.

District	Year	Homicide count
Rinkeby-Kista	2010	0
	2011	0
	2012	0
	2013	0
	2014	1
	2015	5
	2016	5
	2017	3
		Total: **14**
Spånga-Tensta	2010	0
	2011	0
	2012	0
	2013	0
	2014	1
	2015	1
	2016	2
	2017	2
		Total: **6**
Hässelby-Vällingby	2010	0
	2011	1
	2012	1
	2013	0

District	Year	Homicide count
	2014	1
	2015	2
	2016	0
	2017	0
		Total: **5**
Bromma	2010	1
	2011	0
	2012	0
	2013	0
	2014	2
	2015	1
	2016	1
	2017	3
		Total: **8**
Enskede-Årsta-Vantör	2010	0
	2011	1
	2012	1
	2013	3
	2014	3
	2015	1
	2016	0
	2017	3
		Total: **12**
Skarpnäck	2010	1
	2011	0
	2012	0
	2013	0
	2014	0

District	Year	Homicide count
	2015	0
	2016	0
	2017	1
		Total: **2**
Farsta	2010	0
	2011	0
	2012	0
	2013	1
	2014	0
	2015	0
	2016	2
	2017	0
		Total: **3**
Älvsjö	2010	0
	2011	1
	2012	0
	2013	0
	2014	1
	2015	0
	2016	0
	2017	0
		Total: **2**
Hägersten-Liljeholmen	2010	0
	2011	0
	2012	0
	2013	0
	2014	2
	2015	0

District	Year	Homicide count
	2016	0
	2017	1
		Total: **3**
Skärholmen	2010	0
	2011	1
	2012	1
	2013	0
	2014	1
	2015	0
	2016	1
	2017	1
		Total: **5**
Norrmalm	2010	0
	2011	3
	2012	0
	2013	1
	2014	2
	2015	0
	2016	0
	2017	5
		Total: **11**
Östermalm	2010	0
	2011	0
	2012	1
	2013	0
	2014	1
	2015	0
	2016	0

District	Year	Homicide count
	2017	0
		Total: **2**
Södermalm	2010	1
	2011	2
	2012	0
	2013	2
	2014	1
	2015	0
	2016	2
	2017	1
		Total: **9**
Kungsholmen	2010	2
	2011	0
	2012	1
	2013	1
	2014	1
	2015	0
	2016	0
	2017	0
		Total: **5**
Huddinge	2010	3
	2011	2
	2012	1
	2013	2
	2014	2
	2015	1
	2016	0
	2017	2

District	Year	Homicide count
		Total: **13**
Ekerö	2010	0
	2011	0
	2012	0
	2013	0
	2014	0
	2015	0
	2016	0
	2017	0
		Total: **0**
Tyresö	2010	0
	2011	0
	2012	1
	2013	0
	2014	0
	2015	0
	2016	0
	2017	0
		Total: **1**
Nacka	2010	0
	2011	0
	2012	0
	2013	0
	2014	0
	2015	0
	2016	1
	2017	2
		Total: **3**

District	Year	Homicide count
Lidingö	2010	0
	2011	0
	2012	0
	2013	0
	2014	1
	2015	1
	2016	0
	2017	0
		Total: **2**
Danderyd	2010	0
	2011	0
	2012	0
	2013	0
	2014	0
	2015	0
	2016	0
	2017	1
		Total: **1**
Sollentuna	2010	2
	2011	0
	2012	1
	2013	0
	2014	1
	2015	0
	2016	2
	2017	0
		Total: **6**
Sundbyberg	2010	1

District	Year	Homicide count
	2011	0
	2012	3
	2013	0
	2014	0
	2015	1
	2016	1
	2017	5
		Total: **11**
Järfälla	2010	3
	2011	0
	2012	1
	2013	0
	2014	0
	2015	1
	2016	1
	2017	3
		Total: **9**
Solna	2010	1
	2011	1
	2012	0
	2013	2
	2014	0
	2015	1
	2016	0
	2017	1
		Total: **6**

Table 6. Homicide data of all districts in Stockholm municipality and the surrounding parishes.

8.2. Hot spot analysis

CrimeStat is a spatial statistics program that can analyse incident location data. It can interface with a majority of desktop GIS and provides a variety of tools for spatial analysis of crime incidents or other point locations. The program was evolved as a complement to GIS since it offers statistical methods suited for more quantitative approaches, and specializes in the analysis of point locations (Levine 2004). The program was chosen for this dissertation due to its hot spotting abilities for point data, which other GIS-software was considered ill-suited for.

Defining the data specifications correctly in CrimeStat is crucial in order to obtain a meaningful analysis. Table 7 defines the data setup including the parameters used.

A primary file is required for any analysis, and needs to be a point file with X and Y coordinates. Homicide crimes were used as the primary file and created using a shapefile in ArcGIS with coordinates for every point.

A reference file that covers the study area must be defined. Shapefiles defining districts' borders created with geoprocessing tools in ArcGIS were used as reference files.

The measurement parameter requires the user to specify area and length of street network of the reference file. It is used as a reference for the Monte Carlo simulation of randomness to test statistical significance.

Study Area	**Primary File**	**Reference file**	**Measurement parameters**	**NNH parameters**
Stockholm Municipality	Coordinate System: SWEREF99 TM	External Polygon file covering area	Area in sq. meters: 216776337	Type of analysis: Nearest Neighbor Hierachical Spatial Clustering
			Length of Street Network in meters: 114458	Type of Search Radius: Random NN distance
			Type of Measurement: Indirect (Manhattan)	Search radius: P>.5
				Minimum Points per Cluster: 3
				Number of Standard Deviations for the ellipses: 1.5
				Simulation runs: 1000
Stockholm Municipality with Adjacent Parishes	Coordinate System: SWEREF99 TM	External Polygon file covering area	Area in sq. meters: 1206910000	Type of analysis: Nearest Neighbor Hierachical Spatial Clustering
			Length of Street Network in meters: 203268	Type of Search Radius: Random NN distance
			Type of Measurement: Indirect (Manhattan)	Search radius: P>.5
				Minimum Points per Cluster: 3
				Number of Standard Deviations for the ellipses: 1.5
				Simulation runs: 1000

Table 7. Hot spot analysis' details used in CrimeStat.

8.3. Regression appendix

The census data for each deprivation variable per year was obtained from the municipalities' official statistics database. Mean values for every deprivation variable were derived for the time interval 2010-2017, and used together with the homicide crime count product for the same interval. Homicide crime rate had to be rounded to nearest integer to work with Poisson regression.

The variables' z-score were summed together to create the NDI representing the districts' mean values for the time period. Table 8 illustrates the NDI and homicide crime rate used in the regression analysis. Table 9 – 18 shows the raw census data.

Homicide Crime rate	Sum (NDI)	Receiving Social Welfare (%)	Unemployment (%)	Low Income (0 - 159.8 tkr) (%)	People without Tertiary Education (%)	
29.23	6.69	1.96	1.68	1.83	1.22	**Rinkeby-Kista**
15.77	3.99	1.18	1.03	0.91	0.87	**Spånga-Tensta**
7.21	0.57	0	0.24	-0.1	0.43	**Hässelby-Vällingby**
11.19	-4.77	-1.02	-1.28	-0.967	-1.5	**Bromma**
12.69	-0.8	-0.21	-0.21	-0.31	-0.07	**Enskede-Årsta-Vantör**
4.46	-2.01	-0.45	-0.35	-0.37	-0.84	**Skarpnäck**
5.56	0	-0.26	0.06	-0.26	0.46	**Farsta**
7.49	-3.7	-0.93	-1.25	-0.99	-0.53	**Älvsjö**
3.78	-4.15	-1	-0.88	-1	-1.27	**Hägersten-Liljeholmen**
14.24	4.13	0.73	0.95	1.25	1.2	**Skärholmen**

Table 8. The districts' variables' z-scores and homicide rate (per 100 000 inhabitants).

Rinkeby-Kista	Population	Tertiary Education (%)	Low income (%)	Unemploymet (%)	Receiving social welfare (%)
2010	45691	71	54.6	22.1	11.6
2011	46792	70.2	53	22.9	11.1
2012	47872	69.8	51.7	24.2	10.2
2013	48366	69.2	51.2	24.5	9.8
2014	48828	69.2	49.4	23.5	9.2
2015	48604	69	47.8	19.4	8.5
2016	49273	68.8	46.1	20.4	7.7
2017	no data	no data	no data	22.1	7.3
Average	47918	69.6	50.54	22.39	10.65

Table 9. Raw census data of the explanatory variables for Rinkeby-Kista.

Spånga-Tensta	Population	Tertiary Education (%)	Low income (%)	Unemployment (%)	Receiving social welfare (%)
2010	37336	68.4	45.6	19.8	10.4
2011	37806	67.7	44.8	19.6	9.9
2012	37918	67	43.5	19.7	9.5
2013	38512	66.1	42.6	21.8	8.9
2014	38591	65.8	42.2	21	8.1
2015	37868	65.4	40.2	17.3	7.3
2016	38236	64.8	39	17.4	6.2
2017	no data	no data	no data	17.5	5.7
Average	38038	66.46	42.56	19.26	8.25

Table 10. Raw census data of the explanatory variables for Spånga-Tensta.

Hässelby-Vällingby	**Population**	**Tertiary Education (%)**	**Low income (%)**	**Unemployment (%)**	**Receiving social welfare (%)**
2010	65083	64.8	36.4	16.9	5.4
2011	66572	63.8	35.6	15.9	5.5
2012	67904	63.1	34.4	19.5	5.3
2013	69251	62.2	33.5	15.3	5.3
2014	70819	61.8	33.4	15.3	5
2015	72561	61.1	32.2	15.6	4.2
2016	73445	60.8	30.6	12.8	3.5
2017	no data	no data	no data	12.7	3.1
Average	69376	62.51	33.73	15.5	4.66

Table 11. Raw census data of the explanatory variables for Hässelby-Vällingby.

Bromma	**Population**	**Tertiary Education (%)**	**Low income (%)**	**Unemployment (%)**	**Receiving social welfare (%)**
2010	65519	46.9	29.5	9.2	1.9
2011	66976	45.9	28.1	8.7	1.8
2012	69492	45.5	26.9	9.7	1.7
2013	71291	44.7	25.9	9.4	1.6
2014	73974	44.1	25.3	8	1.5
2015	76068	43.4	24.4	7.8	1.3
2016	77295	43	23.5	6.9	1.2
2017	no data	no data	no data	6.2	1.2
Average	71516	44.79	26.23	8.24	1.53

Table 12. Raw census data of the explanatory variables for Bromma.

Enskede-Årsta-Vantör	Population	Tertiary Education (%)	Low income (%)	Unemployment (%)	Receiving social welfare (%)
2010	90370	60.6	35.4	17.1	5.4
2011	91827	59.6	34.5	14.4	4.9
2012	93245	58.7	33	14.4	4.1
2013	94503	58	31.3	13.7	3.9
2014	96470	56.9	31.1	12.9	3.8
2015	97587	55.9	29.8	12.5	3.7
2016	97993	55.2	28.3	10.7	3.3
2017	no data	no data	no data	11.2	3
Average	94571	57.84	31.91	13.36	4.01

Table 13. Raw census data of the explanatory variables for Enskede-Årsta-Vantör.

Skarpnäck	Population	Tertiary Education (%)	Low income (%)	Unemployment (%)	Receiving social welfare (%)
2010	43419	54.8	35.3	15.9	4.4
2011	43844	53.6	34.1	13.7	3.9
2012	44456	52.1	32.6	12.8	3.5
2013	44856	50.5	31.1	14.6	3.3
2014	45150	49.4	30.4	12.4	3.2
2015	45716	48.1	29	12	3
2016	46145	47.2	27.1	10.3	2.5
2017	no data	no data	no data	9.8	2.4
Average	44798	50.81	31.37	12.69	3.28

Table 14. Raw census data of the explanatory variables for Skarpnäck.

Farsta	**Populati-on**	**Tertiary Education (%)**	**Low income (%)**	**Unemployment (%)**	**Receiving social welfare (%)**
2010	50601	65.8	36	16.5	5.1
2011	51852	64.9	35	16.7	4.8
2012	53063	64	33.3	15.6	4.5
2013	54401	62.8	31.7	15.2	3.9
2014	55058	61.7	31.3	13.8	3.6
2015	55693	60.4	30.2	13	3.3
2016	56481	59.4	28.6	11.8	2.9
2017	no data	no data	no data	14.7	2.8
Average	53 878	62.71	32.3	14.66	3.86

Table 15. Raw census data of the explanatory variables for Farsta.

Älvsjö	**Population**	**Tertiary Education (%)**	**Low income (%)**	**Unemployment (%)**	**Receiving social welfare (%)**
2010	25056	57.3	29	12.3	2.3
2011	25594	56.2	27.8	8.2	2.2
2012	26250	55.1	26.7	11.3	2.1
2013	26935	53.9	25.4	10.1	1.7
2014	27159	52.7	25.1	9.5	1.6
2015	27710	51.8	24	8.4	1.6
2016	28141	50.9	23.5	7.1	1.5
2017	no data	no data	no data	10.6	1.4
Average	26692	53.99	25.93	8.36	1.8

Table 16. Raw census data of the explanatory variables for Älvsjö.

Hägersten-Liljeholmen	**Population**	**Tertiary Education (%)**	**Low income (%)**	**Unemployment (%)**	**Receiving social welfare (%)**
2010	73957	50.6	29.6	13.2	2
2011	76598	49.1	28.2	10.9	1.9
2012	78691	47.8	26.6	10.8	1.7
2013	80705	46.4	25.3	10.3	1.5
2014	83100	45.7	24.7	9.8	1.5
2015	84914	44.7	23.8	8.2	1.5
2016	87026	44	22.5	7.6	1.3
2017	no data	no data	no data	10.3	1.3
Average	80713	46.9	25.81	10.14	1.59

Table 17. Raw census data of the explanatory variables for Hägersten-Liljeholmen.

Skärholmen	**Population**	**Tertiary Education (%)**	**Low income (%)**	**Unemployment (%)**	**Receiving social welfare (%)**
2010	33565	71.8	48.7	21.8	8.6
2011	34317	70.9	48	20.2	8.5
2012	34746	70.3	46.8	19.1	7.8
2013	35283	69.3	45.7	21	6.8
2014	35585	68.8	44.5	18.6	6.5
2015	35863	68.2	43.4	18.4	6.2
2016	36378	67.4	41.9	15	5.7
2017	no data	no data	no data	17.2	5.1
Average	35105	69.53	45.57	18.91	6.9

Table 18. Raw census data of the explanatory variables for Skärholmen.

8.4. Specifications for explanatory variables

The NDI consisted of four single deprivation variables; income; unemployment; tertiary education; and social welfare. The variables' details are explained in this section.

8.4.1. Income (low income)

Income was defined as total earned income and is based on reports from the Swedish National Tax Board (Skatteverket). Total earned income consisted of income from service and income from business. Income from service included except income from salaries, also income from pension, sickness benefit, and other taxable benefits from the Swedish Social Insurance Agency (Försäkringskassan).

The available statistics divided the population in five income classes including people from 20 years and up (Thousand Swedish crowns); 0 tkr; 0.1 – 159.9 tkr; 160 – 319.9 tkr; 320 – 499.9 tkr; 500 tkr -. The product of the share between the two lowest income classes were chosen as representative for the income variable (0 – 159.9 tkr). The values for year 2017 was yet not obtainable and therefore not included in the calculation of the mean value.

8.4.2. Unemployment

Unemployment was reported according to the definition of The Swedish Public Employment Service (Arbetsförmedlingen) of openly unemployed i.e. job-seekers without jobs which actively were in the search for jobs and immediately could begin to work. Unemployment rates for the districts were available as four age groups; 18 – 19; 20 – 24; 25 – 54; 55 – 64. Due to the available statistics, the age groups' rates were summed together to create a total.

8.4.3. Education (people without tertiary education)

The statistics for the educational level in the districts were available in three categories between the ages 16-74; elementary school; high school; and college. The percentage of the population with at least 120 credits on university level were qualify into the college category, and used as the measure of education for the districts. To fit the regression model, the values were reversed to specify people without tertiary education. The values for year 2017 was yet not obtainable and therefore not included in the calculation of the mean value.

8.4.4. Social welfare (receiving social welfare)

This variable refers to anyone which is included in a household and has received social welfare under the year. Homeless people are not included in the districts' statistics. The statistics were not official by the time and had to be requested by the municipality.

Series from Lund University

Department of Physical Geography and Ecosystem Science

Master Thesis in Geographical Information Science

1. *Anthony Lawther:* The application of GIS-based binary logistic regression for slope failure susceptibility mapping in the Western Grampian Mountains, Scotland (2008).
2. *Rickard Hansen:* Daily mobility in Grenoble Metropolitan Region, France. Applied GIS methods in time geographical research (2008).
3. *Emil Bayramov:* Environmental monitoring of bio-restoration activities using GIS and Remote Sensing (2009).
4. *Rafael Villarreal Pacheco:* Applications of Geographic Information Systems as an analytical and visualization tool for mass real estate valuation: a case study of Fontibon District, Bogota, Columbia (2009).
5. *Siri Oestreich Waage:* a case study of route solving for oversized transport: The use of GIS functionalities in transport of transformers, as part of maintaining a reliable power infrastructure (2010).
6. *Edgar Pimiento:* Shallow landslide susceptibility – Modelling and validation (2010).
7. *Martina Schäfer:* Near real-time mapping of floodwater mosquito breeding sites using aerial photographs (2010).
8. *August Pieter van Waarden-Nagel:* Land use evaluation to assess the outcome of the programme of rehabilitation measures for the river Rhine in the Netherlands (2010).
9. *Samira Muhammad:* Development and implementation of air quality data mart for Ontario, Canada: A case study of air quality in Ontario using OLAP tool. (2010).
10. *Fredros Oketch Okumu*: Using remotely sensed data to explore spatial and temporal relationships between photosynthetic productivity of vegetation and malaria transmission intensities in selected parts of Africa (2011).
11. *Svajunas Plunge:* Advanced decision support methods for solving diffuse water pollution problems (2011).
12. *Jonathan Higgins:* Monitoring urban growth in greater Lagos: A case study using GIS to monitor the urban growth of Lagos 1990 - 2008 and produce future growth prospects for the city (2011).
13. *Mårten Karlberg:* Mobile Map Client API: Design and Implementation for Android (2011).
14. *Jeanette McBride:* Mapping Chicago area urban tree canopy using color infrared imagery (2011).
15. *Andrew Farina:* Exploring the relationship between land surface temperature and vegetation abundance for urban heat island mitigation in Seville, Spain (2011).
16. *David Kanyari*: Nairobi City Journey Planner: An online and a Mobile Application (2011).
17. *Laura V. Drews:* Multi-criteria GIS analysis for siting of small wind power plants - A case study from Berlin (2012).
18. *Qaisar Nadeem:* Best living neighborhood in the city - A GIS based multi criteria evaluation of ArRiyadh City (2012).

19. *Ahmed Mohamed El Saeid Mustafa:* Development of a photo voltaic building rooftop integration analysis tool for GIS for Dokki District, Cairo, Egypt (2012).
20. *Daniel Patrick Taylor*: Eastern Oyster Aquaculture: Estuarine Remediation via Site Suitability and Spatially Explicit Carrying Capacity Modeling in Virginia's Chesapeake Bay (2013).
21. *Angeleta Oveta Wilson:* A Participatory GIS approach to *unearthing* Manchester's Cultural Heritage '*gold mine'* (2013).
22. *Ola Svensson:* Visibility and Tholos Tombs in the Messenian Landscape: A Comparative Case Study of the Pylian Hinterlands and the Soulima Valley (2013).
23. *Monika Ogden:* Land use impact on water quality in two river systems in South Africa (2013).
24. *Stefan Rova:* A GIS based approach assessing phosphorus load impact on Lake Flaten in Salem, Sweden (2013).
25. *Yann Buhot:* Analysis of the history of landscape changes over a period of 200 years. How can we predict past landscape pattern scenario and the impact on habitat diversity? (2013).
26. *Christina Fotiou:* Evaluating habitat suitability and spectral heterogeneity models to predict weed species presence (2014).
27. *Inese Linuza:* Accuracy Assessment in Glacier Change Analysis (2014).
28. *Agnieszka Griffin:* Domestic energy consumption and social living standards: a GIS analysis within the Greater London Authority area (2014).
29. *Brynja Guðmundsdóttir:* Detection of potential arable land with remote sensing and GIS - A Case Study for Kjósarhreppur (2014).
30. *Oleksandr Nekrasov:* Processing of MODIS Vegetation Indices for analysis of agricultural droughts in the southern Ukraine between the years 2000-2012 (2014).
31. *Sarah Tressel:* Recommendations for a polar Earth science portal in the context of Arctic Spatial Data Infrastructure (2014).
32. *Caroline Gevaert:* Combining Hyperspectral UAV and Multispectral Formosat-2 Imagery for Precision Agriculture Applications (2014).
33. *Salem Jamal-Uddeen:* Using GeoTools to implement the multi-criteria evaluation analysis - weighted linear combination model (2014).
34. *Samanah Seyedi-Shandiz:* Schematic representation of geographical railway network at the Swedish Transport Administration (2014).
35. *Kazi Masel Ullah:* Urban Land-use planning using Geographical Information System and analytical hierarchy process: case study Dhaka City (2014).
36. *Alexia Chang-Wailing Spitteler:* Development of a web application based on MCDA and GIS for the decision support of river and floodplain rehabilitation projects (2014).
37. *Alessandro De Martino:* Geographic accessibility analysis and evaluation of potential changes to the public transportation system in the City of Milan (2014).
38. *Alireza Mollasalehi:* GIS Based Modelling for Fuel Reduction Using Controlled Burn in Australia. Case Study: Logan City, QLD (2015).
39. *Negin A. Sanati:* Chronic Kidney Disease Mortality in Costa Rica; Geographical Distribution, Spatial Analysis and Non-traditional Risk Factors (2015).

40. *Karen McIntyre:* Benthic mapping of the Bluefields Bay fish sanctuary, Jamaica (2015).
41. *Kees van Duijvendijk:* Feasibility of a low-cost weather sensor network for agricultural purposes: A preliminary assessment (2015).
42. *Sebastian Andersson Hylander:* Evaluation of cultural ecosystem services using GIS (2015).
43. *Deborah Bowyer:* Measuring Urban Growth, Urban Form and Accessibility as Indicators of Urban Sprawl in Hamilton, New Zealand (2015).
44. *Stefan Arvidsson:* Relationship between tree species composition and phenology extracted from satellite data in Swedish forests (2015).
45. *Damián Giménez Cruz*: GIS-based optimal localisation of beekeeping in rural Kenya (2016).
46. *Alejandra Narváez Vallejo:* Can the introduction of the topographic indices in LPJ-GUESS improve the spatial representation of environmental variables? (2016).
47. *Anna Lundgren:* Development of a method for mapping the highest coastline in Sweden using breaklines extracted from high resolution digital elevation models (2016).
48. *Oluwatomi Esther Adejoro:* Does location also matter? A spatial analysis of social achievements of young South Australians (2016).
49. *Hristo Dobrev Tomov:* Automated temporal NDVI analysis over the Middle East for the period 1982 - 2010 (2016).
50. *Vincent Muller:* Impact of Security Context on Mobile Clinic Activities A GIS Multi Criteria Evaluation based on an MSF Humanitarian Mission in Cameroon (2016).
51. *Gezahagn Negash Seboka:* Spatial Assessment of NDVI as an Indicator of Desertification in Ethiopia using Remote Sensing and GIS (2016).
52. *Holly Buhler:* Evaluation of Interfacility Medical Transport Journey Times in Southeastern British Columbia. (2016).
53. *Lars Ole Grottenberg*: Assessing the ability to share spatial data between emergency management organisations in the High North (2016).
54. *Sean Grant:* The Right Tree in the Right Place: Using GIS to Maximize the Net Benefits from Urban Forests (2016).
55. *Irshad Jamal:* Multi-Criteria GIS Analysis for School Site Selection in Gorno-Badakhshan Autonomous Oblast, Tajikistan (2016).
56. *Fulgencio Sanmartín:* Wisdom-volkano: A novel tool based on open GIS and time-series visualization to analyse and share volcanic data (2016).
57. *Nezha Acil:* Remote sensing-based monitoring of snow cover dynamics and its influence on vegetation growth in the Middle Atlas Mountains (2016).
58. *Julia Hjalmarsson:* A Weighty Issue: Estimation of Fire Size with Geographically Weighted Logistic Regression (2016).
59. *Mathewos Tamiru Amato:* Using multi-criteria evaluation and GIS for chronic food and nutrition insecurity indicators analysis in Ethiopia (2016).
60. *Karim Alaa El Din Mohamed Soliman El Attar:* Bicycling Suitability in Downtown, Cairo, Egypt (2016).
61. *Gilbert Akol Echelai:* Asset Management: Integrating GIS as a Decision Support Tool in Meter Management in National Water and Sewerage Corporation (2016).
62. *Terje Slinning:* Analytic comparison of multibeam echo soundings (2016).

63. *Gréta Hlín Sveinsdóttir:* GIS-based MCDA for decision support: A framework for wind farm siting in Iceland (2017).
64. *Jonas Sjögren:* Consequences of a flood in Kristianstad, Sweden: A GIS-based analysis of impacts on important societal functions (2017).
65. *Nadine Raska:* 3D geologic subsurface modelling within the Mackenzie Plain, Northwest Territories, Canada (2017).
66. *Panagiotis Symeonidis*: Study of spatial and temporal variation of atmospheric optical parameters and their relation with PM 2.5 concentration over Europe using GIS technologies (2017).
67. *Michaela Bobeck:* A GIS-based Multi-Criteria Decision Analysis of Wind Farm Site Suitability in New South Wales, Australia, from a Sustainable Development Perspective (2017).
68. *Raghdaa Eissa*: Developing a GIS Model for the Assessment of Outdoor Recreational Facilities in New Cities Case Study: Tenth of Ramadan City, Egypt (2017).
69. *Zahra Khais Shahid*: Biofuel plantations and isoprene emissions in Svea and Götaland (2017).
70. *Mirza Amir Liaquat Baig*: Using geographical information systems in epidemiology: Mapping and analyzing occurrence of diarrhea in urban - residential area of Islamabad, Pakistan (2017).
71. *Joakim Jörwall*: Quantitative model of Present and Future well-being in the EU-28: A spatial Multi-Criteria Evaluation of socioeconomic and climatic comfort factors (2017).
72. *Elin Haettner*: Energy Poverty in the Dublin Region: Modelling Geographies of Risk (2017).
73. *Harry Eriksson*: Geochemistry of stream plants and its statistical relations to soil- and bedrock geology, slope directions and till geochemistry. A GIS-analysis of small catchments in northern Sweden (2017).
74. *Daniel Gardevärn:* PPGIS and Public meetings – An evaluation of public participation methods for urban planning (2017).
75. *Kim Friberg:* Sensitivity Analysis and Calibration of Multi Energy Balance Land Surface Model Parameters (2017).
76. *Viktor Svanerud:* Taking the bus to the park? A study of accessibility to green areas in Gothenburg through different modes of transport (2017).
77. *Lisa-Gaye Greene*: Deadly Designs: The Impact of Road Design on Road Crash Patterns along Jamaica's North Coast Highway (2017).
78. *Katarina Jemec Parker*: Spatial and temporal analysis of fecal indicator bacteria concentrations in beach water in San Diego, California (2017).
79. *Angela Kabiru*: An Exploratory Study of Middle Stone Age and Later Stone Age Site Locations in Kenya's Central Rift Valley Using Landscape Analysis: A GIS Approach (2017).
80. *Kristean Björkmann*: Subjective Well-Being and Environment: A GIS-Based Analysis (2018).
81. *Williams Erhunmonmen Ojo*: Measuring spatial accessibility to healthcare for people living with HIV-AIDS in southern Nigeria (2018).
82. *Daniel Assefa*: Developing Data Extraction and Dynamic Data Visualization (Styling) Modules for Web GIS Risk Assessment System (WGRAS). (2018).
83. *Adela Nistora*: Inundation scenarios in a changing climate: assessing potential impacts of sea-level rise on the coast of South-East England (2018).

84. *Marc Seliger*: Thirsty landscapes - Investigating growing irrigation water consumption and potential conservation measures within Utah's largest master-planned community: Daybreak (2018).
85. *Luka Jovičić*: Spatial Data Harmonisation in Regional Context in Accordance with INSPIRE Implementing Rules (2018).
86. *Christina Kourdounouli*: Analysis of Urban Ecosystem Condition Indicators for the Large Urban Zones and City Cores in EU (2018).
87. *Jeremy Azzopardi*: Effect of distance measures and feature representations on distance-based accessibility measures (2018).
88. *Patrick Kabatha*: An open source web GIS tool for analysis and visualization of elephant GPS telemetry data, alongside environmental and anthropogenic variables (2018).
89. *Richard Alphonce Giliba*: Effects of Climate Change on Potential Geographical Distribution of Prunus africana (African cherry) in the Eastern Arc Mountain Forests of Tanzania (2018).
90. *Eiður Kristinn Eiðsson*: Transformation and linking of authoritative multi-scale geodata for the Semantic Web: A case study of Swedish national building data sets (2018).
91. *Niamh Harty*: HOP!: a PGIS and citizen science approach to monitoring the condition of upland paths (2018).
92. *José Estuardo Jara Alvear*: Solar photovoltaic potential to complement hydropower in Ecuador: A GIS-based framework of analysis (2018).
93. *Brendan O'Neill*: Multicriteria Site Suitability for Algal Biofuel Production Facilities (2018).
94. *Roman Spataru*: Spatial-temporal GIS analysis in public health – a case study of polio disease (2018).
95. *Alicja Miodońska*: Assessing evolution of ice caps in Suðurland, Iceland, in years 1986 - 2014, using multispectral satellite imagery (2019).
96. *Dennis Lindell Schettini*: A Spatial Analysis of Homicide Crime's Distribution and Association with Deprivation in Stockholm Between 2010-2017 (2019).