Masters
Program
in **Geospatial Technologies**

Street luminosity influence on reported thefts from vehicles during night-time

José Mário Roberto Ventura

Dissertation submitted in partial fulfilment of the requirements
for the Degree of *Master of Science in Geospatial Technologies*

NOVA IMS Information Management School
Universitat Jaume I
Westfälische Wilhelms-Universität Münster

Street luminosity influence on reported thefts from vehicles during night-time

Dissertation supervised by

Professor Ana Cristina Costa, Ph.D.

Information Management School

Universidade Nova de Lisboa

Lisbon, Portugal

Dissertation co-supervised by

Professor Michael Gould, Ph.D.

Institute of New Imaging Technologies

Universitat Jaume I

Castellón de la Plana, Spain

Mario Pesch, M.Sc.

Institute for Geoinformatics

Westfälische Wilhelms - Universität Münster

Münster, Germany

February 2022

DECLARATION OF ORIGINALITY

I declare that the work described in this document is my own and not from someone else. All the assistance I have received from other people is duly acknowledged and all the sources (published or not published) are referenced.

This work has not been previously evaluated or submitted to NOVA Information Management School or elsewhere.

LISBON, FEBRUARY 22, 2022

JOSÉ MÁRIO ROBERTO VENTURA

ACKNOWLEDGMENTS

First, I would like to acknowledge all the Professors, instructors, researchers, administrative and operational staff at Nova Lisbon´s Information Management School and at Münster´s Institute for Geoinformatics for conducting great courses that one way or another contributed to this dissertation, and for . Thank you to my colleagues for some nice discussions during the initial phases of this work.

My most sincere gratitude to Professor Ana Cristina Costa for being my advisor, always guiding in an objective and considerate approach. Additionally, thank you for making the effort to accommodate my eight-hour time difference!

Thank you to Professor Michael Gould and to Mario Pesch for agreeing to be co-advisors to my thesis and for their remarks. Mario thank you for your kind help and assistance throughout the process of getting parts, assembling and coding the senseBox.

For always supporting and encouraging in all phases of my life a big thank you and love to my mother, nothing could ever be possible without you. To my aunt Eva and uncle Vasco a big thank you for the help and care throughout the course.

For putting up to my sometimes-crazy hours, extended work schedule, and specially for their love and support, thank you to my wife Joana and daughter Alice. I certainly wouldn´t have done without you. You are the loves of my life.

Street luminosity influence on reported thefts from vehicles during night-time

ABSTRACT

Criminality across many urban settings has created the need to develop tools that help better understand the social and physical determinants of crime. One specific aspect is how certain urban characteristics may influence criminal activities. One facet of the built environment, street lighting, influences the perception of safety for a potential victim, and it also influences a perpetrator´s risk analysis, affecting how it envisions both costs and rewards for committing a crime.

The study looked into the relationship between street illuminance levels, through street light pole density as a proxy, and other crime determinants and the prevalence of reported Night-Time Theft from Vehicle crimes in Vancouver, BC, Canada, through exploratory spatial data analysis and by fitting Geographically Weighted Poisson Regressions. To test if street lighting pole density is a usable proxy for street illuminance it also modeled the relationship between street lighting pole and tree densities and measured night time street illuminance by fitting an Ordinary Least Squares regression. Night time street illuminance was measured using a specially built georeferenced mobile illuminance collection station based on the senseBox.

Findings suggest that while a citywide effect is evident for some of the explanatory variables, there is an evident nonstationary relation between the explanatory variables and Night-Time Theft from Vehicle crimes in Vancouver. Regarding street lighting, regressions suggest it may not be an important covariate with Night-Time Theft from Vehicle crime. Coefficients are quite heterogenous throughout with most of the study area showing a mix of weak to mild positive association, specially on the East side, and weak to mild negative associations. The OLS regression showed a moderately weak relation between light poles and tree densities to collected street illuminance. The question of street lighting pole density being a usable proxy for street illuminance could not be answered with confidence.

KEYWORDS

Spatial analysis

Crime analysis

Crime patterns

Theft from vehicle crime

Street illuminance

Street light poles

Geographically Weighted Poisson Regression

Ordinary Least Squares Regression

Vancouver

SenseBox

ACRONYMS

AIC – Akaike Information Criterion

AICc – Corrected Akaike Information Criterion

CDAs – Census Dissemination Areas

ESDA – Exploratory Spatial Data Analysis

GIS – Geographic Information System

GWR – Geographically Weighted Regression

GWPR – Geographically Weighted Poisson Regression

NT-TFV – Night-Time Theft from Vehicle

OLS – Ordinary Least Squares

TFV – Theft from Vehicle

VIF – Variance Inflation Factor

VPD – Vancouver Police Department

INDEX OF THE TEXT

INDEX OF TABLES

INDEX OF FIGURES

1. INTRODUCTION

The increasing rate of criminality across many urban settings has created the need to better understand the social and physical determinants of crime. One specific aspect is how certain urban characteristics may influence criminal activities. As Cornish and Clarke (1987) point out, offenders are decision makers acting on the evaluation of certain aspects. Research in this field is particularly important as, in one hand, it helps identify specific urban characteristics that foster criminal activities, eventually establishing foundations for urban policies and recommendations regarding infrastructure and urban design and, on another hand, it helps recognize specific areas where the built environment is fostering criminal activities, allowing for localized mitigation measures to be considered. The fundamental findings are that crime is influenced by the surrounding environment, that its distribution is not random, and that it is important to understand its patterns for crime prevention, control and investigation (Wortley and Townsley 2016).

Recent environmental criminology literature has indicated a need for criminologists to focus on site-level, as well as community-level, variables that influence property crimes as different, and even contradictory, conclusions have been pushed forward regarding the environment and crime prevention (Mao et al. 2021). The study of the potential relationship between the urban environment and crime incidence in an increasingly data driven world has, to date, to be fully developed. Research using geospatial technologies promote a better understanding about the relationship between the built environment, allied with economic and social indicators. Specifically, research in the field has yet to extensively investigate the relationship between street lighting levels and criminal activities with more detail, as there are only a few studies that make use of geospatial technologies for that purpose (Tagliabue et al. 2020); (Bappee et al. 2020).

One particular facet of the built environment, street lighting (and its possible effects on crime) has been the topic of innumerous research efforts in the last decades (see (Farrington and Welsh 2002, 2008); (Struyf 2020)) for reviews on the subject). Specific conditions may contribute to the selection of criminal targets by offenders and, in that sense, we could consider street luminance (or the lack of it) a micro-level "crime attractor" as set forth by Brantingham and Brantingham (1995): places that create criminal opportunities for specific crimes, so that motivated offenders are attracted to them. Ambient light influences the perception of safety for a potential victim, and it also influences a perpetrator´s risk analysis, affecting how it envisions both costs and rewards for committing a crime. This can happen as the perceived belief of apprehension increases, either from the enhanced possibility of witnesses (or electronic surveillance) being able to better ascertain its actions or from law enforcement having more leverage towards an arrest (Chalfin et al. 2020). Policy makers and law enforcement have presumed that improved lighting increases the anticipated costs of crime and the likelihood of capture (Doleac and Sanders 2015).

In particular, theft from vehicle (TFV) is a very common type of crime. In the study area of this thesis, the city of Vancouver, Canada, it was the most common crime until 2008, and it is still the second most common type of crime until today, after theft of property, accounting for more than 12 thousand reported incidents during 2016 (Vancouver Police Department 2021a). It is also an expensive crime. The United States of America, in 2016, reported 1,301,447 thefts from vehicle and an average value of 869 USD (U.S. Department of Justice - Federal Bureau of Investigation 2017). In Canada, the estimated cost of pain and suffering for non-fatal injuries from the four types of property crimes was $3.63 billion in 1999 ($1.02 billion for breaking and entering, $383 million for motor vehicle theft, $1.44 billion for non-vehicle theft, and $788 billion for vandalism) (Leung 2004). TFV is also an outdoor crime, as it tends to happen in public streets, and lighting levels could play a role in its incidence.

How crime determinants affect the spatial distribution of criminal occurrences is the main question that the proposed research will try to address. The study will particularly look into the relationship between street illuminance levels, through street light pole density as a proxy, and the prevalence of reported Night-Time Theft from Vehicle (NT-TFV) crimes. Illuminance is the measurement of the amount of light falling into and illuminating a given surface in lux.

1.1 Objectives

The research questions that will guide the thesis objectives are:

Do GWPR models benefit from including social disorganization (demographic) variables?

Does street illuminance affect reported thefts from vehicles during night-time in the city of Vancouver's neighborhoods? If yes, in which areas?

What is the local relationship between street lighting pole and tree densities and collected night-time street illuminance?

Is street lighting pole density a usable proxy for street illuminance?

Generally speaking, this thesis aims to examine the spatial relationship between crime determinants and NT-TFV. The primary interest is to investigate how a particular characteristic, public street illuminance, influences NT-TFV in Vancouver, and how. Since data on street illuminance is not available a proxy, street light poles, is used in the investigation. It is also an objective to understand how appropriate this proxy is.

The specific objectives of the research are to:

- Investigate the spatial distribution of Night-Time Theft from Vehicle (NT-TFV) in the city of Vancouver, British Columbia, Canada, through exploratory spatial data analysis.
- Identify and collect data on potential explanatory variables for NT-TFV.
- By fitting a Geographically Weighted Poisson Regression (GWPR), model a possible non-stationary association between "street lighting", and a series of other crime determinants, and NT-TFV, in two models. The first model includes social disorganization (demographic), routine activities (land use) and street level physical variables, while the second model only considers routine activities (land use) and street level physical variables.
- Devise a hardware configuration and associated programing sketch that enables an easy-to-use georeferenced mobile illuminance collection station based on the senseBox.
- By fitting an Ordinary Least Squares (OLS) model the relationship between street lighting pole and tree densities and measured night time street illuminance.

1.2 Organization

This thesis is organized in five chapters: Introduction, Literature Review, Data and Methods, Results and Conclusions. The introduction simply lays out the justifications for such a study, the research questions, and specific objectives. The second chapter, the Literature Review, investigates the fundamental theories and existing research on the general topic of environmental criminology and into specific street, neighborhood and general demographic level characteristics and their relationship with crime, leading into the rationale on the variable selection. Chapter three, Data and Methods, lays out information on the study area, how data was gathered and manipulated to fit the research needs. It also describes the methodology utilized to analyze the data in two parts: NT-TFV and Illuminance. The fourth chapter will present and discuss the results of the study, including the exploratory spatial data analysis and regressions. Lastly, chapter five will be dedicated to the conclusions and to discuss limitations and future research.

2. LITERATURE REVIEW

The following sections are dedicated to the literature review. Starting with a general review of environmental criminology concepts and theories it follows specific sections dedicated to street lighting and to other known demographic and environment crime explanatory variables.

2.1 Environmental criminology

Where are crimes happening? What is the spatial nature of criminal activities? What variables can explain it? As a non-random phenomenon, crime happens somewhere for a reason and, if not, it might be transferred somewhere else. The term environmental criminology can be seen as a broad term that houses a set of theories that focus on the time and place feature of crimes (Andresen 2010), or the circumstances in which those crimes occurred (Wortley and Townsley 2016). The intention is to investigate the environment where a crime occurred, rather than just looking into the characteristics of offenders and/or victims. Early social ecologists, such as Gerry and Quetelet, were interested in how social conditions (such as age and family structures, education levels, population diversity, etc.) influenced crime levels in specifically organized population aggregates (Anselin et al. 2000). An important development in the early 20th century were the ecological works of the Chicago School and, specifically the work of Shaw and Mckay (2010) that led to the social disorganization theory. This theory looks into the sociological influences on law-breaking actions, or how neighborhood characteristics relates to crime (Andresen 2010). The fundamental idea is that the lack of social cohesion characteristics in a neighborhood, such as sparse friendship networks, unsupervised teenage peer groups, and low organizational participation, mediates effects of community structural characteristics such as economic status, ethnic heterogeneity, residential mobility, and family disruption, leading to an increase in crime rates (Sampson and Groves 1989). Social deprivation at some level is generally the strongest positive correlate with crime (Andresen 2006).

Research on the topic of environmental criminology boomed in the 1970s (Brantingham and Brantingham 1993) and has seen a wide range of investigations since then. Seminal works by Jacobs (1961), Angel (1968), Jeffery (1971), and Newman (1972), helped shape the theoretical framework regarding the urban environment and crime.

Specifically, Jeffery coined the expression "crime prevention through environmental design" (CPTED) to depict a series of factors that could somehow influence criminal behavior (both environmental and social), altering the criminal model into a high-risk and low reward endeavor (Jeffery 1971). Newman introduced the concept of "defensible space" to propose, through architectural and urban design, physical aspects that impact crime susceptibility by enhancing both a territorial behavior and a sense of community (Newman 1972). Three fundamental environmental criminology theories, known as opportunity theories of crime, are important to mention: the routine activities, the crime pattern, and the rational choice.

The routine activity theory (Cohen and Felson 1979) focus on the circumstances in which offenders execute crimes, rather than on the offenders characteristics. The premise is that, for a crime to occur, a suitable target and a motivated offender (allied to the non-existence of guardians) must come together in space and time. This conditions, and consequently the criminal act, are more prone to occur during daily events. The diffusion of human activities away from households intensifies the opportunity for crime and, therefore, produces more crime. This condition explains crime upsurge in western countries after the second world war, where an improved family income helped foster activities outside the households (like going to the cinema or eating at a restaurant), allied with societal changes (such as an increased presence of woman in the labor force and young people leaving home to study) led to an amplified exposure of people to less protective environments (Andresen 2010). As different places reflect different demographic, economic, social and cultural characteristics, different routine activities are expected. For this, crime is not random or uniformly distributed (Andresen 2006).

The crime pattern, or geometry of crime, theory focuses on the perceived opportunities for crime that exist in a dynamic urban structure – a backcloth that includes physical attributes, social, cultural standards, legal and political considerations, etc. (Brantingham et al. 2016). Its dynamism has different frequencies, depending on what is being observed (Andresen 2010). For example, a road network changes in a relative slow rate, while the presence of cars on that same road network may change drastically between day and night of the same day. It also connects the places in which we spend time (such as work or school) and the pathways between them, to criminal activities. Brantingham and Brantingham (1993) argue that offenders tend to focus their criminal actions on nodes and paths that are part of their daily traveling patterns. In that sense, criminal opportunities tend to present themselves where both offenders and victims interact, and criminals are more confident in spaces where they have a better understanding of targets and possible risks in one hand (Wortley and Townsley 2016) and also, as it takes less time and effort, criminal prospects will tend to coincide with the offender's regular activities (Andresen 2010). Crimes tend to cluster near spaces where victims and criminals interact, with higher intensities at the nodes. When a ready offender finds a suitable target and faces the right circumstances (where perceived benefits exceed apparent risks), a crime may occur (Brantingham et al. 2016). This idea is expressed in the form of crime generators (nodes with high flows of individuals attracted for reasons other than crime but that create criminal opportunities) and crime attractors (nodes where criminals are attracted for known opportunities for a specific type of crime). Nodes are complemented with paths (traveling corridors) and edges (boundaries). Edges usually involve a mix of activities, making it easier for criminals to blend in. Crime generators and crime attractors are places that become crime hotspots (Brantingham et al. 2016).

The rational choice theory/perspective presumes that offenders will act in response to the perceived opportunities (cost vs. benefit) of specific offenses, recognizing the influence of the environment on the offenders behavior (Cornish and Clarke 1987). It assumes a great variety of motivations to commit (or not) a certain crime, such as psychological, familial, social, and economic factors, and a set of questions that must be answered by the offender: whether or not to commit crime at all, whether or not to select a particular target, how frequently to offend, and whether or not to desist from crime (Andresen 2010). The situational crime prevention aims to reduce criminal opportunities by increasing risks and decreasing rewards with a methodology for challenging specific crime problems and opportunity reducing techniques (Clarke 1995). One of the premises is that decision processes and information used should vary depending on the type of crime as the means certain crimes are conducted have implications for its prevention.

From one perspective, the social disorganization theory is more focused to the origins of motivation than the opportunity theories, as it looks into the social structural factors of an environment that may be favorable to criminal activities but, on another perspective, it fails to consider that individuals need an actual and tangible opportunity to commit a crime. This opportunistic condition is not persistent across all times and places, offenders will not necessarily find a suitable chance. Inversely, opportunity theories tend to assume the existence of a suitable reserve of motivated offenders that will find the right opportunities to commit a crime (Rice and Smith 2002). Social disorganization and opportunity theories have been suggested to be complementary and should be integrated to improve the analytical capability of spatial analyses (Braga and Clarke 2014).

2.2 Street lighting

As a mechanical adjustment to the environment, street lighting is a strategy included by the Crime Prevention Through Environmental Design (CPTED) approach, but it is certainly not a physical obstacle to criminal activities (Painter 1994). The CPTED approach has suffered some criticism as, obstacles to a specific crime will simply result in the "displacement" of such crime, and the overall crime is not really being reduced but rather displaced (Cornish and Clarke 1987). As an opportunity reducing technique, street lighting is included in the "natural surveillance" category of Twelve Techniques of Situational Prevention (Clarke 1995). The effect of lighting on crime can be ambiguous as, in one hand, it can help the effectiveness of possible witnesses (or even surveillance technologies), and it can help promote the use of outdoor spaces at night (as they may be perceived as safer). On the other hand, more people outside also creates more criminal opportunities. Improved lighting can also increase visibility in a way that can allow for a better assessment of potential victims by offenders (Farrington and Welsh 2008).

Literature on the subject of the impacts of street lighting and crime is, to an extent, ambiguous. Jacobs (1961) argues that sidewalks are the main urban scenario that keep a city safe. In that context, street lighting is important for the reassurance it offers but, unless there is availability of surveillance (eyes), lighting is not that useful. Ramsay (1991) examined the impact of improvements to street lighting on both crime and the people's sense of fear, suggesting that lighting improvements are in general more likely to have a positive impact on the public's fear of crime than on the incidence of crime itself. Likewise, Atkins et al. (1991) could not find support to the hypothesis that better street lighting reduces reported crime. Nevertheless, the introduction of lighting in, so called, "black spots" could be beneficial to decreasing the occurrence of crime. Painter

(1994) investigated how lighting could, potentially, impact crime and found consistent evidence that lighting improvements can reduce crime. Similarly, Painter and Farrington (1999) suggest that the effects of better street lighting on crime had two different causal routes: the increased visibility, street use and surveillance after dark, decreased perceived opportunities and increased perceived risks for criminal perpetrators; and, on the other hand, led to an improved community unity that yielded informal social control (Zeng et al. 2021).

In a classic meta-analysis of the effectiveness of street lighting on crime, Farrington and Welsh (2002) summarized the results of thirteen studies, conducted in the United Kingdom and the United States of America, that occurred between the mid-1970s to late 1990s. The more recent British studies agree that improved lighting positively affects crime, while the American studies have mixed outcomes. This review suffered some criticism on its methods. In particular, Marchant (2004) comment on the chosen statistical methods that yielded narrow confidence intervals, due to overdispersion and additionally, as the result of the chosen study areas, a "regression to the mean" effect is expected.

In a more recent review of studies on the relationship of the street environment, and its two elements (paths and nodes along paths) with crime, Mao et al. (2021) conclude, based on their review, that appropriate street lighting levels are beneficial for the prevention of crime as it promotes better surveillance during the night and, as expected, inadequate lighting will trigger crime. Mao et al. (2021) also stress the lack of studies regarding what is proper lighting in terms of illuminance, height of source and light color.

Doleac and Sanders (2015) investigated how daylight-savings time policies in the United States of America affected criminal activity by exploiting the amount of daylight changing unevenly while other factors suffer a smooth transition over the years. The premise is that greater amount of light will increase the probability of arrest, decreasing the propensity to follow through. On the other hand, more light will influence how much time individuals remain outdoors, increasing the number of potential victims. The study found daylight savings lowers robbery rates by 7%, with most decline during the hours affected by the daylight-savings.

Kaplan (2019) examined how a low light dosage might affect crime by investigating how moonlight and proportion of clouds affected outdoor night crime. The study found that brighter nights, where the moon is full and there are no clouds in the skies, present more total crime than nights without a moonlight. Moreover, the research determined that violent crimes were most effected, with a 7.9% increase, and that property crimes were not affected.

Employing a novel supervised machine learning approach, Dakin et al. (2020) tried to use street view images of home facades to obtain risk predictability, based on the dependency between physical features and crime incidents. Among other features, street lighting is an example of a feature with relevance to crime. For each crime site within the study area the study determined a statistical measure of dependency to establish what (if any) features are correlated to different types of crimes, using a chi-squared statistic. According to the study, streetlight only indicated a strong dependency value (a feature identified as to be most frequent for a certain type of crime) with violence and sexual offenses and showed the least dependency to vehicle crimes considering the included features. However, as the analysis mentions, streetlights can have a greater likelihood of being left out of street view images due to their height (or tree coverage). Nevertheless, the study could not find, in general, a strong dependency between the features and crime types tested.

Using eight years of lighting outage data in Chicago, Chalfin et al. (2020) investigated nigh-time crime on street segments affected by light outages, finding evidence that outdoor night-time crimes are susceptible to street lighting conditions. This study considered impacts on both the outage affected street segment and neighboring areas alike. The study found little evidence that lighting outages impact crime in the specific segment where it occurred but found evidence that crime, in general, increases in adjacent areas. Street light outages that go without repair for some time can have echoing effects throughout a neighborhood. The only exception are motor vehicle thefts that, according to the study, there was evidence of an increase on the specific street segment experiencing the street lighting outage. Chalfin et al. (2020) also conclude that neglections regarding the repair of street light outages contribute to uneven crime rates among Chicago's neighborhoods as they found evidence that areas with higher crime rates suffer more light outages and, additionally, outages also take longer to repair.

Bappee et al. (2020) proposed a data-driven approach to test how prediction accuracy can be improved by integrating different data such as streetlight infrastructure, demographic characteristics, and social media human mobility data. Specifically, the research used streetlight poles obtained from the Streetlight Vision, a smart city central management software, and tested two classifiers, Random Forest and Gradient Boosting, against a deep neural network baseline model in several ensembles that featured different combinations of data. Results suggest that both demographic characteristics and streetlight have strong associations with crime, yielding performance improvements in prediction, and that Gradient Boosting outperforms the other classifiers.

Chalfin et al. (2021) conducted a novel experiment in selected New York City public housing projects, where they found evidence that areas given more lighting experienced decreased night time outdoor crimes. The authors also discuss the efficiency of control strategies such as lighting enhancements. The idea is that such strategies can have a deterrence effect, when a possible offender decides not to commit a crime, or they can have an incapacitation effect, when such strategies lead to arresting, adjudication, and confinement. The last being a costly endeavor, the ideal scenario will yield large enough deterrence effects to the point where both crime and incarceration are reduced. The study concluded that the lighting interventions employed in the experiment had deterrence as the dominant mechanism that reduced both crime and arrests, and that the physical environment is important for crime reduction.

2.3 Street trees

As a variable that is deeply connected, in the physical sense, with street lighting, street trees are separated from other explanatory variables. Vegetation has been found to be significantly related to crime in several studies ((Kuo and Sullivan 2001); (Ye et al. 2018)). Vegetation can be seen as having an ambiguous effect, much like street lighting does. In one hand, the presence of vast and well cared vegetation reflects social cohesion that, based on social disorganization theory, will have a negative effect on crime as guardianship in such areas tends to be higher. On the other hand, it also stimulates the use of outdoor areas, creating more criminal opportunities. Vegetation can also disturb surveillance, as it can offer more hiding places or impact the effectiveness of street lighting during dark hours.

Kuo and Sullivan (2001) conducted a novel research to examine the relationship between vegetation and crime in public housing buildings in Chicago. An Ordinary Least Squares (OLS) regression indicated that vegetation (grass and widely spaced high canopy trees) is significantly related to crime, accounting for 7% to 8% of the variance in the number of crimes. The greener a building surroundings the fewer crimes (property and violent). The authors suggest that vegetation that does not affect visibility can constrain crime by increasing the surveillance and by mitigating psychological precursors to violence. Other investigations that incorporated an OLS regression include Troy et al. (2012), that investigated statistical associations between crimes and tree coverage in Baltimore city and Baltimore county, a very heterogeneous region with regard to crime rates and land use, and Gilstad-Hayden et al. (2015), that examined the relationship between crime and vegetation in New Haven and concluded that tree canopy coverage was associated with lower crime rates (violent, property and total crime), independent of selected demographic indicators.

Ye et al. (2018) investigated spatial patterns of property and theft crimes in downtown Vancouver by fitting OLS, spatial lag, and Geographically Weighted Regression (GWR) regression models that included vegetation, road density, light density, presence of graffiti, and several demographic variables. The results show an association between tree coverage and property crime rates. Specifically, the GWR model shows significant negative correlations in the downtown area of the city.

More recently, Lin et al. (2021) described crime-facilitating mechanisms pertaining to urban trees and classify them into “crime-facilitating” and “crime-deterring” categories. Trees can be facilitating crime as they may reduce visibility, provide camouflage or barriers for criminals. Trees can also be seen as neglected or have poor maintenance, denoting a poor social organization. On the other hand, well kept trees can pass a sense of orderliness and security, denoting a cohesive neighborhood, and can also attract an increased public usage of a street, increasing surveillance on one hand, and opportunities on the other. The investigation included OLS and a spatial Durbin model (accounting for the spatial spillover effect) with the later model achieving a better data fit.

2.4 Other covariates with crime

Explanatory variables are, in general, associated with either the social disorganization or the routine activities theories. In that sense, many studies have been conducted with quite heterogenous sets of explanatory variables.

Shen and Andresen (2021) summarized both social disorganization and routine activities theories possible explanatory variables that influence crime, along with the expected directional outcome (positive or negative).

Regarding the routine activities, the authors suggested the following groups and variables, with direction in parenthesis:

- Target quantity: resident population (+), ambient population (+), ambient–resident population ratio (+), number of dwellings (+), number of commercial outlets (+).
- Target value: government assistance (-), low income (-), unemployment rate (-), median house value (+), median monthly rent (+), median household income (+), postsecondary education (+).
- Guardianship: renters (+).

Regarding the social disorganization theory, the authors suggested the following groups and variables, with direction in parenthesis:

- Ethnic heterogeneity: aboriginal identity (+), immigrants (+), recent immigrants (+), visible minorities (+), ethnic heterogeneity (+).
- Population turnover: moved 1 year ago (+), renters (+).
- Economic deprivation: government assistance (+), low income (+), unemployment rate (+), housing, major repairs (+), post-secondary education (-), median house value (-), median monthly rent (-), median household income (-).
- Populations at risk: resident population (+), ambient population (+), ambient-resident population ratio (+), number of dwellings (+), number of commercial outlets (+).

As mentioned before, both theories have been suggested to be complementary and should be combined. Specifically looking at the street level crime, Jones and Pridemore

(2019) suggest that street level concepts will have expected main effects but those effects will be conditioned by neighborhood effects, supporting a multilevel theory of crime concentration. Shen and Andresen (2021) also indicate that, depending on the theory being analyzed, the expected direction of the relation for a certain variable will be different. For instance, for the routine activities theory, socioeconomic status essentially create more opportunities for crime, more valuable targets. But, for the social disorganization theory, an improved socioeconomic status will decrease crime, less necessity.

2.5 Rationale on independent variables

The following paragraphs will address known covariates of crime that were selected for the regression analysis presented in this work.

Potential explanatory variables included in the presented GWPR models (see sections 4.2 and 4.3) can be classified in many ways. The two variables previously discussed, street light and street trees, are physical attributes of street segments that are aggregated into Census Dissemination Areas (CDAs). Other variables are related to the general land use classification or specific activity of certain areas, or, in other words, what sort of routine activity is pursued in a specific area of the city. The variables are mixed land use rate, commercial land use rate, distance to closest rapid station, distance to parking land use, distance to liquor stores, and number of bus stops aggregated into CDAs. Other three variables are related to the social disorganization theory namely total median income, recent immigration rate, and rate of commuting drivers in the population.

The idea of transit stations and land use (particularly certain types of commerce) to have an influence on crime has been investigated by several researchers. The risk of crime associated with the presence (or absence) of bus stops and transit stations alike is thought to be strongly associated on the types of commercial activities to be found in those activity nodes (Hart and Miethe 2014). It is also a consensus in the literature that criminals tend to commit offenses close to where they live; there is a distance-decay function where a criminal is less likely to commit a crime the further away from his residency. However, transit systems can reduce the distance for a criminal to reach a certain target, subjecting new areas to their actions through an easier “journey to crime”

(Liggett et al. 2003). Regarding crimes against vehicles, the premise is that bus stops tend to be located in business areas (crime attractors and generators) that feature abundant parking spaces (either in parking facilities or street parking) and that motor vehicle related crimes will tend to gather around these areas (Yu 2009). Commercial and mixed use land use, with an assortment of services and products, can act as crime attractors and generators. Areas dedicated to parking have a high quantity of (possibly) unattended vehicles and are a crime attractor. Nevertheless, even though some land uses facilitate crime, not all specific units in a certain area will have the combined characteristics that will somehow attract potential offenders and, for that reason, will not experience crime (Kinney et al. 2008).

Block and Davis (1996) compared the geographic distribution of street crime in four police precincts in Chicago, where two precincts had low robbery rates and the other two very high robbery rates. The authors compared the risk of robbery in areas around rapid transit stations against areas further away, finding two distinct patterns. Rapid transit stations, and commercial areas connected to them, attract and bring together both targets and offenders (criminals that either live in the area or travel to it). In areas with high rates of robbery, crime was dispersed along main commercial streets, whereas areas with lower rates of robbery had crime concentrated near rapid transit stations.

Kinney et al. (2008) analyzed assault and motor vehicle theft patterns in relation to the distribution of land uses in Burnaby, British Columbia, and concluded that specific land uses act as major crime attractors and generators. Generally, although the city is dominated by residential land uses, commercial and civic-institutional-recreational land uses concentrate crime. Specifically, commercial land use is the main generator and attractor of crime, especially for motor vehicle thefts. Consistent with environmental criminology, the study found that the land use subcategory with the higher motor vehicle thefts rate was the shopping center, probably due to a high concentration of vehicles in parking lots and people in the shopping malls. The study also notes the multi-use characteristics of the surroundings of many of the shopping malls, including transit stations and high-rise condominiums. Vehicle theft rates were also found to be high in educational land uses.

Yu (2009) investigated the influence of bus stops and commercial establishments in crime opportunity levels of robbery, aggravated assault, motor vehicle theft, and theft from motor vehicle in Newark, New Jersey, by fitting four regression models that accounted for a different combination of explanatory variables, and tested with different regression techniques including OLS, spatial lag and spatial error models. The hypothesis was that commercial land use and bus stops would increase the opportunity for motor vehicle related crimes. The study found that, for all types of crimes examined, the number of bus stops was associated with increased crime and that, regarding the business types affecting crime, in general, retail businesses exerted a greater influence than services. Specifically, certain businesses were associated with an increase in thefts from motor vehicle criminal opportunities: food stores, automotive dealers, gasoline stations, and automotive related services. Also, mixed land use and the presence of vacant land were related to increases in both motor vehicle theft and theft from motor vehicle, and higher education institutions were related to a lower theft from motor vehicle risk.

In Henderson, Nevada, Stucky and Smith (2017) investigated whether bus stops have an autonomous effect on crime or if the characteristics of its surroundings are the true drivers of crime. Using a negative binomial regression that included socioeconomic, land use (types of activity nodes) and bus stops as explanatory variables, the authors found that bus stops were associated with both violent and property crimes and that the impacts of bus stops on crime are enhanced in commercial and industrial areas and lessened in high density residential areas.

Sypion-Dutkowska and Leitner (2017) examined the types of land use that influence the spatial distribution of several crimes, including car crimes, using multiple ring buffer and crime location quotient techniques, in Szczecin, Poland. Generally, a strong attraction of total crimes occurs within a fairly short distance (0 to 50 meters) to some land uses such as alcohol outlets, discos and clubs, cultural facilities, municipal housing, and commercial buildings. Land uses that detract crimes are sport facilities, sacral buildings, industrial plants, buildings of justice, and university buildings. Specifically for car crimes, alcohol outlets strongly attract, and hotels attract, those types of crimes, while bus and tram stations detract car crimes.

Andresen (2006), looking through the social disorganization and routine activity theories perspectives, investigated crime in Vancouver employing a spatial autoregressive regression procedure. Results show that unemployment rate, population change, and the standard deviation of average family income have positive relationships with the three crime rates investigated: automotive theft, break and enter, and violent crime. Specifically for automotive theft, population change had a significant and positive impact, while ethnic heterogeneity had a strong negative effect.

In a reanalysis of the work of Andresen (2006), Shen and Andresen (2021) examined if (and how) neighborhood characteristics, from the social disorganization and routine activity theories, could still explain crime in Vancouver, finding solid motives to advocate for the use of social disorganization and routine activities approaches to understand spatial patterns of property crime.

Hodgkinson et al. (2016) investigated the changes in the spatial patterns of auto theft in Vancouver, between 2003 and 2013. Using hot spot analysis to follow concentration shifts, and spatial point pattern tests to assess the degree of similarity between two datasets to confirm either a possible spatially homogeneous auto theft reduction or if decline occurred in particular locations only. Results show that auto theft in Vancouver is spatially concentrated but, the concentration location shifted over the ten years from wealthier neighborhoods to an extremely socially disorganized area, downtown east side.

More recently, Ristea et al. (2018) made use of social media data during professional hockey games to analyze the spatial patterns of six property crimes in Vancouver. The research focused on the differences between home games, away games and no game days, and found an increase in criminal events during home games that varies depending on the type of crime. Specifically, for Theft from Vehicle (TFV), the study found increased concentrations around the arena and the downtown area during home games. The research also ran a series of GWR models to examine the spatial performance of local models. Generally, the geographic proximity to the arena and to the alcohol district led to increases in R^2 values. Explanatory variables included tweets, population, parks, public roads, street parking, disability parking, motor vehicles parking, street light poles, rapid transit stations, traffic signals, pubic washrooms, and liquor stores. With large numbers of spectators drawn to the arena during game days there is a normal increase in crime related tweets, an important explanatory variable for all crimes during games. The GWR models revealed that theft-from-vehicle and mischief had a stronger connection with crime-tweets.

Similarly, Andresen et al. (2021) investigated spatial variations in the effects of long- and short-run unemployment relationships using GWR, using a large set of control explanatory variables: Unemployment (long run and short run), population change, rented homes, major repairs in houses, old houses, recent moves, postsecondary education, low income, government assistance, average dwelling value, average rent, median family income, aboriginal identity, immigrants, recent immigrants, visible minorities and ethnic heterogeneity.

Recent investigations that involve a GWPR model include Deng et al. (2021), that analyzed the association of eight streetscape indexes of the street built environment (acquired using Google Street View images) and crime in New York City using a GWPR model to investigate influences on crime activities; Tavares and Costa (2021) that investigated crimes against property in Portugal to determine what factors (demographic and socioeconomic) may influence property crimes in different municipalities by using a series of regression models that included GWPR; and Arnio (2021) that examined how levels of violence, racial and ethnic composition and socioeconomic disadvantage are related to shootings involving police officers in Houston, using GWPR models to evaluate local differences in this relationship.

3. DATA AND METHODS

3.1 Study area

The city of Vancouver is the largest city in British Columbia, Canada, with, according to the 2016 Census, an estimated population of 631,486 habitants, a population density of 5,492 habitants/km^2, in a land area of approximately 115 km^2 (Statistics Canada 2017). The city can be divided into 22 local planning areas (commonly known as neighborhoods) and it contains 17,033 public street segments (City of Vancouver 2021). In 2016 the city had a passenger vehicle population of 280,414 vehicles, 31,325 commercial vehicles, 1,507 motorhomes, 10,006 motorcycles, and 8,904 trailers (Insurance Corporation of British Columbia 2021).

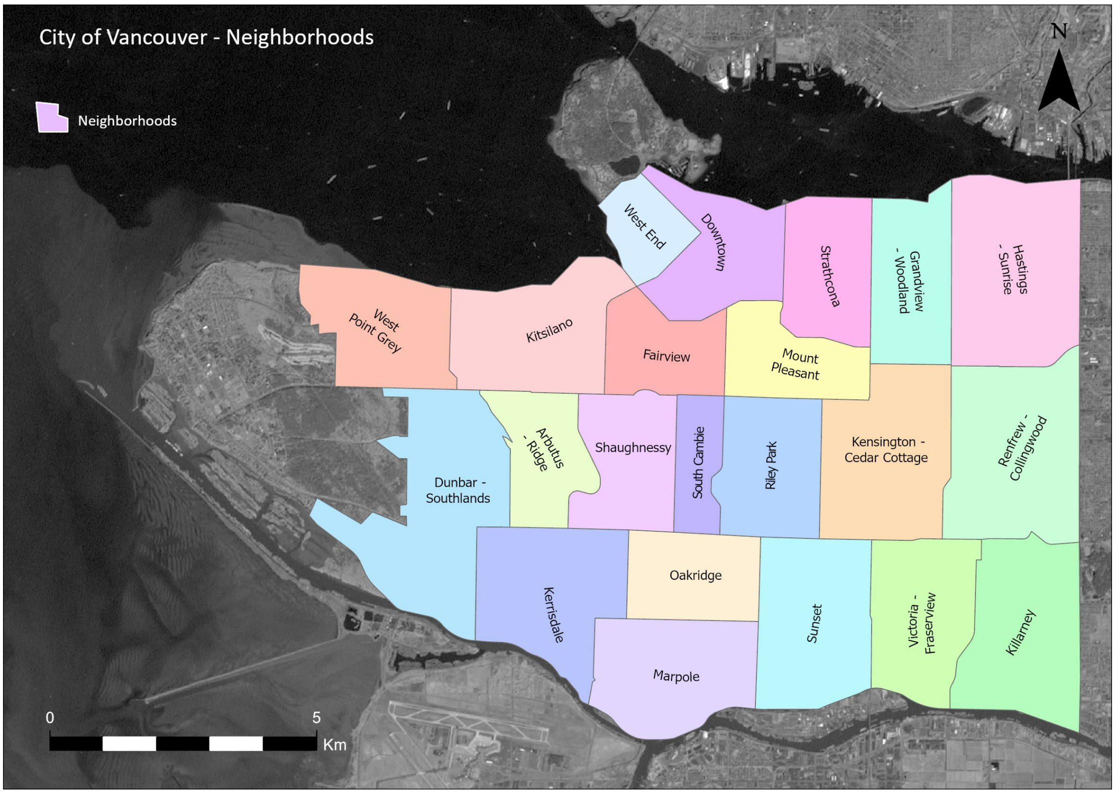

Figure 1 - Study area and neighborhoods

Crime in Vancouver had a constant decrease in its crime rate per 1000 inhabitants from 2004 until 2012. Similarly, the TFV rate also decreased since 2004 but started to go up one year earlier, in 2011. For 2012 to 2016, the time span where crime data was collected for this study, total crime was relatively stable from 2012 to 2015, experiencing a rise in 2016, while both property crime rates and TFV rates were stable in the first two years and endured a considerable increase in the subsequent years (Table 1) (Vancouver Police Department 2021a).

Year / Crime rate / 1k habitants	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Total crime	95.3	90.1	83.1	77.6	75.8	74.6	74.7	72.66	73.58	78.95
Property Crime	68.9	64.2	57.8	53.1	50.2	51.1	51.3	55.57	56.90	61.99
TFV crime	20.1	18.5	15.9	13.6	11.7	12.4	12.6	14.90	15.58	18.63

Table 1 - Crime rates (2007 to 2016) per 1000 inhabitants for total crime, property crimes, and theft from auto (adapted from (Vancouver Police Department 2021a))

Between 2012 and 2016 the most common crime in Vancouver was “Other theft” (the theft of property that includes personal items) with 53706 reported incidents, closely followed by “Theft from Vehicle” (theft of property from a vehicle) with 49934 reported events. These two categories account for more than 58% of the crime reported to the Vancouver Police Department (VPD) (Figure 2).

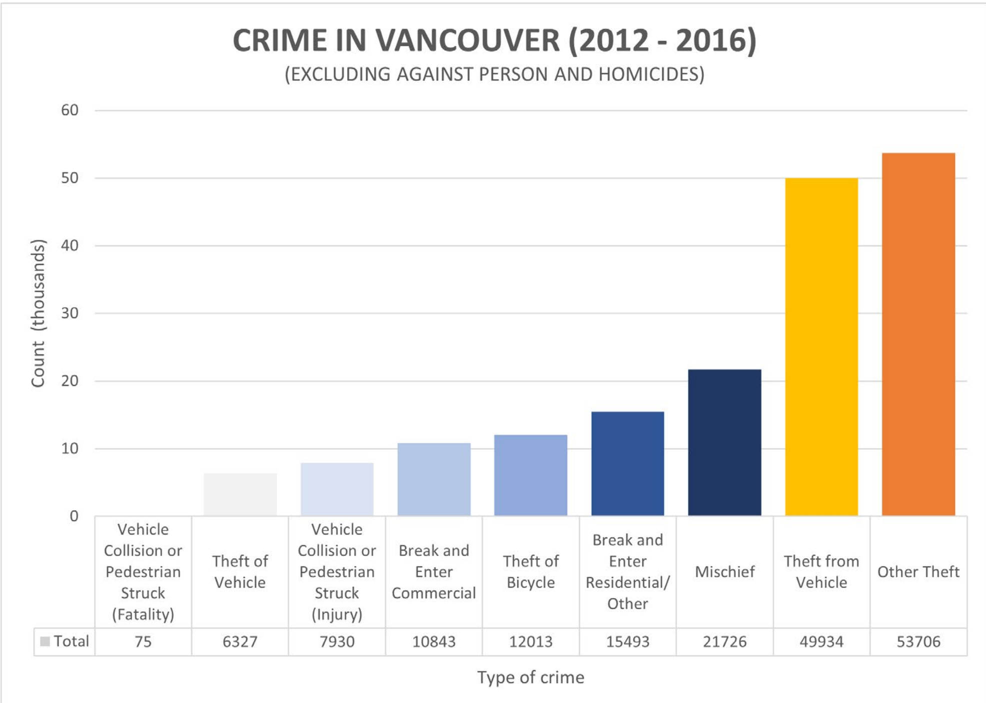

Figure 2 - Crime in Vancouver (2012 to 2016)

TFV crime rates were relatively stable in 2012 and 2013, after several years of decline, and endured a considerable increase in the subsequent years. In 2012 there were 7933 TFV reported events and by 2016 this number raised to 12720, a 60% increase in a short time (Figure 3).

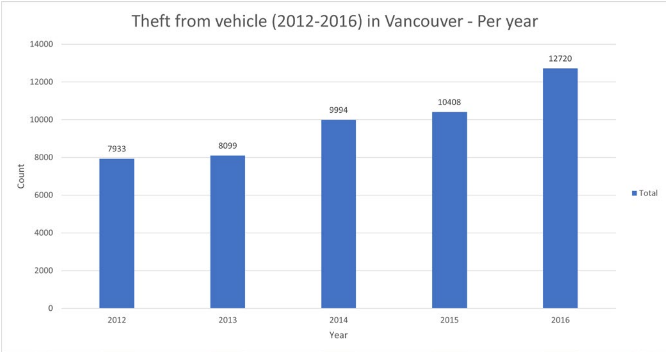

Figure 3 – TFV between 2012 and 2016

Reported TFV distribution was stable during the year, peaking in the month of November and, as expected, with lower counts during the month of February (Figure 4).

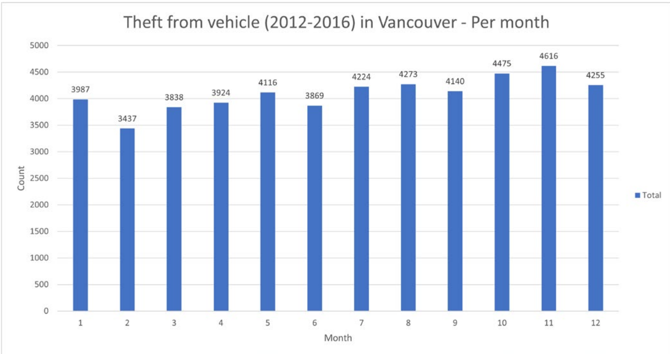

Figure 4 – TFV between 2012 and 2016, per month

The hourly distribution of TFV was quite interesting, showing an intense daily variation. Census data shows that, in the Vancouver census subdivision, roughly 51% of the employed labor force individuals leave to work between 7 am and 8:59 am, and 45.4% of these drive a car, truck, or van as the main mode of commuting (Statistics Canada 2017). There was a notable increase in TFV in the study area at 7am, declining after 10 am. At noon, another increase in reported TFV, followed by a decrease and subtle increase until 5pm. From 5pm to 6.59 pm there is a sharp increase in the incidence of reported TFV, followed by a stable, but high, incidence through the night hours, peaking, between midnight and 1 am and then sharply dropping (Figure 5). The sharp increase at mid night is curious, as it does not follow the expected pattern. This might be the result of when victims found about thefts occurring on their vehicles rather than the

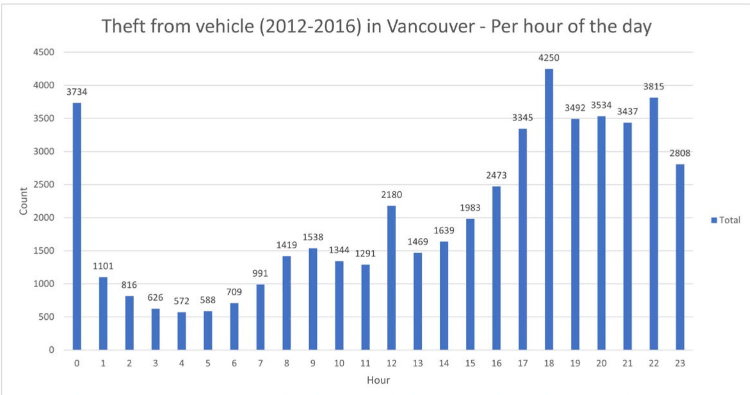

Figure 5 - TFV between 2012 and 2016, per hour

actual time of the occurrence. Alcohol serving establishments close at 1am from Sunday to Thursday and at 2am on Fridays and Saturdays. As noted in the literature review there is a known link between alcohol outlets and crime and the Downtown area (known as the alcohol district) concentrates the majority of TFV crimes in the city, followed by adjacent neighborhoods such as East End and Strathcona (Figure 6). Regarding days of the week, there is an almost uniform distribution throughout the days of the week, with a minor increase from Friday to Monday (Figure 7).

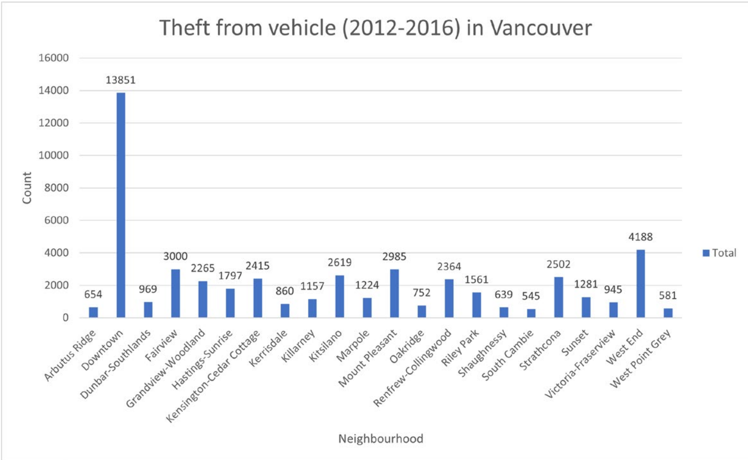

Figure 6 – TFV between 2012 and 2016, per neighbourhood

Figure 7 – TFV between 2012 and 2016, per day of the week

3.2 Data

Data for the analysis is divided into four categories: demographic, environment, crime, and illuminance data. The following sections will, for each data category, provide a basic description and source of each data set, and depict any transformations it may have undergone in order to be utilized by the research.

3.2.1 Demographic data

Demographic data, available from Statistics Canada latest available Census, 2016, is organized around Census Dissemination Areas (CDAs), a small area composed of neighboring dissemination blocks. The study area has 994 CDAs with an average population of 635.4 individuals and an average area of 0.11 Km^2. Data was aggregated into dissemination areas boundary shapefiles provided by Statistics Canada (Statistics Canada 2017). The analysis used data from the following 2016 Census topics: "Population, Age, Immigration and Ethnicity", "Journey to work", and "Income".

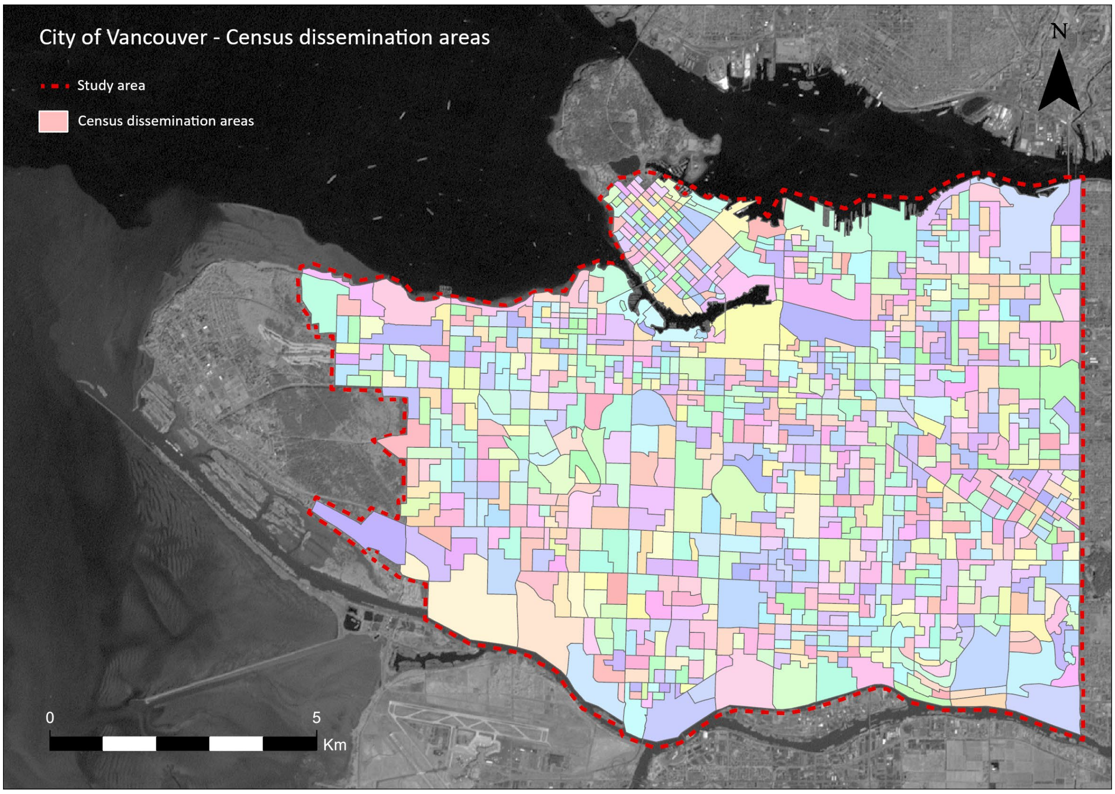

Figure 8 – Study area and 2016 Census Dissemination Areas

In the "Population, Age, Immigration and Ethnicity" topic, the "Population" (count of individuals per CDAs) is used to calculate "NT-TFV per 100 thousand habitants", to calculate the "Recent immigration rate", and "Commuter driver to population rate".

In the "Population, Age, Immigration and Ethnicity" topic, the "Immigrant status 2011 to 2016" (count of individuals in private households where the period of immigration was between 2011 and 2016, per CDAs) is used to calculate "Recent immigration rate" by dividing the number of recent immigrants by the number of habitants.

In the “Journey to work” topic, the “Drive as main mode of commuting” (count of individuals aged 15 or more, in private households, that drives a car, truck or van as the main mode of commuting to work, per CDAs) is used to calculate “Commuter driver to population rate” by dividing the number of commuting drivers by the number of habitants.

In the “Income” topic, the “Median total income” (median total income in 2015 among recipients aged 15 years and over in private households, per CDAs) is used as one of the independent variables in the regression analysis, alongside with calculate “Recent immigration rate” and “Commuter driver to population rate”.

3.2.2 Environment data

Environment data includes the following datasets: public streets, street trees, street lighting poles, location of bus stops, location of rapid transit stations, location of liquor stores, and land use.

Public streets, location of street trees, location of street lighting poles and location of rapid transit stations datasets are provided by the City of Vancouver Open Data Portal in WGS 1984 (City of Vancouver 2021) and were projected to NAD 1983 UTM Zone 10N coordinate system.

“Public streets” is a vector dataset containing the city’s street centerline network in segments broken at the intersection of two or more city streets or alleyways. It contains 17,033 street segments, including the “Shape length” for each record.

“Location of street trees” is a vector dataset containing the location of the city’s public trees, not including park trees. It contains 124,719 street trees. A basic spatial join was conducted so that each tree record is aggregated into one, and only one, specific CDA, yielding a “Tree count” attribute for each CDA. Furthermore, an additional spatial join was conducted so that each tree record is aggregated into one, and only one, specific street segment, yielding a “Tree count” attribute for each street segment. The “Tree count” attribute was further used to create a “Trees per meter” attribute by dividing the number of trees in a street segment by the segment’s length. “Trees per meter” was used in the basic correlation analysis: “street light poles per meter” (see below) and “trees per meter” to “average illuminance per street segment”.

“Location of street lighting poles” is a vector dataset containing the location of the city’s street lighting poles. It contains 56,422 street lighting poles. A basic spatial join was conducted so that each street lighting pole record is aggregated into one, and only one, specific CDA, yielding a “Pole count” attribute for each CDA. Moreover, an additional spatial join was conducted so that each street lighting pole record is aggregated into one, and only one, specific street segment, yielding a “Pole count” attribute for each street segment. The “Pole count” attribute was further used to create a “Poles per meter” attribute by dividing the number of trees in a street segment by the segment’s length. “Poles per meter” was used in the basic correlation analysis: “street light poles per meter” and “trees per meter” to “average illuminance per street segment”.

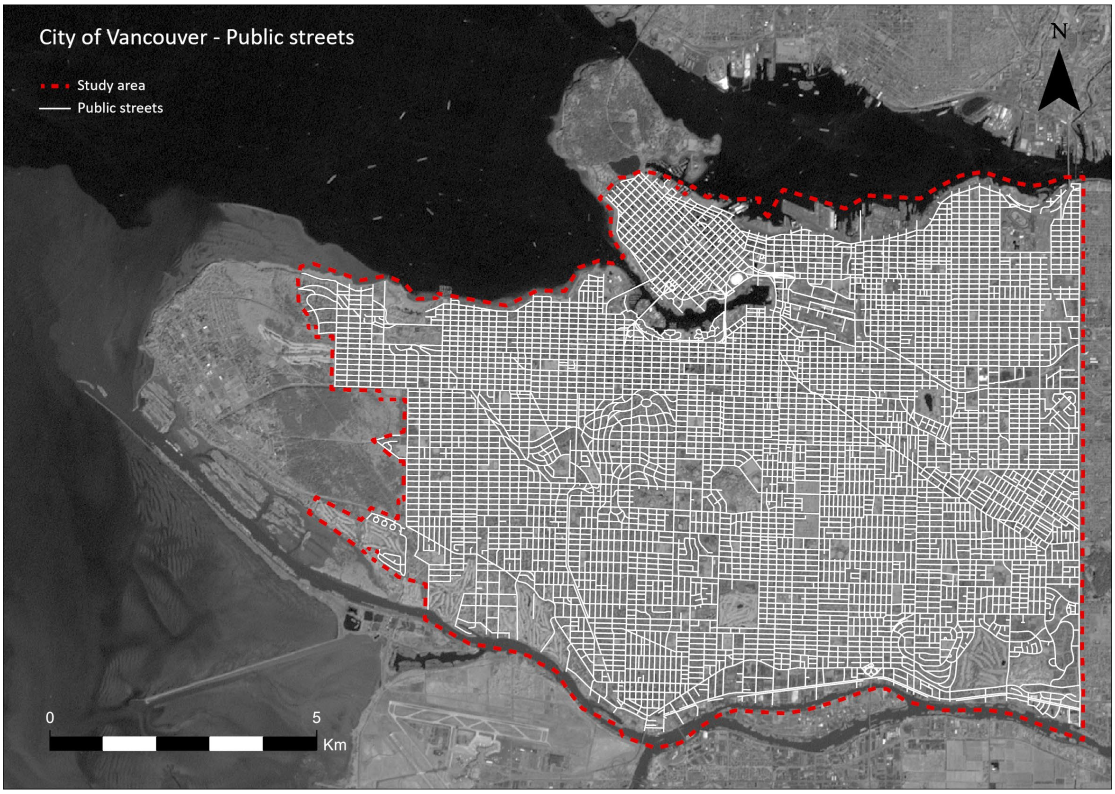

Figure 9 - Study area and public streets network

“Location of rapid transit stations” is a vector dataset containing the location of metro Vancouver rapid transit stations. After selecting stations within the municipality of Vancouver, it contains 22 rapid transit stations. A near table was conducted so that each CDA has a distance to the nearest rapid transit station, yielding a “Distance to rapid transit” attribute for each CDA.

"Location of bus stops" dataset is provided by Translink in WGS 1984 (TransLink (South Coast British Columbia Transportation Authority) and Lesack 2017) and was projected to NAD 1983 UTM Zone 10N coordinate system. "Location of bus stops" is a vector dataset containing the location of metro Vancouver bus stops. After selecting bus stops within the municipality of Vancouver, 1,895 bus stops are accounted for. A basic spatial join was conducted so that each bus stop record is aggregated into one, and only one, specific CDA, yielding a "Bus stop count" attribute for each CDA.

"Location of liquor stores" is a vector dataset provided by the British Columbia Data Catalogue (B.C. Government 2021a) in "comma separated values" format containing the civic addresses of private and provincial government owned liquor stores in the city of Vancouver. It contains the location of 72 liquor stores. The civic address was geocoded with the government of British Columbia Address Geocoder application (B.C. Government 2021b) and the location was added to an ArcGIS Pro map through its X and Y coordinates, and further projected to NAD 1983 UTM Zone 10N coordinate system. A near table was conducted so that each CDA has a distance to the nearest liquor store, yielding a "Distance to liquor store" attribute for each CDA.

"2016 Generalized Land Use Classification" dataset is provided by Metro Vancouver in NAD 1983 UTM Zone 10N coordinate system (Metro Vancouver 2021). "2016 Generalized Land Use Classification" is a vector dataset containing the assigned land use classification of each land parcel in Vancouver's metropolitan region. After selecting those parcels within the municipality of Vancouver and excluding polygons that are assigned a use of "Road Right-of-way", there are 15,942 parcels in the analysis dataset. A summarization was conducted on each CDA, yielding a "commercial land use area", "mixed land use area", and "parking land use area" attribute for each CDA. The summarized "commercial land use area" was divided by the shape area to yield a "commercial land use rate" attribute. The summarized "mixed land use area" was divided by the shape area to yield a "mixed land use rate" attribute. A near table was conducted so that each CDA has a distance to the nearest "parking land use area" (59 areas), yielding a "Distance to parking land use area" attribute for each CDA. These attributes are some of the explanatory variables used in the regression analysis, together with "Tree count", "Pole count", "Distance to rapid transit", "Bus stop count", and "Distance to liquor store".

3.2.3 Crime data

Crime incident data was acquired from the VPD GeoDASH Open Data platform. This dataset includes crime data since 2003, updated on a weekly basis, and it contains the following attributes for crimes reported to the VPD: type, year, month, day, hour, minute, hundred block, neighborhood, X, and Y (Vancouver Police Department 2021b). Crime data for the years 2012 through 2016 was selected for this study as, for the time being, 2016 is the latest census data available.

For the work presented in this thesis, crime data classified as Theft from Vehicle (TFV) was added to an ArcGIS Pro map through its X and Y coordinates, projected to NAD 1983 UTM Zone 10N coordinate system. For property types of crime, VPD provides the location to the hundred block level of the crime and around the general area of the block (Vancouver Police Department 2021b). The level of accuracy provided by the dataset does not impose any downside to this research as TFV crimes are further aggregated to the larger spatial units of analysis such as census dissemination areas or public street segments.

Additionally, TFV crimes were further organized into night-time Theft from Vehicle (NT-TFV) crimes by querying crimes that occurred between sunset and sunrise, according to the first day of each month´s Nautical Twilight start hour and end hour for the year of 2016 (National Research Council Canada 2021). The analysis dataset contained 27,392 records of the study variable (NT-TFV) ranging from January 1st 2012 to Dec 31st 2016. These records were aggregated into two distinct spatial units of analysis: CDAs and street segments.

To aggregate into CDAs, a basic spatial join was conducted so that each record is aggregated into one, and only one, specific CDA. Additionally, a new attribute was created, “NT-TFV per 100 thousand habitants”, by dividing the number of NT-TFV by the number of habitants and multiplying by 100 thousand. This new attribute is the dependent variable used in the regression analysis. To aggregate into street segments a near table was computed for each of the records and exclusively to the closest street segment. The near points were further spatially joined to one, and only one, street segment, yielding a “NT-TFV count” attribute for each street segment. The “NT-TFV count” attribute was further used to create a “NT-TFV per meter” attribute by dividing the number of NT-TFV in a street segment by the segment’s length.

3.2.4 Street illuminance data

Street illuminance data was collected with the SenseBox, a citizen science hardware toolkit. Illuminance data was collected in 1913 points along randomly selected 374 street segments that included residential, arterial and collector streets in varied neighborhoods of the study area. For each measure point, illuminance in Lux is saved with X and Y coordinates. The data was added to an ArcGIS Pro map through its X and Y coordinates and projected to NAD 1983 UTM Zone 10N coordinate system. To aggregate into street segments a near table was computed for each of the illuminance records and exclusively to the closest street segment. The near points were further spatially joined to one, and only one, intersecting street segment, following a "mean" merge rule. This yielded an average measured illuminance for each of the sampled street segments. The sampled illuminance is used in the basic correlation analysis: "street light poles per meter" and "trees per meter" to "average illuminance per street segment".

3.3 Methodology

The research will be conducted as a case study in Vancouver, BC, Canada with two parts. Part one, pertaining to NT-TFV, has four phases: data collection and processing, exploratory spatial data analysis, regression modeling, and analysis. Part two, the relationship between pole density and illuminance, has three phases: assembly of the sensing apparatus, illuminance collection, and analysis. The following sections will detail the methodology for each of the objectives.

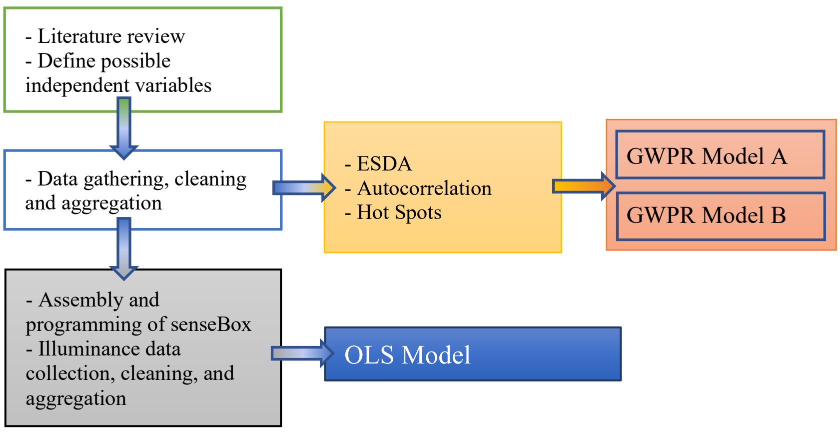

Figure 10 - Methodology flowchart

3.3.1 Spatial analysis and modelling of NT-TFV

The first part, the analysis of the distribution of NT-TFV is built around CDAs, the smallest unit where Census data is available. Some maps are shown using the street segment or a Kernel density as the unit of analysis just as an illustration of how NT-TFV is distributed in the study area when using those spatial units. Data was acquired either in that scale of analysis or spatially joined into it.

A set of techniques called Exploratory Spatial Data Analysis (ESDA) is applied to visualize the spatial distribution, identify spatial outliers and uncover patterns of spatial association, clusters, or hot spots. The set of techniques includes spatial autocorrelation analysis, the Global Moran's I Index, and the occurrence (or not) of a pattern is revealed by the co-incidence of similar values. The Local Indicators of Spatial Association (LISA) methods, proposed by Anselin (1995), are used to identify local crime clusters (positive local autocorrelation) or local outliers (negative local autocorrelation) with the Local Moran's I Index. High-high clusters represent hot spots with high NT-TFV counts that are surrounded by other high NT-TFV counts, while low-low clusters represent cold spots where low counts of NT-TFV are surrounded by other low counts of NT-TFV. Also, regions with high rates of NT-TFV crime that are surrounded by low rates of NT-TFV are high-low cluster, and the inverse are low-high cluster (Anselin et al. 2000). Getis and Ord (1992) proposed a method of spatial association, the G statistic, to analyze patterns at a local scale for non-stationary processes, that is widely used in spatial analysis. The Getis-Ord Gi*, similarly to the Local Moran's I statistic, assess the presence of clusters with similarly high values (hot spots) or low values (cold spots).

Independent variables for the regression analysis were selected based on social disorganization or routine activity theories (such as income and immigration status), allied with some physical street attributes that can influence victims and/or offenders conduct in a micro place (light poles and street trees), or that can act as crime attractors or generators (such as transit stations and land use). The intention was to keep the number of independent variables around eight according to the principle of parsimony. Initially 18 possible independent variables were tested using the exploratory regression tool in ArcGIS Pro. This tool evaluates possible combinations of independent variables to develop appropriate OLS models, looking for Adjusted R^2, VIF values and Jarque-Bera's test p-values, also running the Spatial Autocorrelation (Global Moran's I) tool on the residuals (ESRI 2021). Although based on a global model, this tool provided a first glimpse on the possible explanatory variable's significance and possible multicollinearity issues. OLS is a simple method but it does not account for spatial autocorrelation, and crime is known for having a spillover effect to surrounding areas and also be affected by the specific area and by adjacent area´s characteristics (Lin et al. 2021).

Variable	Description	Source
drive_pop_rate	Commuter driver to population rate	(Statistics Canada 2017)
near_parking	Distance to parking land use area	(Metro Vancouver 2021)
near_rap_trans	Distance to rapid transit	(City of Vancouver 2021)
poles_count	Pole count	(City of Vancouver 2021)
poles_rate_sqkm	Poles per sq. Km	(City of Vancouver 2021)
liquor_dist	Distance to liquor store	(B.C. Government 2021a)
pop_dens	Population density	(Statistics Canada 2017)
pop	Population	(Statistics Canada 2017)
bus_stops	Bus stop count	(TransLink and Lesack 2017)
rec_imm_rate	Recent immigration rate	(Statistics Canada 2017)
trees_count	Tree count	(City of Vancouver 2021)
trees_rate_sqkm	Trees per sq. Km	(City of Vancouver 2021)
mix_landuse_rate	Mixed land use rate	(Metro Vancouver 2021)
comm_landuse_rate	Commercial land use rate	(Metro Vancouver 2021)
labor_empy	Employed labor force	(Statistics Canada 2017)
value_med	Median value of dwellings	(Statistics Canada 2017)
rented_rate	Rate of rented private households	(Statistics Canada 2017)
med_tot_in	Median total income	(Statistics Canada 2017)

Table 2 - Independent variables tested using the exploratory regression

Following the exploratory regression, a GWPR analysis looked for spatial variations in the relationship between the NT-TFV and the explanatory variables. The basic concept is that relationships of the dependent variable with demographic (and other) characteristics may be different between locations, and this heterogeneity should be investigated. Osgood (2000) suggested a novel approach to the problem of analyzing aggregate crime rates based on low population and low base rates, the Poisson regression model. The issue is one that, in OLS, crime rates that involve small counts of crimes will have larger errors of predictions in crime rates based on small populations than in those based on large populations, as the precision of the estimates depend on population size. Aggregate units that, like in this study, have a mix of population sizes across the study area will tend to violate the assumption of homogeneous error variance. Andresen and Ha (2020) reviewed GWR applications in criminology concluding, from the different studies analyzed, that many explanatory variables have relationships that vary in space, and that spatial processes are of great relevance, thus important to guide prevention policies.

The GWPR models estimated in this research used an adaptive bi-square kernel with a “user defined” bandwidth of 42 neighbors, considering the lowest Corrected Akaike Information Criterion (AICc). The dependent variable was “NT-TFV per 100 thousand inhabitants” for both models. Model A features “Recent immigration rate”, “Commuter driver to population rate”, “Median total income”, “Tree count”, “Pole count”, “Distance to rapid transit”, “Bus stop count”, and “Commercial land use rate”. Model B features “Tree count”, “Pole count”, “Distance to rapid transit”, “Bus stop count”, “Distance to liquor store”, and “Commercial land use rate”, “Mixed land use rate”, and “Distance to parking land use area”.

3.3.2 Data collection and modelling of Illuminance

In the second part, the analysis of relationship between two physical attributes, street lighting poles and tree density, and measured night-time street illuminance is built around street segments. Data was acquired in vector points and spatially joined into street segments.

The senseBox is an open-source hardware toolkit for building environmental monitoring devices that has been used for digital education, in the context of citizen science and for professional environmental data collection. As web-based companions to the senseBox, the Blockly environment is used to program the board, and the openSenseMap data portal is used to visualize and store the collected data (Pesch and Bartoschek, 2019). SenseBox is based on the combination of sensors, connection components (XBees), and accessories that are arranged around a micro controller unit (MCU) to create an environmental measuring station, and other types of sensing devices. The Arduino compatible MCU contains a processor and an array of connections that will receive sensors and XBees. There is a large array of sensors available that include temperature, air pressure, humidity, illuminance, UV, particulate matter, CO_2, and others. The XBees connect the senseBox to a means to save data such as an SD card to save data locally, or to the internet to save data on the web. Each project requires a specific arrangement of sensors, XBees and accessories depending on the specific nature of what is being measured and how the collected data is to be stored or visualized, all that allied to a specific Arduino code.

For this research the objective was to devise a simple illuminance collecting unit that could be used while driving a car. The idea was to mount the unit right on the car top and to be able to collect data on the go. It is a simple experiment that will provide a basic understanding about street illuminance on selected street segments of the study area and could, in future work, be refined to offer better reliability in other scientific experiments. The senseBox-based georeferenced mobile illuminance collecting station used for the illuminance collection in this study has the following hardware configuration (Figure 11):

- A microcontroller: senseBox MCU, based on the ARM Cortex-M0+ processor. Available interfaces include I2C, UART and digital I/Os with a robust JST connector system and XBee compatible sockets.
- GPS module: senseBox GPS module, based on the u-blox CAM-M8Q Multi GNSS module.
- Port for miniSD card
- The light intensity sensor: AMS-TAOS TSL45315 sensor
- Display: senseBox OLED display
- Power bank and cable

The unit was programmed using the Blockly for senseBox graphical programming environment. The web-based interface allows for programming sketches and compiling them into machine language to later be transferred to the MCU. The Blockly sketch and associated Arduino source code are available in Appendix 1: Blockly assembly and Arduino source code.

After some tests, the mobile station components were assembled in a way that minimizes interferences between components, specially to the GNSS module, and to provide for a simple operation. After the senseBox station is assembled, the street illuminance collection has the following steps: bring the unit outdoors, connect to power bank, mount the unit on specific position on vehicle top (Figure 12), wait for a FixType (quality of signal) of level 3 (3D: latitude, longitude and altitude), and, lastly, flood the light intensity sensor with a high beam of light (to mark the start of the illuminance collection on the CSV data file). Illuminance and position were constantly sensed and saved every two seconds.

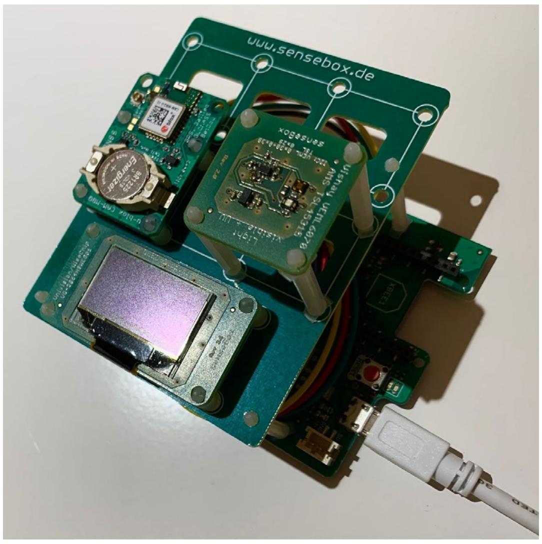

Figure 11 - Components of sensing unit

Figure 12 - Sensing unit mounted in car top

The chosen method to analyze the relationship between street lighting poles and tree density and measured night-time street illuminance is the OLS regression. For this analysis, a simple linear regression method is adequate as there is no expectation of nonstationary. The dependent variable is the "average illuminance per street segment" and the explanatory variables are "street light poles per meter" and "trees per meter". The expectation is that "street light poles per meter" should have a somehow strong and positive relationship and that "trees per meter" should have a moderate negative relationship with the dependent variable.

4. RESULTS

The following sections will present the results of the spatial analysis and modelling of NT-TFV and modelling of illuminance.

4.1 Exploratory spatial data analysis

Figure A 1 in the appendix shows the spatial distribution of NT-TFV crime in Vancouver's CDAs between 2012 and 2016. Higher counts are concentrated in the downtown area and disperse into lower counts, in an almost radial spreading, towards the East, West and South edges of the study area. Similarly, Figure 13 shows the spatial distribution of NT-TFV crime per 100 thousand habitants in Vancouver's CDAs between 2012 and 2016. The rates are higher in the downtown area and adjacent neighborhoods of Strathcona and Mount Pleasant. To the East, a high rate CDA can be seen in the Renfrew - Collingwood neighborhood, and a series CDAs with average and somewhat above average rates of NT-TFV can be see across the study area, especially on the West side of the city.

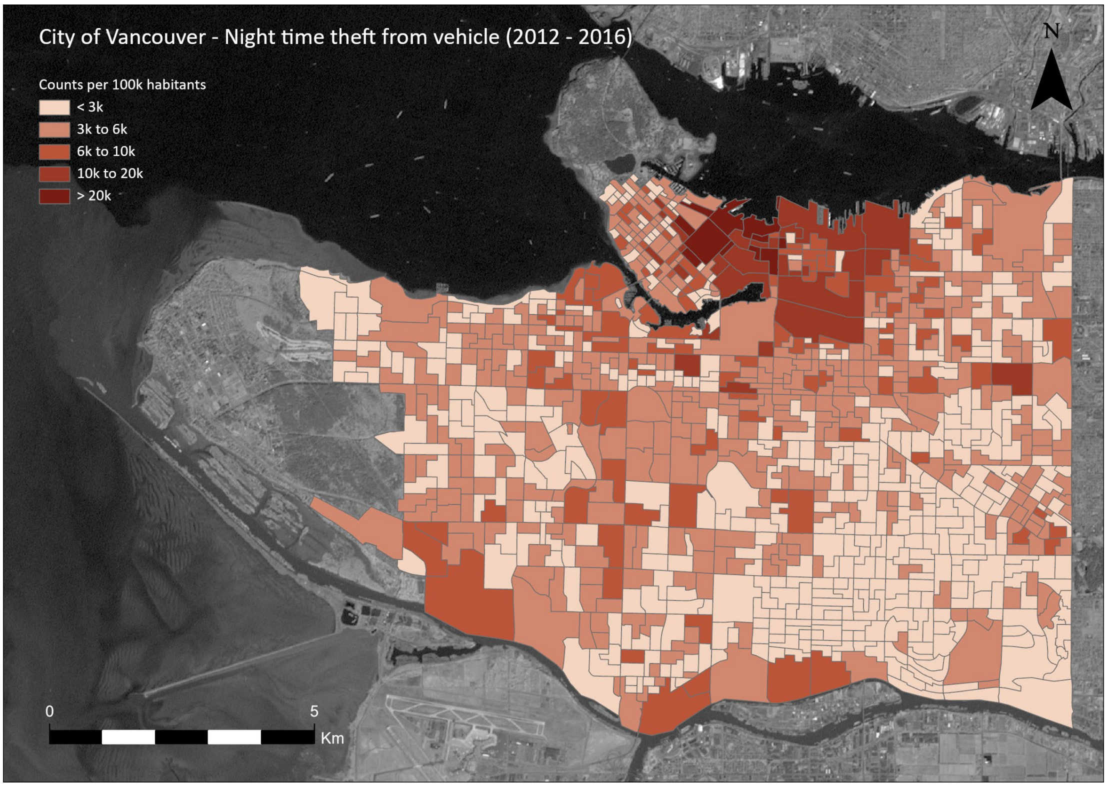

Figure 13 - Spatial distribution of NT-TFV crime per 100 thousand inhabitants, 2012 to 2016

Between 2012 and 2016, 27392 counts of NT-TFV were reported to the VPD, an average of 27.55 per CDA. Both CDAs with the highest count and highest ratio are located in the downtown area of the city.

Figure A 2 in the appendix presents the histogram of NT-TFV crime rate per 100 thousand inhabitants per CDA from 2012 to 2016. The histogram shows an asymmetric, right tail, positively skewed (skewness of 6.47), thin and high peak leptokurtic (kurtosis of 57.27) distribution with an average of 4039 NT-TFV crimes per 100k habitants. The majority of the CDAs, 664 out of 994, have below mean NT-TFV rates, and its right long tail indicates the presence of outliers.

Figure 14 shows the spatial distribution of NT-TFV crime rate in Vancouver's public streets. Higher counts per meter of street segment are concentrated in the downtown area and disperse towards the municipal edges. An average of 1.6 NT-TFV crimes per street segment were reported between 2012 and 2016, with an average of 0.017 NT-TFV crimes per meter of street segment. The segment with the most reported crimes in the period had 218 NT-TFV crimes with 1.32 NT-TFV crimes per meter.

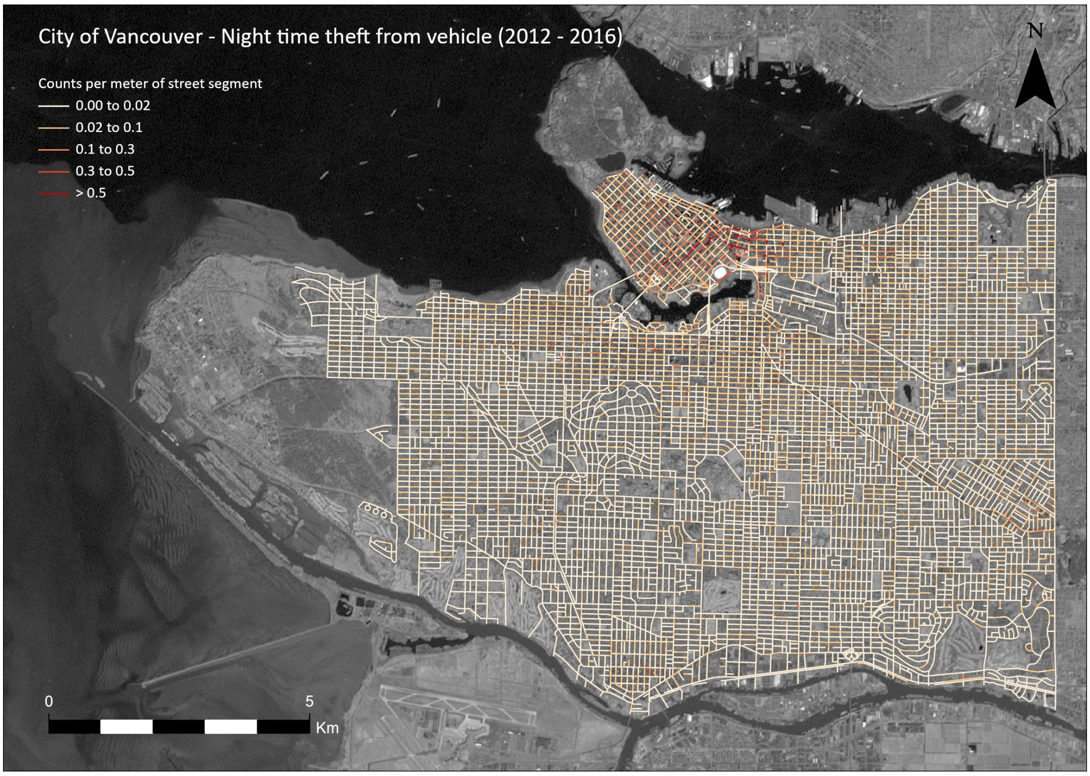

Figure 14 - Spatial distribution of NT-TFV crime rate per meter of street segment, 2012 to 2016

Figure 15 shows how different types of streets (selected) compare regarding NT-TFV crimes. Arterial types of streets had a much higher ratio of crime per street segment than residential and secondary arterial streets.

Figure 16 shows the Kernel density of NT-TFV points. The concentration is very high in the downtown area and its adjacent neighborhoods, in line with the spatial distributions previously shown.

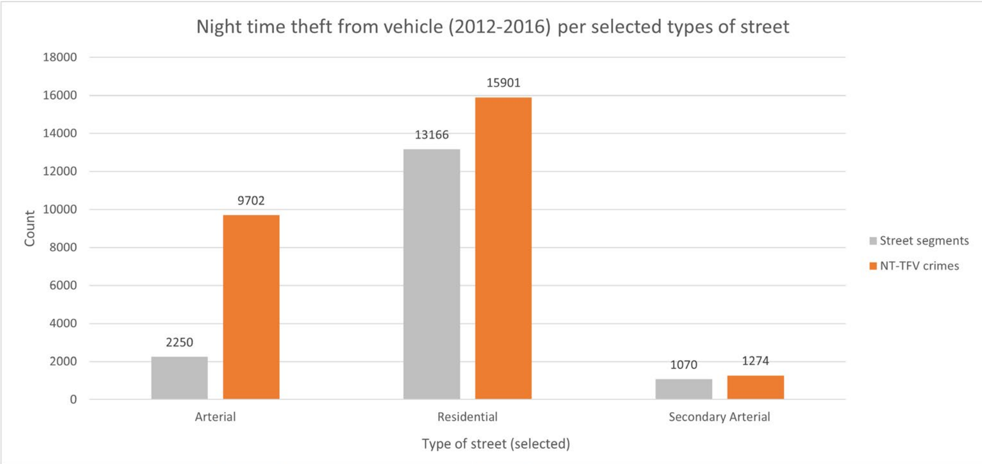

Figure 15 - NT-TFV crime per selected street segment, from 2012 to 2016

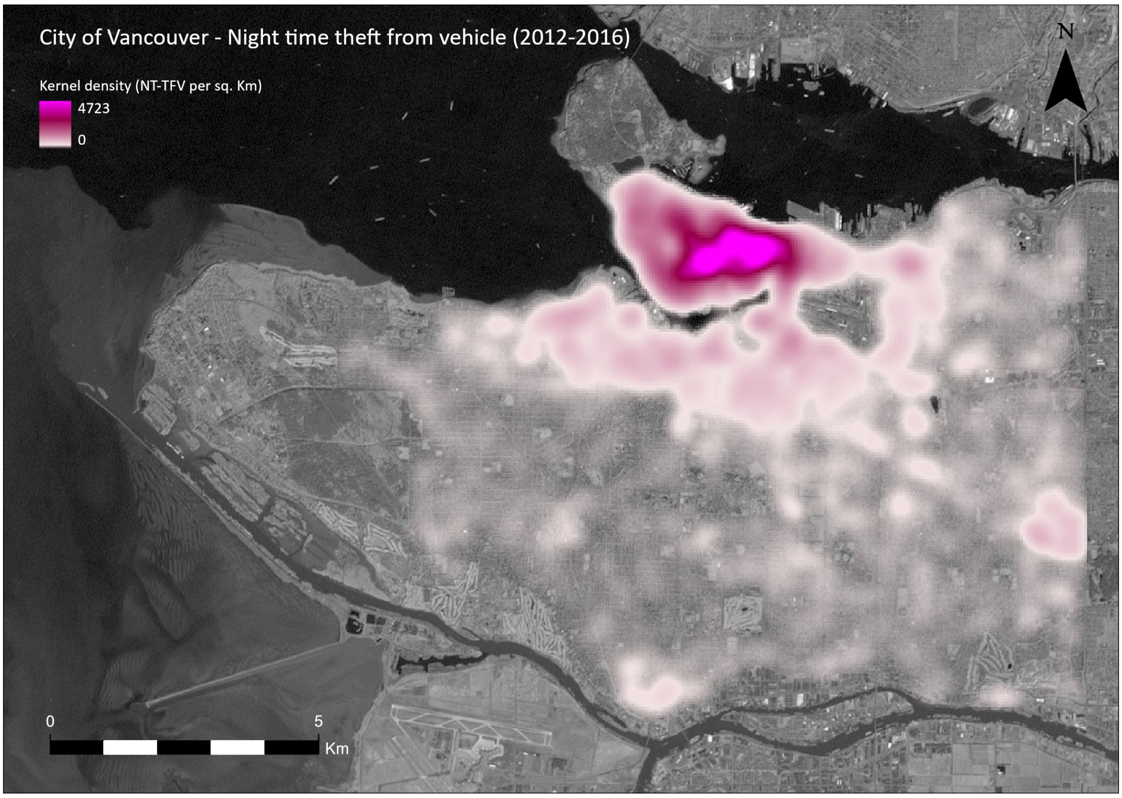

Figure 16 - Kernel density of NT-TFV crime per Km2 from 2012 to 2016

Table 3 presents the descriptive statistics for the variables used in both regression models, for the dependent variable, and for NT-TFV crime per CDAs. The spatial distribution maps are presented on Appendix 2: Maps.

Variable	Mean	Min.	Max.	Std. Deviation
NT-TFV rate (NT-TFV crimes per 100 thousnad inhabitants)	4039.48	0	54411.70	4388.18
Street trees (count)	125.47	0	1050	84.30
Street lighting poles (count)	56.76	5	1127	55.94
Commuter drivers rate (drivers per inhabitant)	0.2296	0	0.509	0.070
Recent immigrant rate (recent immigrants per inhabitant)	0.0585	0	0.247	0.041
Median total income (Canadian dollar)	33355	0	79872	10488
Bus stops (count)	1.90	0	42	2.57
Distance to nearest rapid transit (meters)	1606	0	6565	1375.80
Commercial land use rate (sq meters of use per total area)	0.031	0	0.68	0.079
Mixed land use rate (sq meters of use per total area)	0.025	0	0.57	0.048
Distance to nearest liquor store (meters)	399	0	1798	317.50

Table 3 - Descriptive statistics of the variables used in the regression analysis

4.1.1 Spatial autocorrelation

The spatial autocorrelation for both NT-TFV rate per CDA and for NT-TFV per meter of street segment, was tested using the Global Moran's I with, respectively, 0.3392 and 0.2474 indexes. With both p-values of zero, we can reject the null hypothesis that both NT-TFV rate and per meter are randomly distributed, and we can assume positive spatial autocorrelation. Both CDAs and street segments with high NT-TFV crimes tend to be surrounded by other similarly high NT-TFV CDAs and street segments, and the same spatial process happens with low NT-TFV CDAs and street segments.

A Local Moran's I test was pursued, in order to provide local indexes that could reveal specific local patterns or spatial regimes, for both aggregated data types, CDAs (Figure 17) and street segments (Figure 18). The tests (with a 95 percent confidence level) reveal clusters of high values, low values, and outliers for both aggregation types.

Clusters of high values can be seen in the downtown area of the city, in line with the distributions and the Kernel density previously shown. This is an area that concentrates both high CDA rates of NT-TFV and high thefts per meter of street segment. The different aggregation types expose a different pattern, expected from a different scale of analysis. While the CDA analysis portraits high clusters gathered around the downtown, with just a few CDAs outside that neighborhood, the street segment analysis presents an expanded high cluster area that goes well beyond the downtown area. The street segment analysis also reveals a much more outlier-permeated pattern, denoting that not all street segment in that neighborhood endure high levels of NT-TFV per meter of street segment. A small area in south-east is also considered a cluster of high values in the street segment aggregation but not significant in the CDA aggregation.

Clusters of low values are concentrated in the south-east region of the city, especially in the Sunset, Victoria-Fraserview and Killarney neighborhoods, when analyzing the CDAs aggregation, while a much larger radial area is significantly considered a low cluster in the street segment analysis. Again, outliers are sparsely around the low value cluster of CDAs while the street segment low clusters have a much-infiltrated pattern, with many outliers across the entire cluster area.

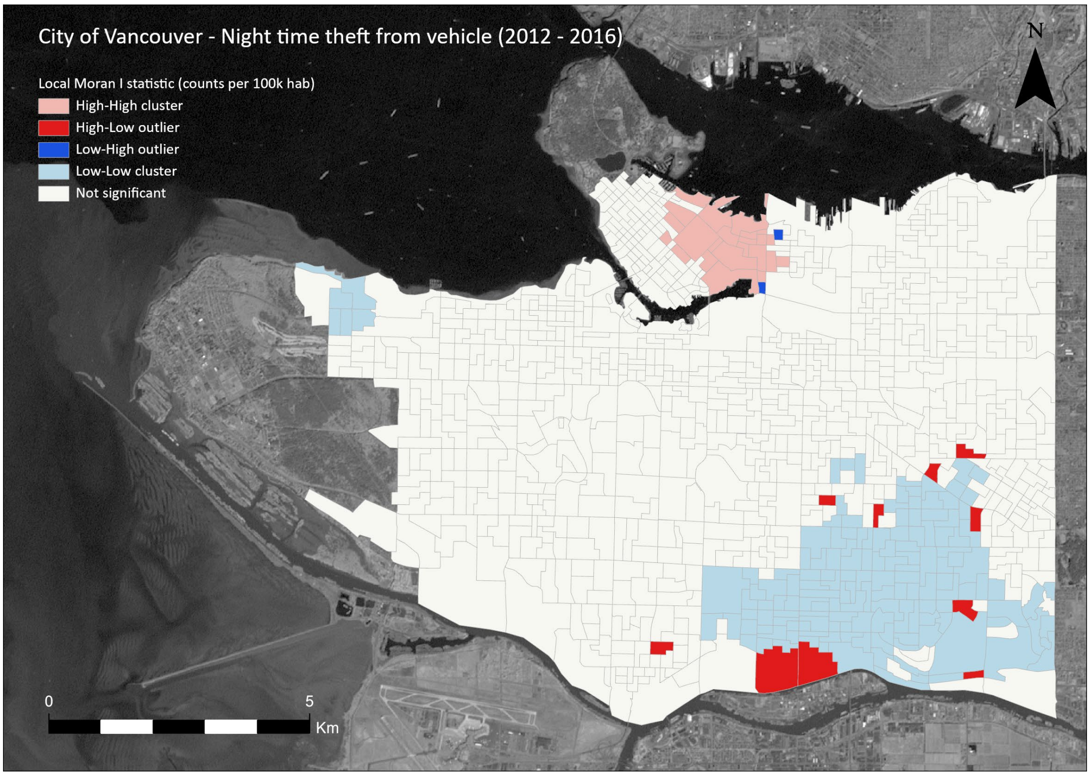

Figure 17 - Local Moran's I clusters and outliers for NT-TFV crime rate per 100k inhabitants

Finally, the areas where results are not significant differ between the two aggregation types. While CDAs aggregation shows a vast area where results were not significant, the street segment aggregation shows a radial belt between the high and low clusters as non-significant. Nevertheless, the analysis of street segments are interesting as high clusters are located in regions of high population density and high commercial and mixed land use, while low clusters are in more low population density residential areas.

The results of the clusters and outliers analysis are in line with the previously shown distribution of NT-TFV crimes.

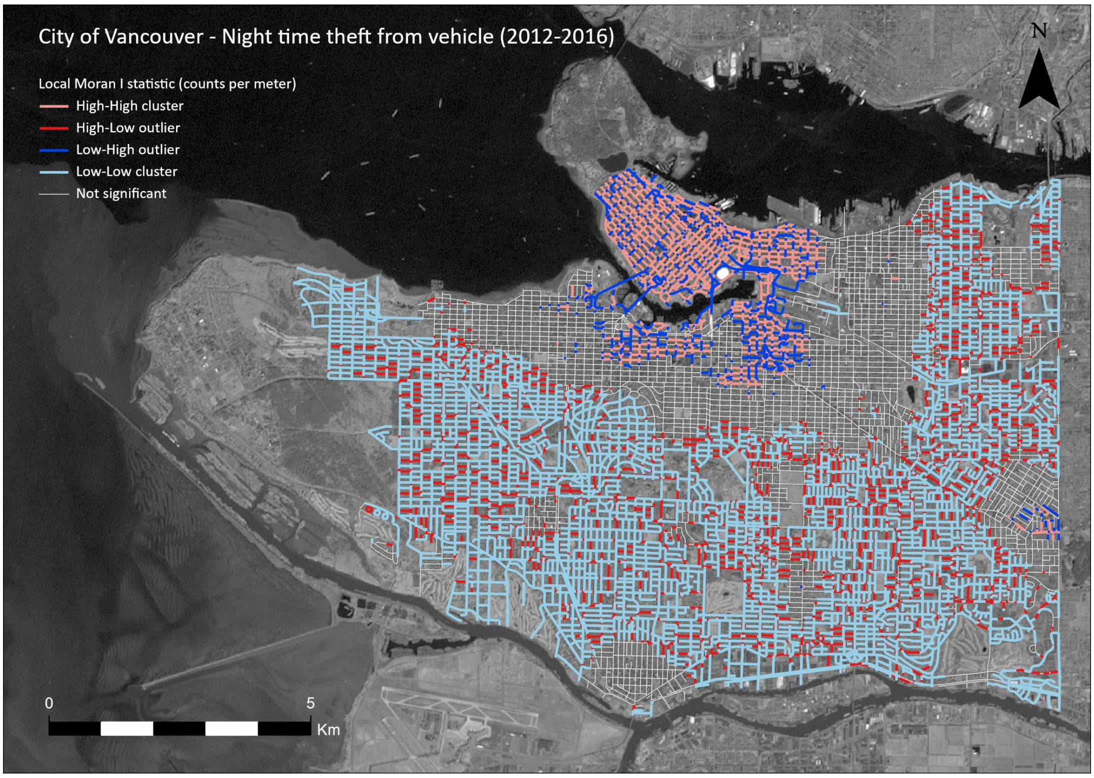

Figure 18 - Local Moran's I clusters and outliers for NT-TFV crimes per meter of street segment

4.1.2 Hot spots

Hot spot analysis allows for the identification of statistically significant hot spots (clusters of high values) and cold spots (clusters of low values) using the Getis-Ord Gi* statistic. Hot spots for the CDAs aggregation type (Figure 19) can be seen in the downtown area of the city, in line with the NT-TFV distributions, the Kernel density and the Local Moran's I test previously shown. Cold spots are concentrated in the south-east region of the city, in Sunset, and Victoria-Fraserview neighborhoods.

Hot spots for the street segment type of aggregation (Figure 20) follow a similar pattern as the street segment Local Moran's I, with an expanded Hot Spot area that goes beyond downtown and a much larger Cold Spot area. Street segments where the statistic is not significant can be found in a radial belt between the Hot and Cold Spots.

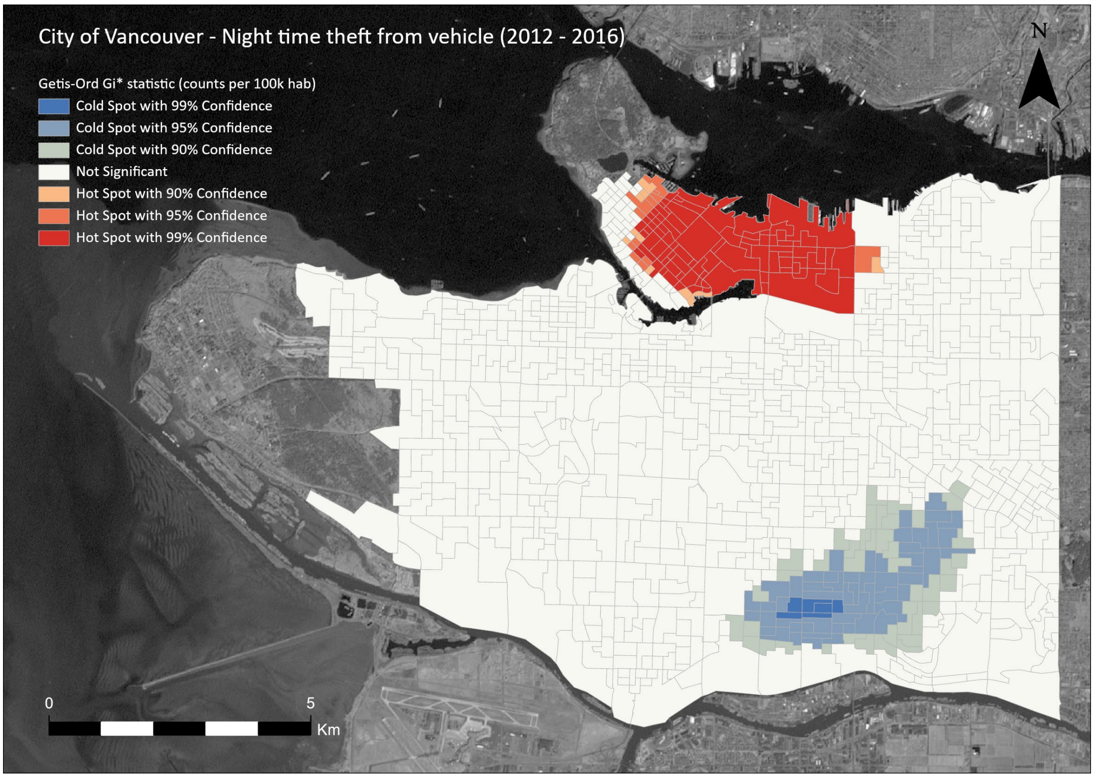

Figure 19 - Hot spot analysis for NT-TFV crime rate per 100k inhabitants

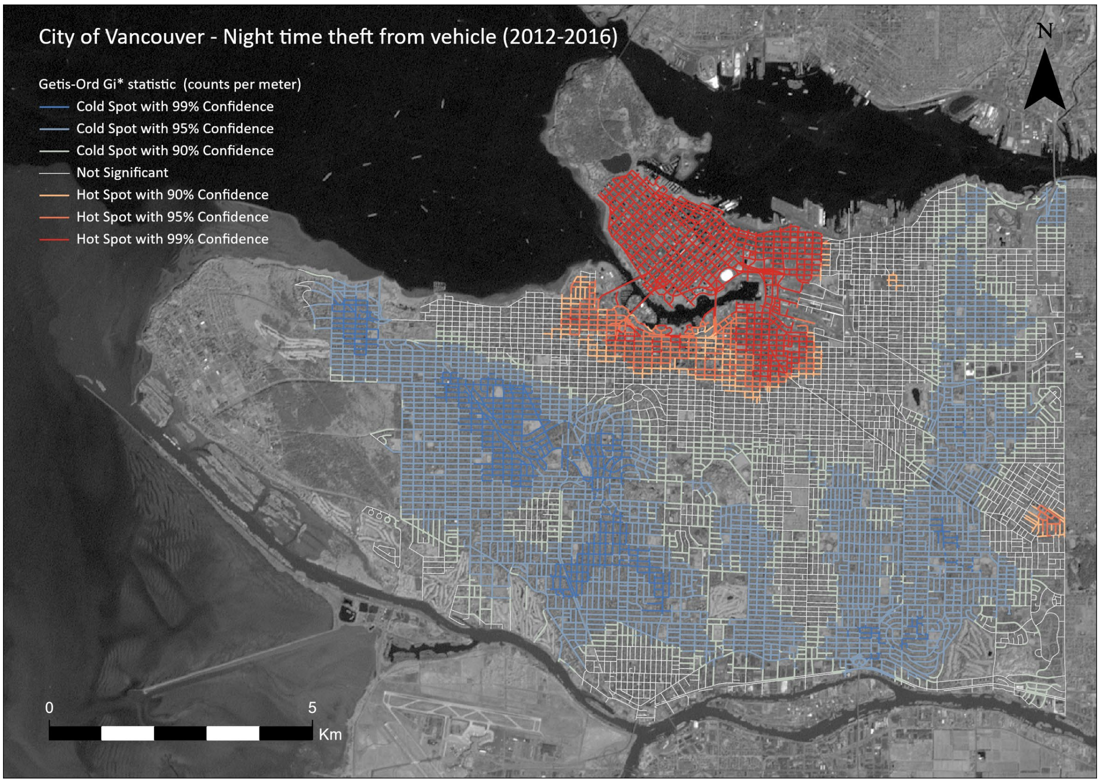

Figure 20 - Hot spot analysis for NT-TFV crimes per meter of street segment

4.2 Geographically Weighted Poisson Regression (GWPR) – Model A

The GWPR model A has "NT-TFV per 100 thousand inhabitants" as the dependent variable and features the following independent variables: "Recent immigration rate", "Commuter driver to population rate", "Median total income", "Tree count", "Pole count", "Distance to rapid transit", "Bus stop count", and "Commercial land use rate". This model is used to explore how the relationship between the crimes and the selected explanatory variables varies within the study area. Model A diagnostics (Table 4) show an 84.27% deviance explained by the local model, against only 35.88% deviance explained by the global model.

Model A Diagnostics	
Deviance explained by the global model (non-spatial)	0.3588
Deviance explained by the local model	0.8427
Deviance explained by the local model vs. global model	0.7547
AICc	393878.5
Sigma-Squared	72484842.6
Sigma-Squared MLE	35485016.8
Effective Degrees of Freedom	486.6

Table 4 - GWPR Model A diagnostics

The map of the local percentage deviance explained (Figure 21) shows a rather heterogenous distribution of explanatory power throughout the study area. The neighborhoods of Downtown and Sunset concentrate the CDAs where percentages of deviance explained are higher than 90%, while Oakridge and Renfrew-Collingwood join the two previous neighborhoods with the concentration of CDAs where percentages of deviance explained are higher than 80%. Interestingly, CDAs where percentages of deviance explained are lower than 30% are all concentrated in the West End neighborhood, a neighborhood adjacent to Downtown.

The next set of maps will show the distribution of GWPR model A local coefficients for each independent variable. Maps depicting the distribution of the independent variables can be found in Appendix 2: Maps. Since, in GWR, there are no statistical significance tests for local coefficients, pseudo-t values were used for that matter (Nakaya et al. 2005). Local coefficients that may be significant at the 5% level are shown ($t < -1.96$ or $t > 1.96$) and omitted otherwise ($-1.96 < t < 1.96$). The models reveal the spatially varying influence of each explanatory variable on NT-TFV rates, with associations that fluctuate from negative to positive depending on the area of the city, demonstrating the existence of nonstationarity of the local coefficients. The light to dark blue colors indicates that the variable has an increasing positive influence on NT-TFV crime (tendency to vary in the same direction), while the yellow-orange-red range indicates an increasing negative influence (tendency to vary in opposite directions).

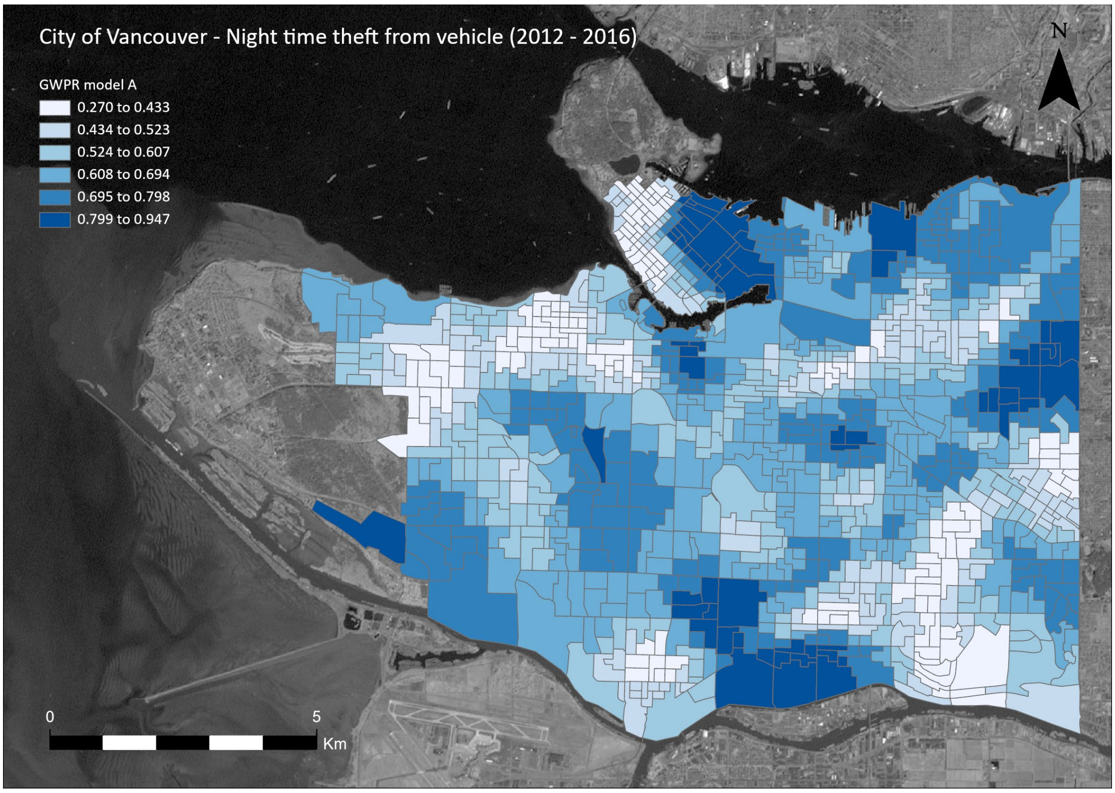

Figure 21 – GWPR model A local percentage deviance explained

Figure 22 shows the local coefficients of the "commuter driver rate" variable. As a measure of the number of inhabitants that commute to work on a private vehicle, it can reflect the availability of targets for NT-TFV crimes. In that sense the expectation is that a positive association would take place. Downtown, the area with most NT-TFV crimes has the lowest rates of commuter drivers in the study area and the model shows a strong negative impact from this variable. This could be related to the fact that the area has a high percentage of condominiums, usually featuring better guardianship capabilities such as private garages with surveillance, making the already low concentration of vehicles being driven to work not available to become targets as if they would be if parked in public streets. Other areas with a negative association are adjacent to the commercial corridor of Kingsway Street, towards the East of the study area. The neighborhood of West End, with a much lower rate of condominiums, shows a positive impact from this variable, as well as adjacent neighborhoods such as Kitsilano and Mount Pleasant. Coefficients are distributed in a very fragmented manner across the rest of the study area, in a mix of positive and negative mild associations.

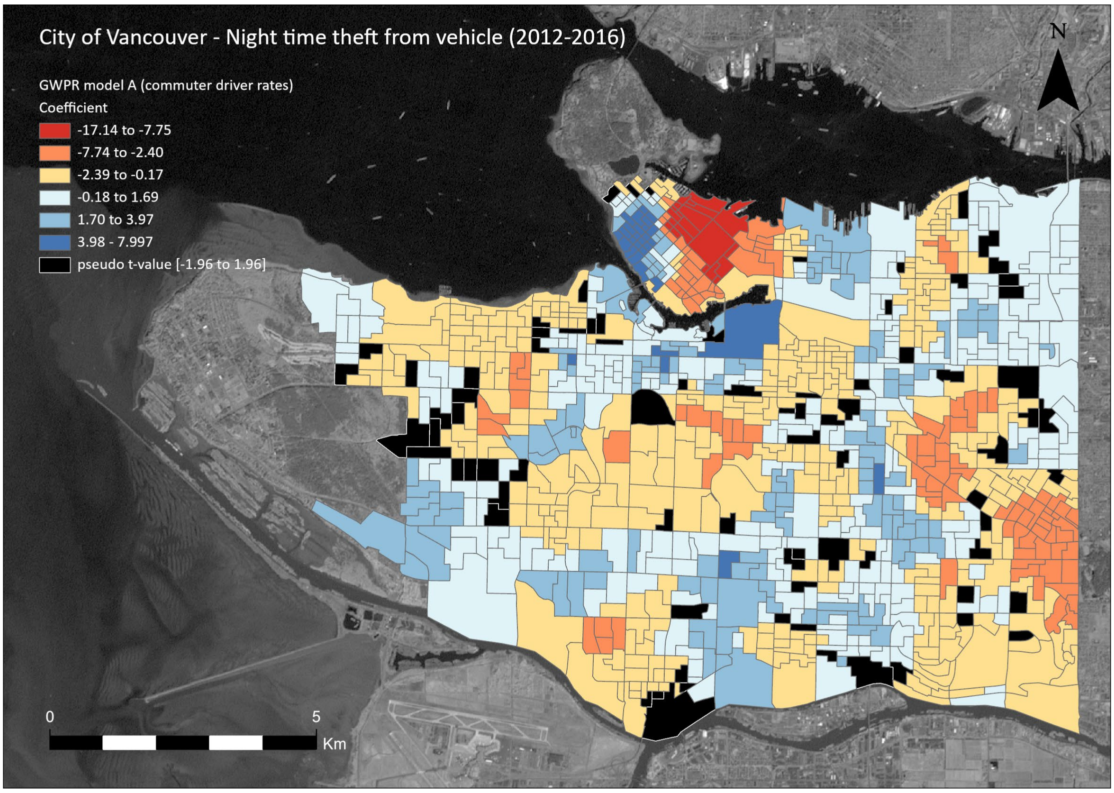

Figure 22 - "commuter driver rate" model A local coefficients and distribution

Figure 23 shows the local coefficients of the "recent immigration rate" variable. As a measure of the number of individuals in private households where the period of immigration was between 2011 and 2016, it is expected to have a positive association with TFV crimes from the social disorganization theory and from previous studies (Andresen and Ha 2020). The most observable association can be seen about Strathcona, just East of Downtown, with a strong negative association between the variables, against what is expected. As with the previous variable, coefficients are also distributed in a very uneven manner through the rest of the study area, in a mix of positive and negative mild associations and some pockets of strong positive association in Kitsilano, Renfrew-Collingwood and bordering areas in the North and South of the study area.

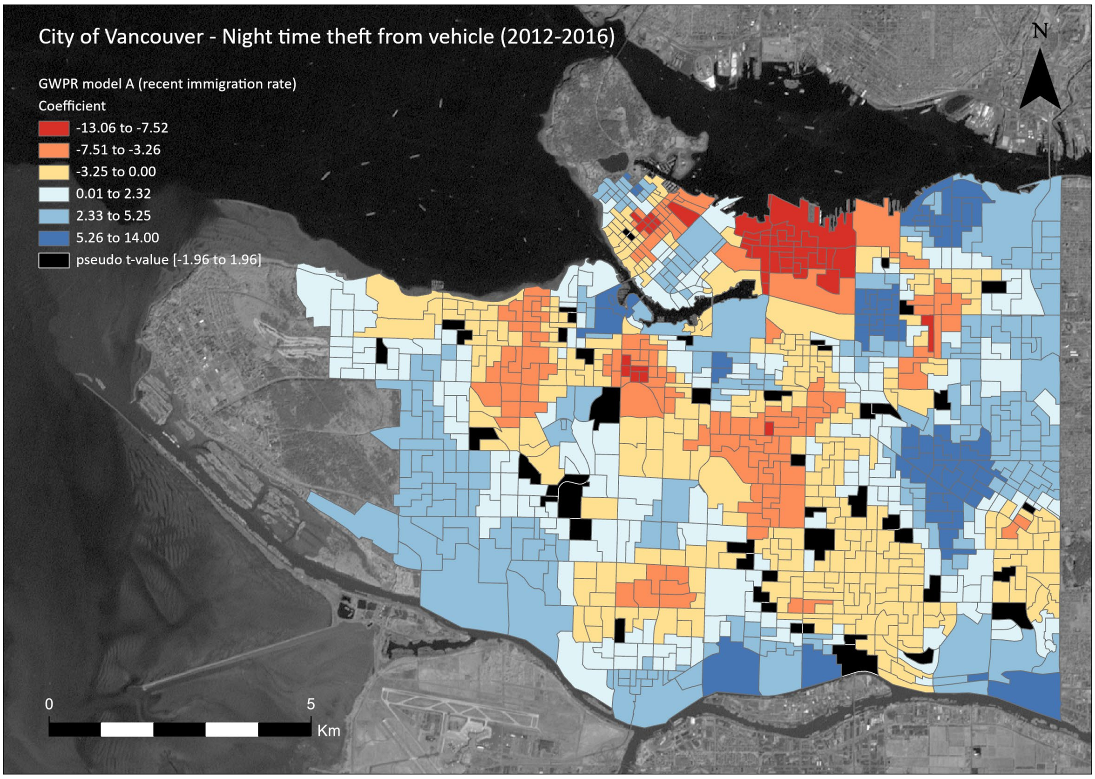

Figure 23 - "recent immigration rate" model A local coefficients and distribution

Figure 24 shows the local coefficients of the "total median income" variable. As a measure of economic wealth, it is expected to have a negative association with crime from the social disorganization theory (less need to commit crimes) and a positive association from the routine activities theory (better target value). Most of the study area shows a weak, either positive or negative, association between income and NT-TFV, with the strongest positive associations on some East neighborhoods and the strongest negative associations in Sunset.

Figure 25 shows the local coefficients of the "commercial land use rate" variable. As a measure of ambient population during business hours, it can reflect the availability of targets for NT-TFV crimes, from the idea of crime attractors and generators, but it can also act as an indirect measure for better guardianship, both from people walking by and from retailers' surveillance mechanisms. In that sense the expectation is that either a positive association (target availability) or negative (guardianship) would take place. Most of the study area shows a mild to strong negative association between commercial land use rates and NT-TFV crimes. In areas with high commercial land use rates, the association tends to be gently negative in most areas and gently positive in other areas.

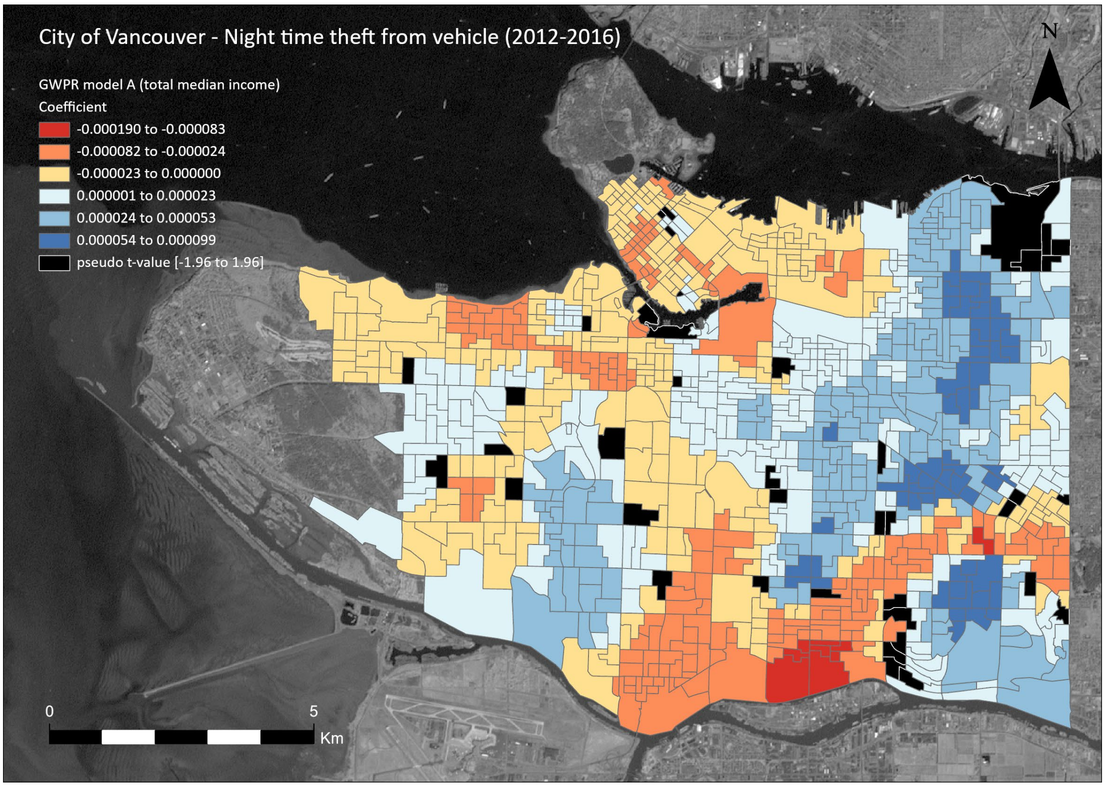

Figure 24 - "total median income" model A local coefficients and distribution

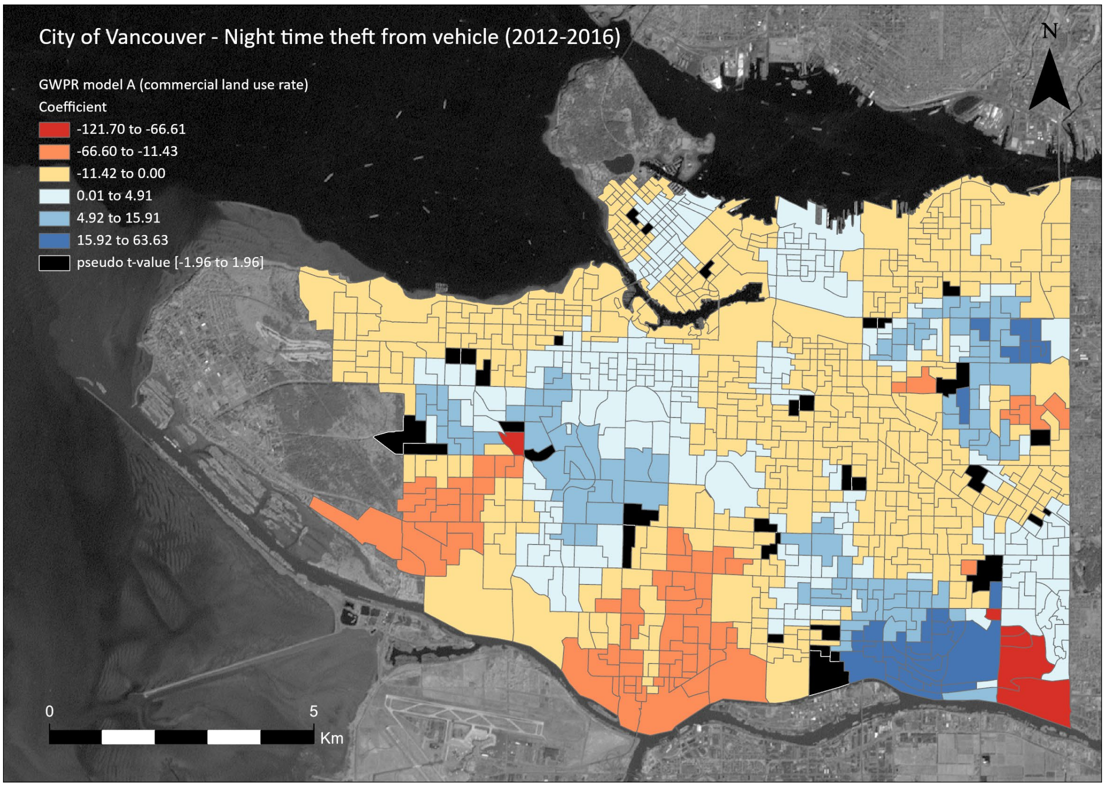

Figure 25 - "commercial land use rate" model A local coefficients and distribution

Figure 26 shows the local coefficients of the “distance to rapid transit” variable. As a measure of how close a rapid transit station is, it reflects the potential of a criminal to reach a certain target, exposing new regions to their actions through an easier “journey to crime” (Liggett et al. 2003). It can also suggest increased population density and commercial land use rates, where people and vehicles tend to gather around, acting as a crime attractor in one hand and increased guardianship in another. On the other hand, Block and Davis (1996) concluded that in areas with high rates of robbery, crime was dispersed along main commercial streets, whereas areas with lower rates of robbery had crime concentrated near rapid transit stations, while Sypion-Dutkowska and Leitner (2017) found that, specifically for car crimes, bus and tram stations detract car crimes. The association is expected to have a mix of positive and negative associations across the study area. Most of the study area shows a mild negative association between rapid transit stations and NT-TFV crimes. Areas showing a positive association are quite weak in strength and scattered across the study area but with a distinguished concentration around some transit stations around Mount Pleasant and Renfrew-Collingwood. The strongest association, negative, is in downtown, where the largest number of stations are found and the smallest rate of commuter drivers are concentrated.

Figure 27 shows the local coefficients of the “count of bus stops” variable. As with rapid transit stations, it reflects the potential of a criminal to reach a certain target but it is not necessarily associated with an increased population density or commercial land use rates. Bus routes are distributed across the entire study area, usually associated with arterial or secondary arterial street types, with bus stops consistently distributed along those routes, associated with all types of neighborhoods and land uses, while rapid transit stations are usually associated with higher rates of commercial and mixed land use. Studies have suggested that bus stops, especially when linked with certain land uses (see Yu 2009; Stucky and Smith 2017) can have a positive association on crime. Coefficients are distributed in a very uneven manner through the study area, mostly in a mix of positive and negative weak to mild associations. There are some noticeable areas of strong positive association, West End, Kitsilano/Fairview and Sunset, and strong negative associations in Strathcona and Riley Park/Kensington-Cedar Cottage.

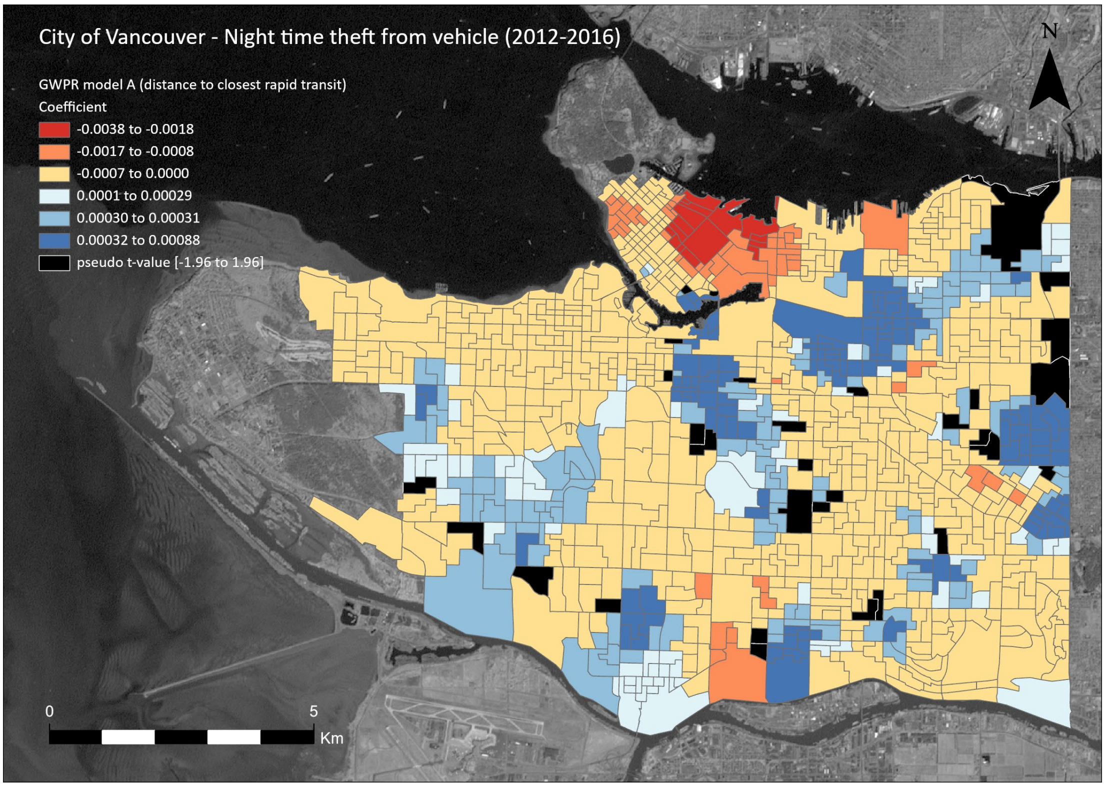

Figure 26 - "distance to closest rapid transit" model A local coefficients and distribution

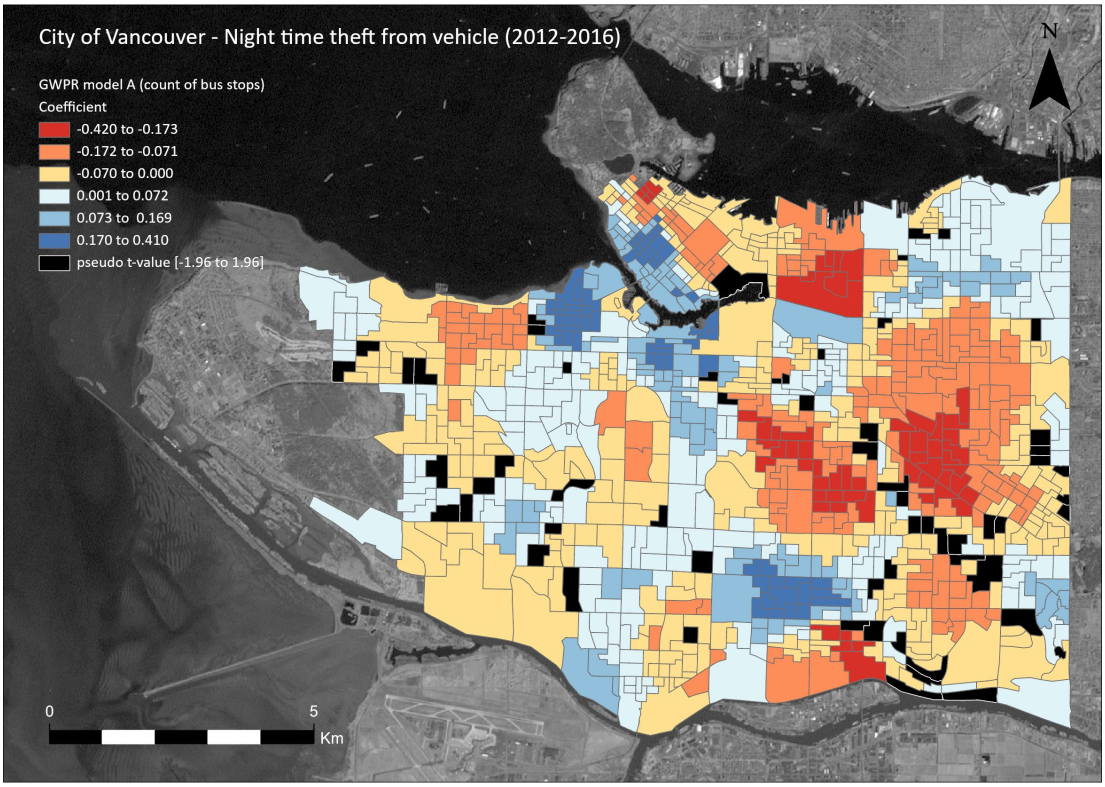

Figure 27 - "count of bus stops" model A local coefficients and distribution

Figure 28 shows the local coefficients of the "count of street trees" variable. Vegetation has been associated with ambiguous effects regarding crime (see section 2.3) and it is expected to have a mix of coefficient types and strengths. Most of the study area shows a weak to mild positive association between bus stops and NT-TFV crime with some pockets in the East showing a strong negative relationship, particularly in Renfrew-Collingwood.

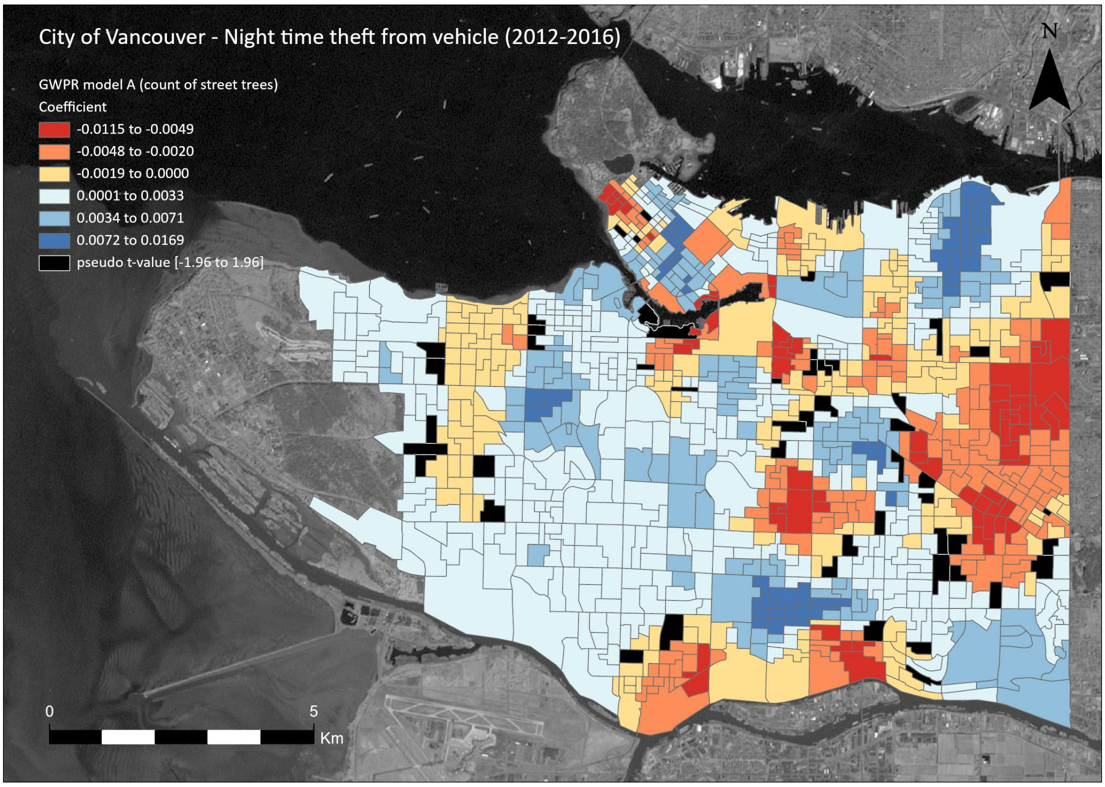

Figure 28 – "count of street trees" model A local coefficients and distribution

Figure 29 shows the local coefficients of the "count of street light poles" variable. Street lighting is seen as having ambiguous effects regarding crime and, as with street trees, it is expected to have a mix of coefficient types and strengths. Most of the study area shows a mix of weak to mild positive association and weak negative associations between street light poles and NT-TFV crime. There are several areas that show a mild to strong negative association, specifically in the Sunset/Oakridge/Kerrisdale and in Kitsilano/Arbutus-Ridge.

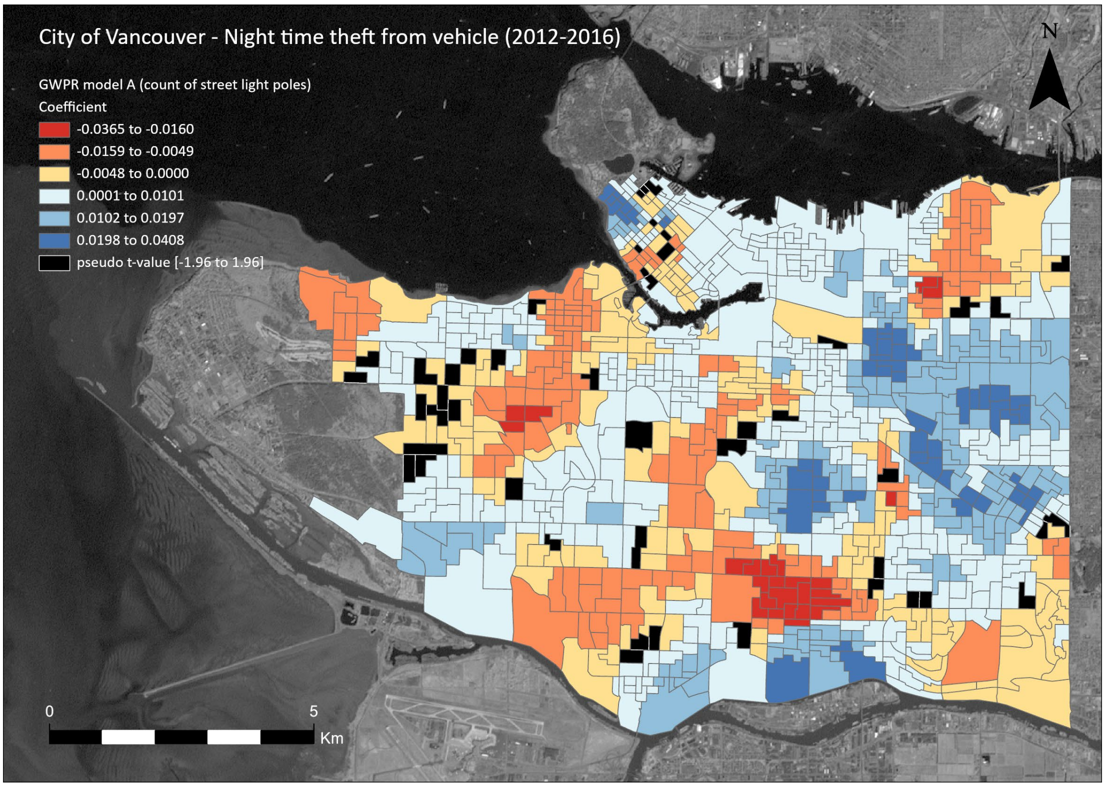

Figure 29 - "count of street light poles" model A local coefficients and distribution

4.3 Geographically Weighted Poisson Regression (GWPR) – Model B

The GWPR model B estimated "NT-TFV per 100 thousand inhabitants" and features the following independent variables: "Tree count", "Pole count", "Distance to rapid transit", "Bus stop count", "Distance to liquor store", and "Commercial land use rate", "Mixed land use rate", and "Distance to parking land use area". Model B diagnostics (Table 5) show an 80.95% deviance explained by the local model, against only 31.56% deviance explained by the global model. These results show Model A has a better goodness of fit with a lower AICc (393878.5) and a superior percent of deviance explained (84.27%).

The map of the local percentage deviance explained (Figure 30) shows a rather heterogenous distribution of explanatory power throughout the study area, in the same fashion of Model A. In fact, generally speaking, between the two models, the neighborhoods show quite similar percentages of deviance explained but, in Model B, CDAs where percentages are higher than 90% are Renfrew-Collingwood and Sunset. Figure 31 compares the local percentage deviance explained between Model B and

Model A in terms of percentage difference. The image shows that in most of the areas the models disagree in less than +/- 20%. However, some small areas in Kitsilano, West End, Dunbar and Sunset have a high percentage of variation. Model A had an average 60% deviance explained by the local model (min of 0.27 and max of 0.95), while model B had an average of 56% (min of 0.18 and max of 0.92).

Model B Diagnostics	
Deviance explained by the global model (non-spatial)	0.3156
Deviance explained by the local model	0.8095
Deviance explained by the local model vs. global model	0.7216
AICc	476602.4
Sigma-Squared	67081102.1
Sigma-Squared MLE	35485161.1
Effective Degrees of Freedom	525.8

Table 5 - GWPR Model B diagnostics

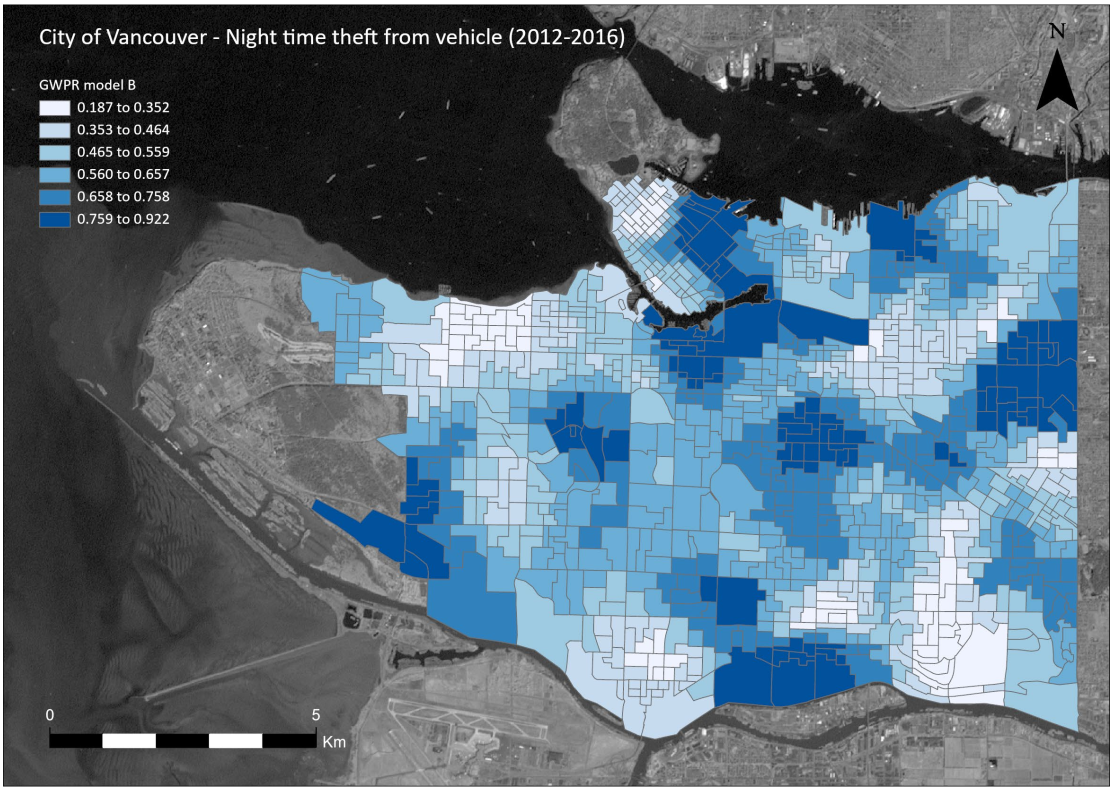

Figure 30 - GWPR model B local percentage deviance explained

The next set of maps will show the distribution of GWPR local coefficients for each independent variable that was included in Model B. The maps of the distribution of independent variables that are new to this model: “Distance to liquor store”, “Mixed land use rate”, and “Distance to parking land use area” can be found in Appendix 2: Maps.

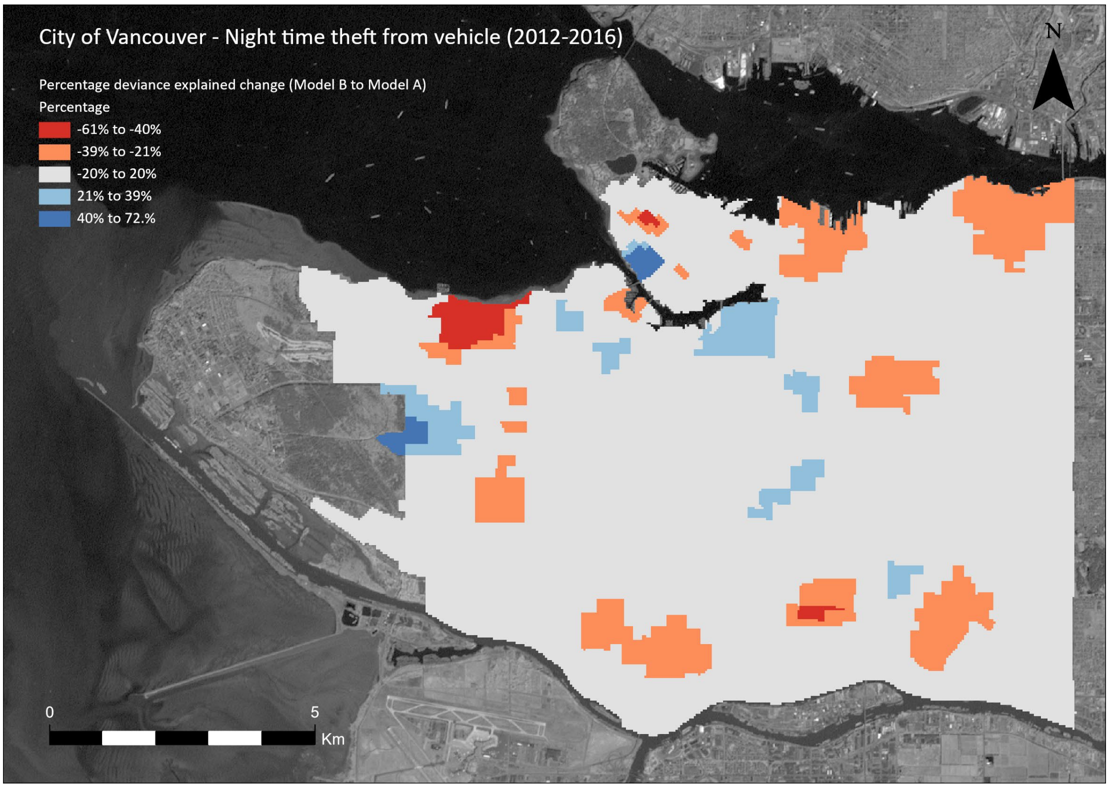

Figure 31 - Percentage of change between Model B and Model A deviance explained

Figure 32 shows the local coefficients of the "commercial land use rate" variable. Most of the study area shows a mild to strong negative association between commercial land use rates and NT-TFV crimes, similar to the spatial pattern found in Model A. A noticeable difference is that model B shows more areas where a strong negative association is found, at the East, West and South borders of the study area, and some areas that revealed a weak negative association and now show a moderate negative association. A decrease in strength of positive associations can also be observed. For the "commercial land use rate" variable, model B shifted the association towards a negative association when compared to model A.

Figure 33 shows the local coefficients of the "distance to rapid transit" variable. Most of the study area shows a weak negative association between rapid transit stations and NT-TFV crimes. Compared to model A, there is a shift in association strength where both negative and positive associations become weaker in model B.

Figure 34 shows the local coefficients of the "count of bus stops" variable. As with model A, coefficients are distributed in a very uneven way in a mix of positive and negative weak to mild associations. Both models show a similar distribution of local coefficients in strength and direction.

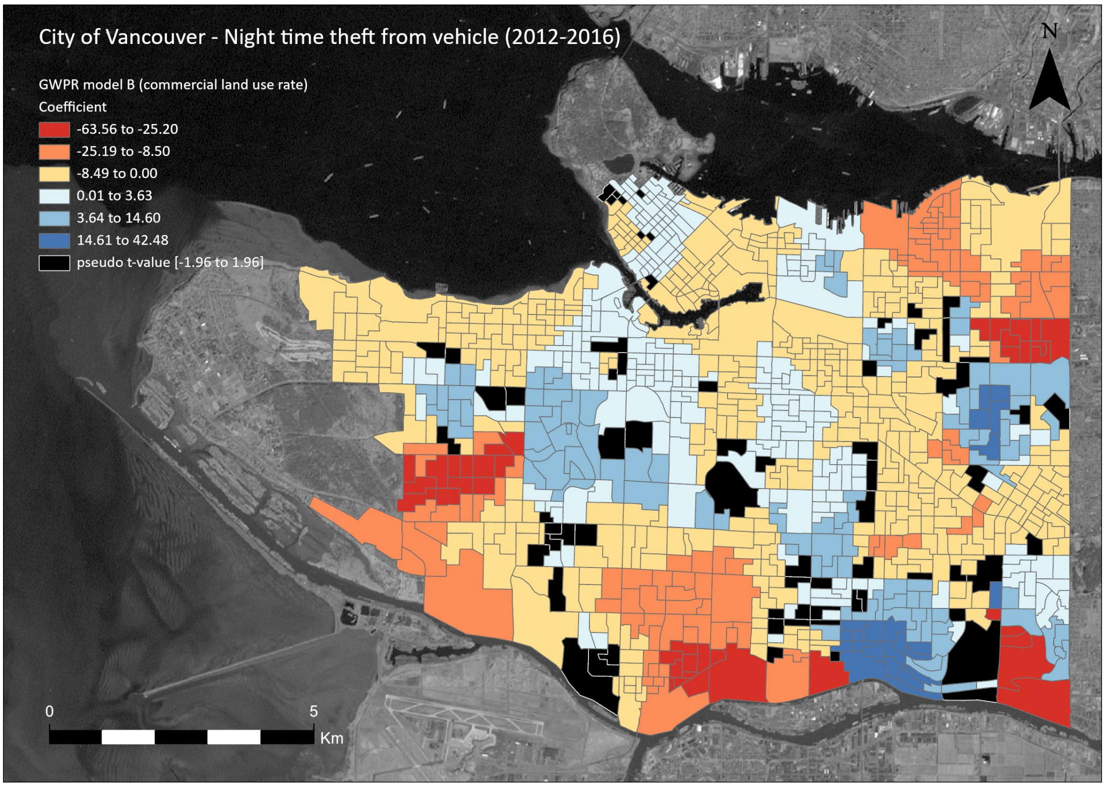

Figure 32 - "commercial land use rate" model B local coefficients.

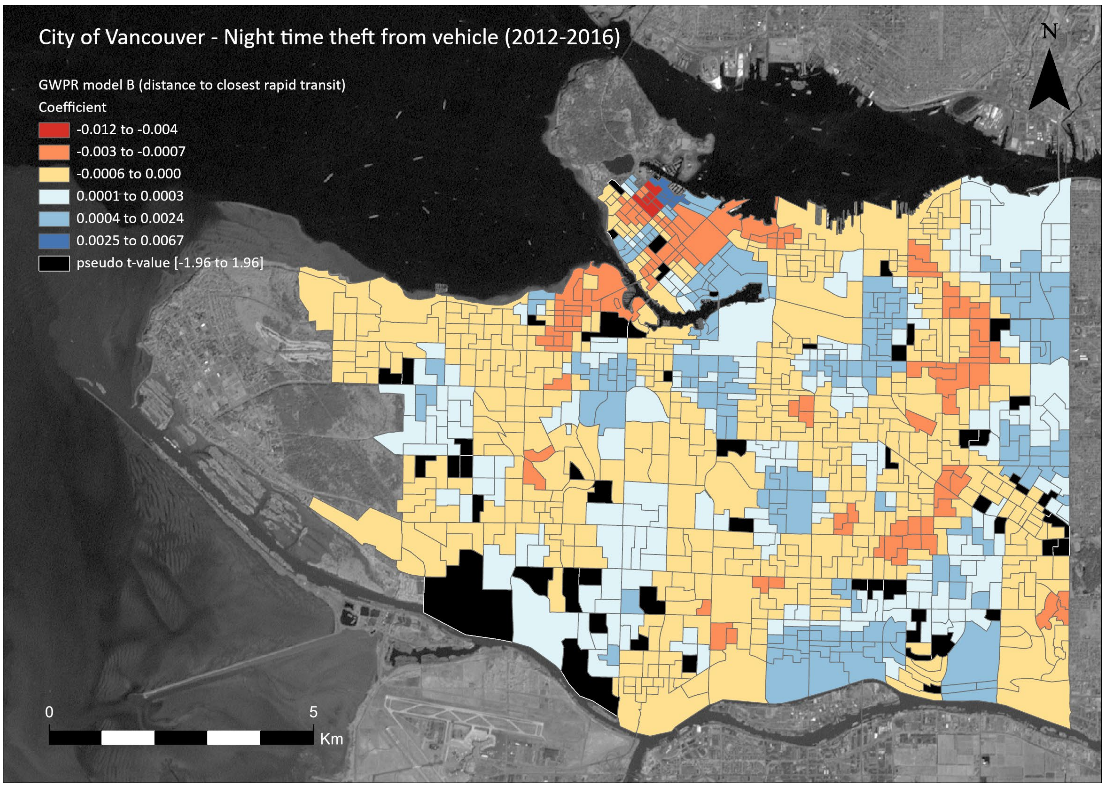

Figure 33 - "distance to closest rapid transit" model B local coefficients.

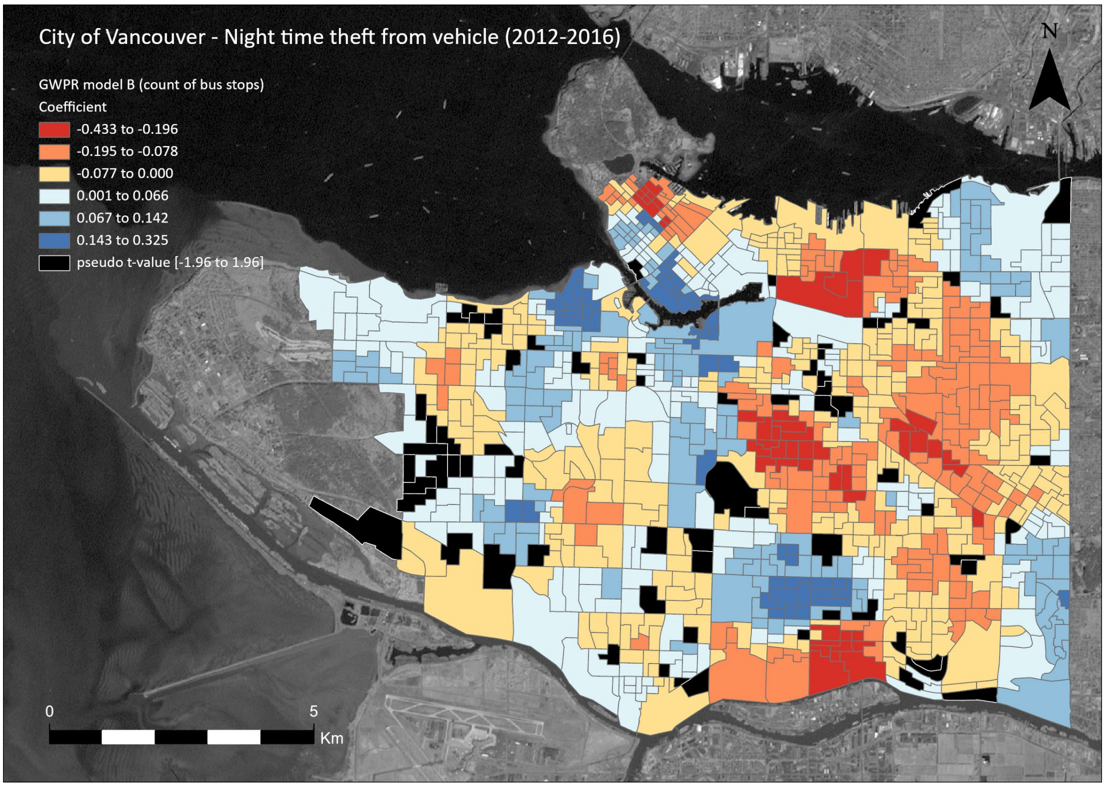

Figure 34 - "count of bus stops" model B local coefficients.

Figure 35 shows the local coefficients of the "count of street trees" variable. Most of the study area shows a weak positive association between bus stops and NT-TFV crime with some pockets in the East and South displaying a strong negative relationship. Compared to model A, there are shifts in association strength but without a specific pattern.

Figure 36 shows the local coefficients of the "count of street light poles" variable. Similar to model A, the area shows a mix of weak to mild positive association and weak to mild negative associations between street light poles and NT-TFV crime. The number of areas with a mild to strong negative association increased, model B shifted the association towards a negative association when compared to model A.

Figure 37 shows the local coefficients of the "distance to parking land use" variable. As a measure of how close an area dedicated to parking is, it can reflect the availability of targets for NT-TFV crimes, and it is a crime attractor. In that sense the expectation is that a positive association (target availability) would take place between parking land uses and crime. A good part of the study area shows a weak negative association, and the area of Downtown shows a stronger negative association. This could be because most parking land use in that neighborhood is underground and with a certain level of guardianship.

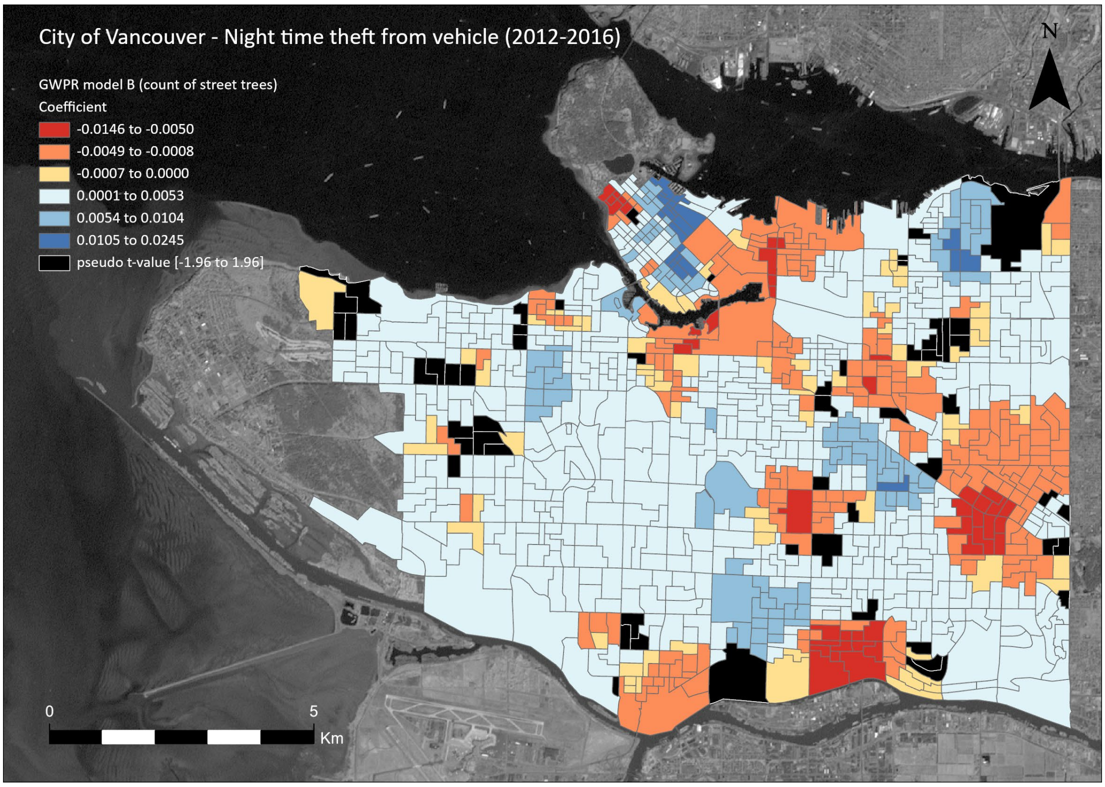

Figure 35 - "count of street trees" model B local coefficients.

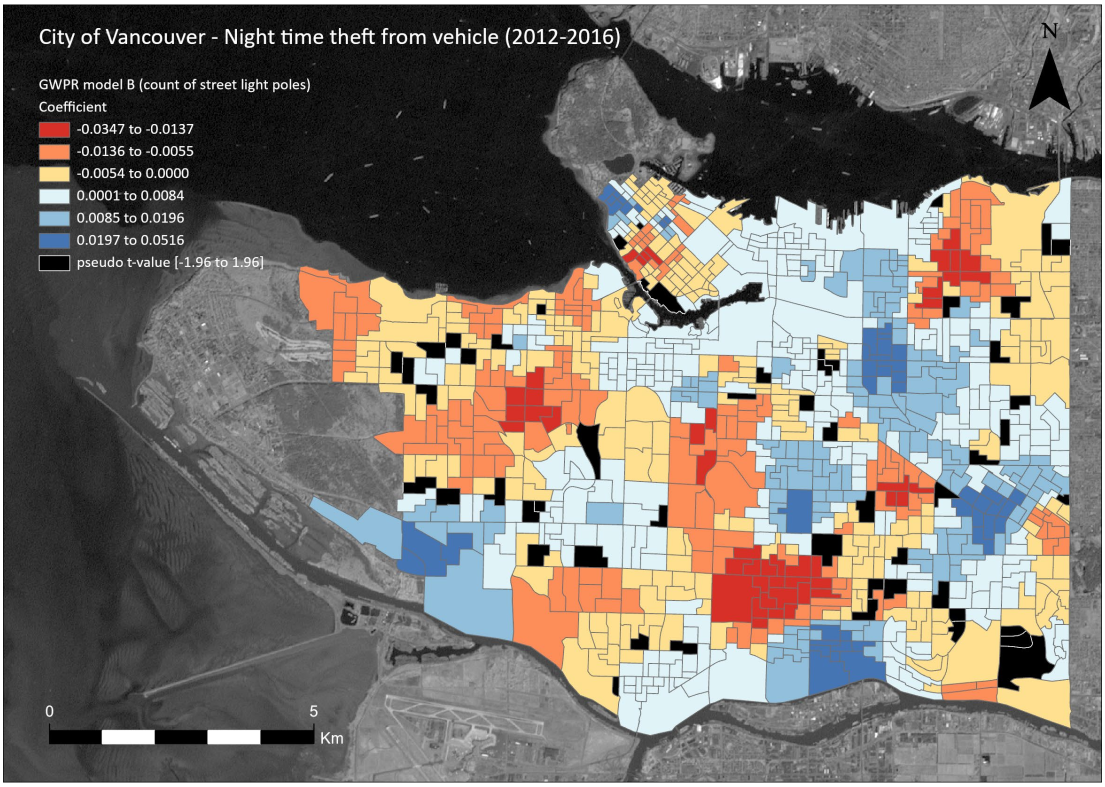

Figure 36 - "count of street light poles" model B local coefficients.

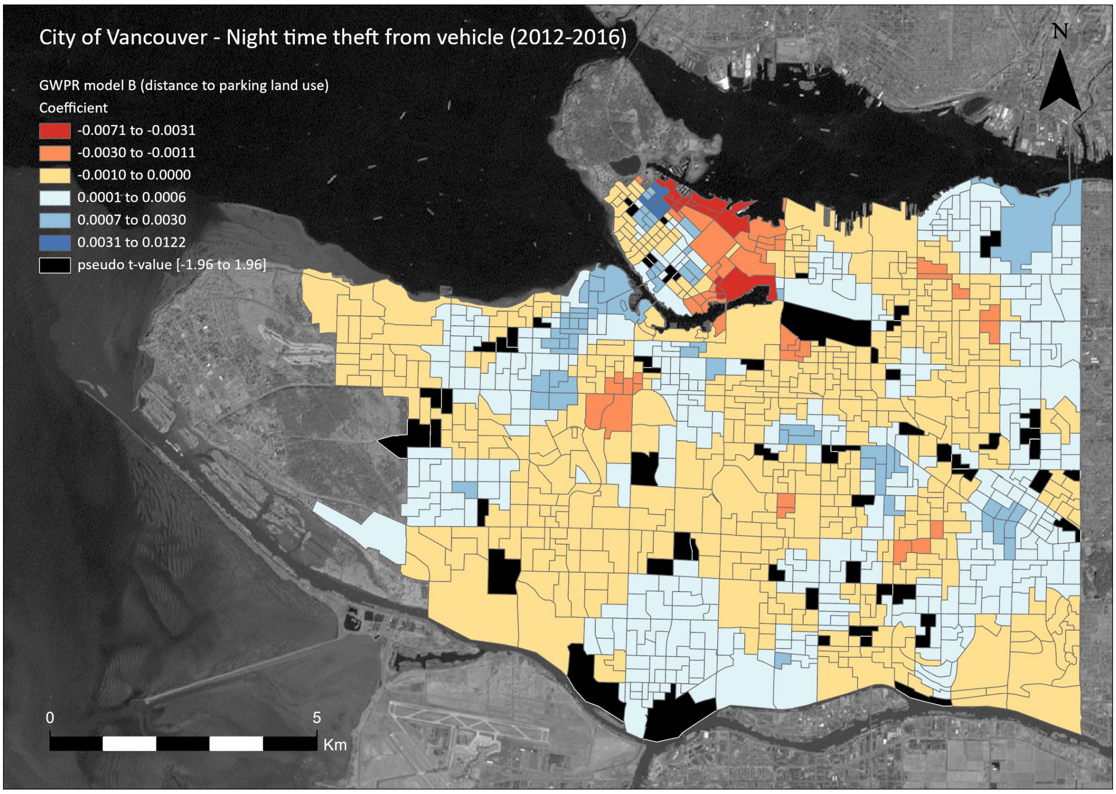

Figure 37 - "distance to parking land use" model B local coefficients and distribution

Figure 38 shows the local coefficients of the "distance to liquor store" variable, a measure of how close a liquor store is. As crime attractors, liquor stores are expected to have a positive association with crime. These stores are usually associated with larger commercial areas, potentially increasing the attractor factor. Distance to liquor stores has a quite mixed association in the study area. It is strongly positively associated with NT-TFV right at city center and parts of West End, while negatively associated in areas immediately bordering those.

Figure 39 shows the local coefficients of the "mixed land use rate" variable. As with the commercial land use, it can be a measure of ambient population during business hours, and it can also reflect the availability of targets for NT-TFV crimes, from the idea of crime attractors and generators. On the other hand, it can act as an indirect measure for better guardianship, both from people walking by and from retailers' surveillance mechanisms. The expectation is that either a positive association (target availability) or negative (guardianship) would take place between mixed land uses and crime. Mixed land use has a quite heterogeneous association with NT-TFV in the study area, with a prevalence on weak and mild positive associations, with some pockets of strong positive association, and one area with a strong negative association.

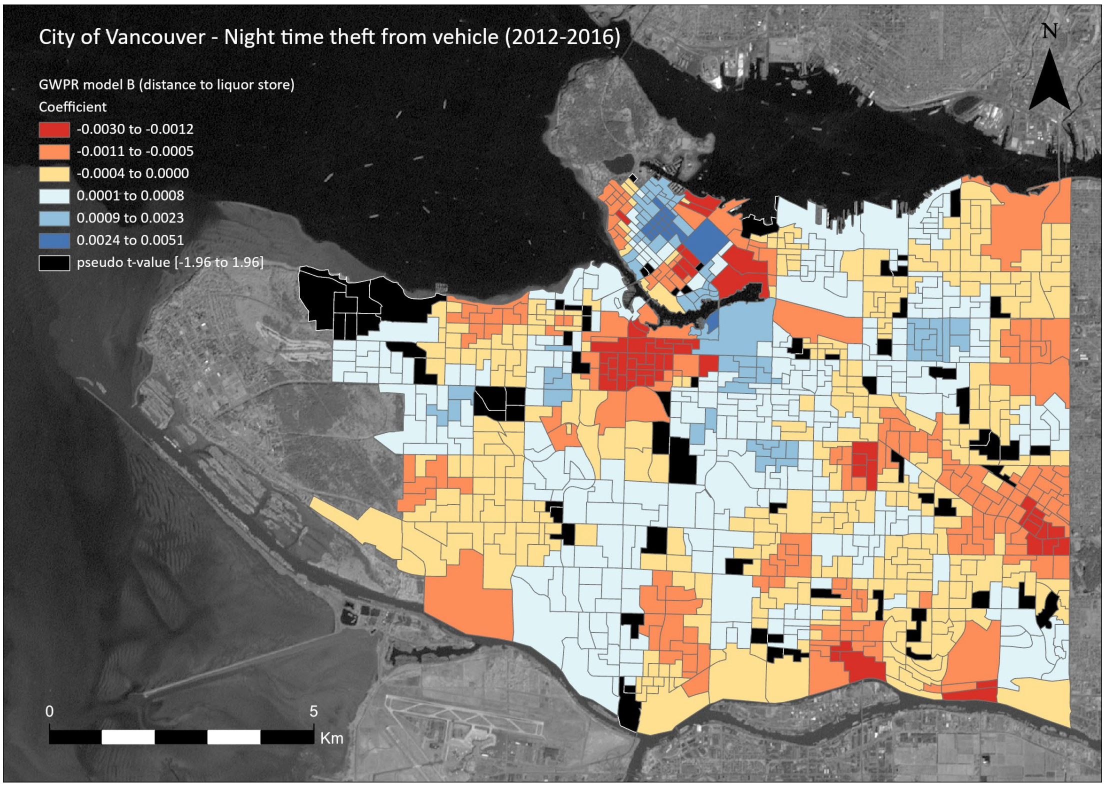

Figure 38 - "distance to liquor store" model B local coefficients and distribution

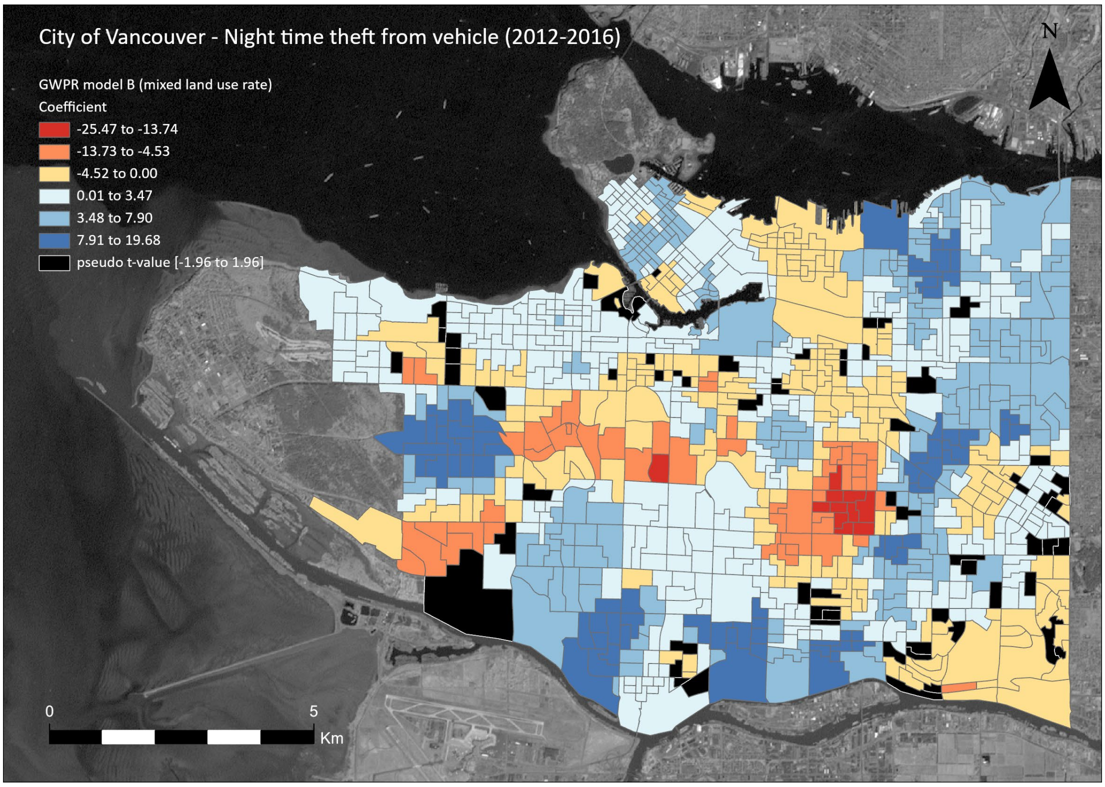

Figure 39 - "mixed land use rate" model B local coefficients and distribution

4.4 Ordinary Least Squares (OLS)

An OLS regression (Table 6) was pursued to analyze the relationship between street lighting poles and tree density and measured night-time street illuminance. The dependent variable is the “average illuminance per street segment” (collected with the sensing station) and the explanatory variables are “street light poles per meter” and “trees per meter”.

Model performance was below the initial expectation with an adjusted R-squared of 0.24. As expected, trees had a negative relationship with illuminance and a relatively strong association, while light pole density showed a strong positive relationship, both variables being statistically significative while VIF are well below the threshold. Additionally, the Koenker (BP) Statistic is not significant, and we can’t reject stationarity, as anticipated.

Table 6 - OLS results

Summary of OLS Results

Variable	Coefficient[a]	StdError	t-Statistic	Probability[b]	Robust_SE	Robust_t	Robust_Pr[b]	VIF[c]
Intercept	7.763355	1.037192	7.484974	0.000000*	1.202764	6.454597	0.000000*	--------
TREES_METER	-20.087801	8.548754	-2.349793	0.019294*	8.442334	-2.379413	0.017832*	1.016604
POLES_PER_METER	148.112289	14.057682	10.536039	0.000000*	19.945623	7.425804	0.000000*	1.016604

OLS Diagnostics

Input Features	street_all_var_lux_Spatial02	Dependent Variable	LUMI_AVG
Number of Observations	374	Akaike's Information Criterion (AICc)[d]	2708.301471
Multiple R-Squared[d]	0.251861	Adjusted R-Squared[d]	0.247828
Joint F-Statistic[e]	62.448718	Prob(>F), (2,371) degrees of freedom	0.000000*
Joint Wald Statistic[e]	68.586413	Prob(>chi-squared), (2) degrees of freedom	0.000000*
Koenker (BP) Statistic[f]	11.524286	Prob(>chi-squared), (2) degrees of freedom	0.003144*
Jarque-Bera Statistic[g]	383.270908	Prob(>chi-squared), (2) degrees of freedom	0.000000*

5. CONCLUSION

This thesis looked into the spatial relationship between crime determinants and NT-TFV. The primary interest was to investigate how public street illuminance, allied to a set of other explanatory variables, influences NT-TFV in Vancouver. The premise is that a set of variables that are supported by both the routine activities and the social disorganization theories should be combined and, to incorporate street illuminance, a set of two physical street level attributes were included: street light poles and street trees. These general objectives generated some questions that will be addressed in this section.

Do GWPR models benefit from including social disorganization (demographic) variables? The idea of selecting crime covariates that are supported by both theories is well supported by the literature. Nevertheless, this study was interested in how demographic characteristics could benefit a regression conducted in such a small unit of data aggregation, the census dissemination area, and to analyze such a specific type of crime, night time theft from vehicle. Two GWPR models were pursued where three demographic variables were swapped for three additional land use variables, maintaining a selection of other five variables. Although both models featured a quite similar distribution (both in location and in explanatory power) of the local percentage deviances, the models were compared in terms of the general explanatory value with 80.95% deviance explained by the local model B, against 84.27% explained by the local model A, that also featured a better goodness of fit. For the type of research in question, it is beneficial to analyze NT-TFV in a GWPR model that includes demographic explanatory variables as the results show a better fit to the data.

Findings suggest that while a citywide effect is evident for some of the explanatory variables, their spatial influence on NT-TFV crime rates significantly varies across neighborhood contexts (and even through adjacent CDAs), there is an evident nonstationary relation between the explanatory variables and the dependent variable. The spatial unit of aggregation is quite fine, and this adds to the visually fragmented distribution of the coefficients. The regions that contain the highest percentages of deviance explained are Downtown, Sunset, Oakridge and Renfrew-Collingwood. Downtown, the most populous neighborhood, is usually negatively influenced (mild to strong relations) by the explanatory variables, with the exception of "street light poles" (weak and positive) and "street trees" (mixed in strength and direction).

Does street illuminance affect reported thefts from vehicles during night-time in the city of Vancouver's neighborhoods? If yes, in which areas? "Count of street light poles" coefficients are quite heterogenous throughout the study area. Street lighting is a characteristic with ambiguous effects regarding crime and is expected to have a mix of coefficient types and strengths, even though there may exist a tendency to assume a negatively strong association. Most of the study area shows a mix of weak to mild positive association, specially on the East side, and weak to mild negative associations between street light poles and NT-TFV crime. Two areas show a mild to strong negative association, the first in North Sunset, Oakridge and Kerrisdale, and the second in Kitsilano and Arbutus-Ridge. Street light poles are physically more concentrated in Downtown, and so are NT-TFV crimes. The fact that, in Downtown, there is not a stronger positive relation between the count of light poles and NT-TFV crime rate indicates that street lighting may not be an important covariate with NT-TFV crime. During the exploratory regression the "street light poles density" was included as a possible variable, but it failed to be included in the final models, as an alternative to simple street light pole count, due to redundancy (high VIF). It would also be interesting to pursue a GWR on the street segment aggregated data.

What is the local relationship between street lighting pole and tree densities and collected night-time street illuminance? The OLS regression showed a moderately weak relation between light poles and tree densities to collected street illuminance. These are the two most explicit characteristics that can be linked to impact street luminance and yet the analysis failed to offer evidence of this possibly strong relation. One of the issues is certainly the illuminance collecting experiment that failed to offer a more reliable mean to analyze the relationship. The question of street lighting pole density being a usable proxy for street illuminance cannot be answered with confidence at this time.

5.1 Limitations

This study was conducted based on reported crime data from the VPD. The type of crime being examined is known for being a crime with high under-reporting rates. The data may not completely represent the phenomena being analyzed.

Although the best effort was made to have a set of temporal compatible variables, the analysis investigated data from non-coincident years. Also, some variables used in this investigation act as proxies for complex social processes that are at the true source for understanding crime.

Collecting illuminance on the car top can minimize the impact of incoming vehicles lights to influence collected illuminance but the experiment failed to cope with a high variance in illumination in the same street segment. As expected, illuminance levels peak under the light poles, and are at a minimum between poles. Since the average illumination was calculated, the amount of collecting points in a street segment and the exact location of these points will have an enormous impact on the results. Also, collection was pursued during winter, where deciduous trees tend to be striped of their leaves, with less impact over light falling down on the streets.

5.2 Future work

The theme of crime spatial analysis is certainly fertile in data and full of possibilities regarding methods, especially statistical ones. Certainly, the first suggestion for future work would be to use the soon to be released Census 2021 data.

Regarding the specific interest of this thesis, street luminance, it would be valuable to, in the future, investigate NT-TFV crimes by separating early nighttime occurrences, when the commerce is still open, and late night time, when the city is already sleeping. Such a study could help determine how much land use activity, especially commercial and alcohol related outlets, can influence crime in certain times of the day, presenting such a more abundant offer of targets that even “good lighting” cannot moderate.

Another potentially informative investigation could separate the city in areas according to their characteristics (residential, mixed, etc.), and model GWPR, and other types of regressions, with specific explanatory variables for each area, as opportunity and risks tend to be different.

The development of a more consistent method of illuminance collection could help advance the pole density usability as a proxy for street illuminance question. Collection needs to solve the issue of irregular illuminance throughout the street segments (that yields many zero Lux) in order to have a more consistent average illuminance per street segment and, consequently, a more reliable correlation test. Another possibility is to use streetlight consumption data as a proxy.

BIBLIOGRAPHIC REFERENCES

Andresen MA (2010) The Place of Environmental Criminology within Criminological Thought. In: Classics in Environmental Criminology. Routledge, Boca Raton, pp 5–28

Andresen MA (2006) A spatial analysis of crime in Vancouver, British Columbia: A synthesis of social disorganization and routine activity theory. Canadian Geographer 50:487–502. https://doi.org/10.1111/j.1541-0064.2006.00159.x

Andresen MA, Ha OK (2020) Spatially varying relationships between immigration measures and property crime Types in Vancouver Census Tracts, 2016. The British Journal of Criminology 60:1342–1367. https://doi.org/10.1093/bjc/azaa029

Andresen MA, Ha OK, Davies G (2021) Spatially Varying Unemployment and Crime Effects in the Long Run and Short Run. Professional Geographer 73:297–311. https://doi.org/10.1080/00330124.2020.1838928

Angel Shlomo (1968) Discouraging crime through city planning. Center for Planning and Development Research, Institute of Urban and Regional Development, University of California, Berkeley

Anselin L (1995) Local Indicators of Spatial Association—LISA. Geographical Analysis 27:93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x

Anselin L, Cohen J, Cook D, et al (2000) Spatial Analyses of Crime. CRIMINAL JUSTICE 51

Arnio AN (2021) Examining the Effects of Neighborhood Contextual Factors on Officer-Involved Shootings. null 38:626–652. https://doi.org/10.1080/07418825.2019.1679862

Atkins S, Husain S, Storey A (1991) The Influence of Street Lighting on Crime and Fear of Crime. London

Bappee FK, Petry LM, Soares A, Matwin S (2020) Analyzing the Impact of Foursquare and Streetlight Data with Human Demographics on Future Crime Prediction. In: 16th International Conference on Data Science (ICDATA20). pp 1–15

B.C. Government (2021a) B.C. Data Catalogue. In: B.C. Data Catalogue. https://catalogue.data.gov.bc.ca/. Accessed 20 Jun 2021

B.C. Government (2021b) BC Address Geocoder. In: BC Address Geocoder. https://www2.gov.bc.ca/gov/content/data/geographic-data-services/location-services/geocoder. Accessed 20 Jun 2021

Block R, Davis S (1996) The environs of rapid transit stations: a focus for street crime or just another risky place? Preventing Mass Transit Crime (Crime Prevention Studies, vol 6)

Braga AA, Clarke RV (2014) Explaining High-Risk Concentrations of Crime in the City: Social Disorganization, Crime Opportunities, and Important Next Steps. Journal of Research in Crime and Delinquency 51:480–498. https://doi.org/10.1177/0022427814521217

Brantingham P, Brantingham P (1995) Criminality of place. European Journal on Criminal Policy and Research 3:5–26. https://doi.org/10.1007/BF02242925

Brantingham P, Brantingham P (1993) Nodes, paths, and edges: considerations on the complexity of crime and the physical environment. Journal of Environmental Psychology 13:3–28. https://doi.org/10.1016/S0272-4944(05)80212-9

Brantingham P, Brantingham P, Andresen MA (2016) The geometry of crime and crime pattern theory. In: Environmental criminology and crime analysis. Routledge, London, pp 98–115

Chalfin A, Hansen B, Lerner J, Parker L (2021) Reducing Crime Through Environmental Design: Evidence from a Randomized Experiment of Street Lighting in New York City. Journal of Quantitative Criminology. https://doi.org/10.1007/s10940-020-09490-6

Chalfin A, Kaplan J, LaForest M (2020) Street Light Outages, Public Safety and Crime Displacement: Evidence from Chicago. CJRN: Neighborhoods (Topic). https://dx.doi.org/10.2139/ssrn.3526467

City of Vancouver (2021) Open data portal. In: City of Vancouver Open Data Portal. https://opendata.vancouver.ca/pages/home/. Accessed 20 Jun 2021

Clarke RV (1995) Situational Crime Prevention. Crime and Justice 19:91–150

Cohen LE, Felson M (1979) Social Change and Crime Rate Trends: A Routine Activity Approach. American Sociological Review 44:588–608. https://doi.org/10.2307/2094589

Cornish DB, Clarke RV (1987) Understanding crime displacement: an application of rational choice theory. Criminology 25:17. https://doi.org/10.1111/j.1745-9125.1987.tb00826.x

Cozens PM, Neale RH, Whitaker J, et al (2003) A Critical Review of Street Lighting, Crime and Fear of Crime in the British City. Crime Prevention and Community Safety 5:7–24. https://doi.org/10.1057/palgrave.cpcs.8140143

Dakin K, Xie W, Parkinson S, et al (2020) Built environment attributes and crime: an automated machine learning approach. Crime Science 9:12. https://doi.org/10.1186/s40163-020-00122-9

Deng M, Yang W, Chen C, Liu C (2021) Exploring associations between streetscape factors and crime behaviors using Google Street View images. Frontiers of Computer Science 16:164316. https://doi.org/10.1007/s11704-020-0007-z

Doleac J, Sanders N (2015) Under the Cover of Darkness: How Ambient Light Influences Criminal Activity. Review of Economics and Statistics 97:1093–1103. https://doi.org/10.1162/REST_a_00547

ESRI (2021) ArcGIS Pro geoprocessing tool reference. https://pro.arcgis.com/en/pro-app/2.8/tool-reference/main/arcgis-pro-tool-reference.htm. Accessed 20 Jun 2021

Farrington DP, Welsh BC (2002) Improved street lighting and crime prevention. Justice Quarterly 19:313–342. https://doi.org/10.1080/07418820200095261

Farrington DP, Welsh BC (2008) Effects of Improved Street Lighting on Crime. Campbell Systematic Reviews 4:1–51. https://doi.org/10.4073/csr.2008.13

Getis A, Ord JK (1992) The Analysis of Spatial Association by Use of Distance Statistics. Geographical Analysis 24:189–206. https://doi.org/10.1111/j.1538-4632.1992.tb00261.x

Gilstad-Hayden K, Wallace LR, Carroll-Scott A, et al (2015) Research note: Greater tree canopy cover is associated with lower rates of both violent and property crime in New Haven, CT. Landscape and Urban Planning 143:248–253. https://doi.org/10.1016/j.landurbplan.2015.08.005

Hart TC, Miethe TD (2014) Street robbery and public bus stops: A case study of activity nodes and situational risk. Security Journal 27:180–193. https://doi.org/10.1057/sj.2014.5

Hodgkinson T, Andresen MA, Farrell G (2016) The decline and locational shift of automotive theft: A local level analysis. Journal of Criminal Justice 44:49–57. https://doi.org/10.1016/j.jcrimjus.2015.12.003

Insurance Corporation of British Columbia (2021) Vehicle population data. https://public.tableau.com/app/profile/icbc/viz/VehiclePopulationIntroPage/VehiclePopulationData. Accessed 20 Jun 2021

Jacobs J (1961) The Death and Life of Great American Cities. Random House, New York

Jeffery CR (1971) Crime Prevention Through Environmental Design. American Behavioral Scientist 14:598. https://doi.org/10.1177/000276427101400409

Jones RW, Pridemore WA (2019) Toward an Integrated Multilevel Theory of Crime at Place: Routine Activities, Social Disorganization, and The Law of Crime Concentration. Journal of Quantitative Criminology 35:543–572. https://doi.org/10.1007/s10940-018-9397-6

Kaplan J (2019) The Effect of Moonlight on Outdoor Nighttime Crime. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3369228

Kinney JB, Brantingham PL, Wuschke K, et al (2008) Crime Attractors, Generators and Detractors: Land Use and Urban Crime Opportunities. Built Environment (1978-) 34:62–74

Kuo FE, Sullivan WC (2001) Environment and Crime in the Inner City: Does Vegetation Reduce Crime? Environment and Behavior 33:343–367. https://doi.org/10.1177/0013916501333002

Leung A (2004) The Cost of Pain Suffering from Crime in Canada. Research and Statistics Division, Department of Justice Canada

Liggett R, Loukaitou-Sideris A, Iseki H (2003) Journeys to Crime: Assessing the Effects of a Light Rail Line on Crime in the Neighborhoods. Journal of Public Transportation 6:85–115. https://doi.org/10.5038/2375-0901.6.3.5

Lin J, Wang Q, Huang B (2021) Street trees and crime: What characteristics of trees and streetscapes matter. Urban Forestry & Urban Greening 65:127366. https://doi.org/10.1016/j.ufug.2021.127366

Mao Y, Yin L, Zeng M, et al (2021) Review of Empirical Studies on Relationship between Street Environment and Crime. Journal of Planning Literature. https://doi.org/10.1177/0885412221992280

Marchant PR (2004) A Demonstration That the Claim That Brighter Lighting Reduces Crime Is Unfounded. The British Journal of Criminology 44:441–447

Metro Vancouver (2021) Open Data Catalogue. In: Open Data Catalogue. http://www.metrovancouver.org/data. Accessed 20 Jun 2021

Nakaya T, Fotheringham AS, Brunsdon C, Charlton M (2005) Geographically weighted Poisson regression for disease association mapping. Statistics in Medicine 24:2695–2717. https://doi.org/10.1002/sim.2129

National Research Council Canada (2021) Sunrise/sunset calculator. https://nrc.canada.ca/en/research-development/products-services/software-applications/sun-calculator/. Accessed 29 Sep 2021

Newman Oscar (1972) Defensible space; crime prevention through urban design.

Osgood DW (2000) Poisson-Based Regression Analysis of Aggregate Crime Rates. Journal of Quantitative Criminology 16:21–43. https://doi.org/10.1023/A:1007521427059

Painter K (1994) The Impact of Street Lighting on Crime, Fear, and Pedestrian Street Use. Security J 5:116–125

Painter K, Farrington DP (1999) Improved Street Lighting: Crime Reducing Effects and Cost-Benefit Analyses. Security Journal 12:17–32. https://doi.org/10.1057/palgrave.sj.8340037

Ramsay M (1991) The Effect of Better Street Lighting on Crime and Fear : a Review. London

Rice KJ, Smith WR (2002) Socioecological Models of Automotive Theft: Integrating Routine Activity and Social Disorganization Approaches. Journal of Research in Crime and Delinquency 39:304–336. https://doi.org/10.1177/002242780203900303

Ristea A, Andresen MA, Leitner M (2018) Using tweets to understand changes in the spatial crime distribution for hockey events in Vancouver. The Canadian Geographer / Le Géographe canadien 62:338–351. https://doi.org/10.1111/cag.12463

Sampson RJ, Groves WB (1989) Community Structure and Crime: Testing Social-Disorganization Theory. American Journal of Sociology 94:774–802

Shaw CR, Mckay HD (2010) Juvenile Delinquency and Urban Areas: A Study of Rates of Delinquency in Relation to Differential Characteristics of Local Communities in American Cities (1969). In: Classics in Environmental Criminology. Routledge, Boca Raton

Shen J-L, Andresen MA (2021) A Tale of Two Theories: Whither Social Disorganization Theory and the Routine Activities Approach? Canadian Journal of Criminology and Criminal Justice 63:1–22. https://doi.org/10.3138/cjccj.2021-0013

Statistics Canada (2017) Census of Population, 2016

Struyf P (2020) Fear of the dark. In: Crime and Fear in Public Places: Towards Safe, Inclusive and Sustainable Cities., Ceccato, V., Nalla, M.K. (Eds.). Routledge, pp 347–361

Stucky TD, Smith SL (2017) Exploring the conditional effects of bus stops on crime. Security Journal 30:290–309. https://doi.org/10.1057/sj.2014.16

Sypion-Dutkowska N, Leitner M (2017) Land use influencing the spatial distribution of urban crime: A case study of Szczecin, Poland. ISPRS International Journal of Geo-Information 6:. https://doi.org/10.3390/ijgi6030074

Tagliabue LC, Re Cecconi F, Moretti N, et al (2020) Security assessment of urban areas through a gis-based analysis of lighting data generated by IoT sensors. Applied Sciences (Switzerland) 10:. https://doi.org/10.3390/app10062174

Tavares JP, Costa AC (2021) Spatial Modeling and Analysis of the Determinants of Property Crime in Portugal. ISPRS International Journal of Geo-Information 10:. https://doi.org/10.3390/ijgi10110731

TransLink (South Coast British Columbia Transportation Authority), Lesack P (2017) TransLink Transit GIS Data, 21 April 2017

Troy A, Morgan Grove J, O’Neil-Dunne J (2012) The relationship between tree canopy and crime rates across an urban–rural gradient in the greater Baltimore region. Landscape and Urban Planning 106:262–270. https://doi.org/10.1016/j.landurbplan.2012.03.010

U.S. Department of Justice - Federal Bureau of Investigation (2017) Crime in the United States, 2016. https://ucr.fbi.gov/crime-in-the-u.s/2016/crime-in-the-u.s.-2016/

Vancouver Police Department (2021a) Crime Incident and Crime Rate Statistics. https://vpd.ca/crime-statistics/. Accessed 20 Jun 2021

Vancouver Police Department (2021b) GeoDASH Open Data. https://geodash.vpd.ca/opendata/. Accessed 20 Jun 2021

Wortley R, Townsley M (2016) Environmental criminology and crime analysis. Situating the theory, analytic approach and application. In: Environmental criminology and crime analysis. Routledge, London, pp 1–25

Ye C, Chen Y, Li J (2018) Investigating the Influences of Tree Coverage and Road Density on Property Crime. IJGI 7:101. https://doi.org/10.3390/ijgi7030101

Yu SV (2009) Bus stops and crime: Do bus stops increase crime opportunities in local neighborhoods? Ph.D., Rutgers The State University of New Jersey

Zeng M, Mao Y, Wang C (2021) The relationship between street environment and street crime: A case study of Pudong New Area, Shanghai, China. Cities 112:. https://doi.org/10.1016/j.cities.2021.103143

ANNEXES

Appendix 1: Blockly assembly and Arduino source code

{ }
Arduino Source Code

```
#include <SPI.h>

#include <Wire.h>

#include <Adafruit_GFX.h>

#include <Adafruit_SSD1306.h>

#include "SenseBoxMCU.h"

#include <SD.h>

#include <SparkFun_u-blox_GNSS_Arduino_Library.h>

#include <avr/dtostrf.h>



char tsBuffer[21];


const long interval = 2000;

long time_start = 0;

long time_actual = 0;


#define OLED_RESET 4

Adafruit_SSD1306 display(OLED_RESET);

File Ltdata;

Lightsensor lightsensor;

SFE_UBLOX_GNSS myGNSS;

const char SENSOR_ID1DF[] PROGMEM = "61c92eca866767001b9c51df";

static const uint8_t NUM_SENSORS = 1;

typedef struct measurement {

    const char *sensorId;

    float value;

  } measurement;

char buffer[750];

measurement measurements[NUM_SENSORS];

   uint8_t num_measurements = 0;
```

```
char* getTimeStamp()

{

  if (myGNSS.getTimeValid() == true)

    {

    sprintf(tsBuffer, "%04d-%02d-%02dT%02d:%02d:%02dZ",

    myGNSS.getYear(), myGNSS.getMonth(), myGNSS.getDay(), myGNSS.getHour(), myGNSS.getMinute(),
myGNSS.getSecond());

    }

return tsBuffer;

}

void addMeasurement(const char *sensorId, float value) {

        measurements[num_measurements].sensorId = sensorId;

        measurements[num_measurements].value = value;

        num_measurements++;

        }

void writeMeasurementsToSdCard(char* timeStamp, int32_t latitudes, int32_t longitudes) {

    // iterate throug the measurements array

        for (uint8_t i = 0; i < num_measurements; i++) {

            char lng[20];

            char lat[20];

            float longitude = longitudes / (float)10000000;

            float latitude = latitudes / (float)10000000;

            dtostrf(longitude, 8, 7, lng);

            dtostrf(latitude, 8, 7, lat);

            sprintf_P(buffer, PSTR("%s,%9.2f,%s,%02s,%02s"),  measurements[i].sensorId, measurements[i].value, timeStamp, lng,
lat);

            // transmit buffer to client

            Ltdata.println(buffer);

            }
```

```
        // reset num_measurements
        num_measurements = 0;
      }



  void saveValues() {
      // send measurements
      writeMeasurementsToSdCard(getTimeStamp(), , );
      num_measurements = 0;
  }




void setup() {
senseBoxIO.powerI2C(true);
delay(2000);
display.begin(SSD1306_SWITCHCAPVCC, 0x3D);
display.display();
delay(100);
display.clearDisplay();
SD.begin(28);

Ltdata = SD.open("Ltdata.csv", FILE_WRITE);
Ltdata.close();

lightsensor.begin();
 Wire.begin();

 if (myGNSS.begin() == false) //Connect to the Ublox module using Wire port
 {
   Serial.println(F("Ublox GPS not detected at default I2C address. Please check wiring. Freezing."));
   while (1);
 }
```

```
  myGNSS.setI2COutput(COM_TYPE_UBX); //Set the I2C port to output UBX only (turn off NMEA noise)

  myGNSS.saveConfiguration(); //Save the current settings to flash and BBR

}

void loop() {

  display.setCursor(0,0);

  display.setTextSize(1);

  display.setTextColor(WHITE,BLACK);

  display.println((String("Illuminance: ") + String(lightsensor.getIlluminance()) + String(" Lux")));

  display.setCursor(0,12);

  display.setTextSize(1);

  display.setTextColor(WHITE,BLACK);

  display.println((String("Lat: ") + String(myGNSS.getLatitude())));

  display.setCursor(0,24);

  display.setTextSize(1);

  display.setTextColor(WHITE,BLACK);

  display.println((String("Long: ") + String(myGNSS.getLongitude())));

  display.setCursor(0,36);

  display.setTextSize(1);

  display.setTextColor(WHITE,BLACK);

  display.println((String("Time: ") + String(getTimeStamp())));

  display.setCursor(0,55);

  display.setTextSize(1);

  display.setTextColor(WHITE,BLACK);

  display.println((String("Signal: ") + String(myGNSS.getFixType())));

display.display();

display.clearDisplay();

time_start = millis();

if (time_start > time_actual + interval) {

  time_actual = millis();
```

```
  Ltdata = SD.open("Ltdata.csv", FILE_WRITE);
     addMeasurement(SENSOR_ID1DF,lightsensor.getIlluminance());
   saveValues();
  Ltdata.close();
}


}
```

Appendix 2: Maps and histogram

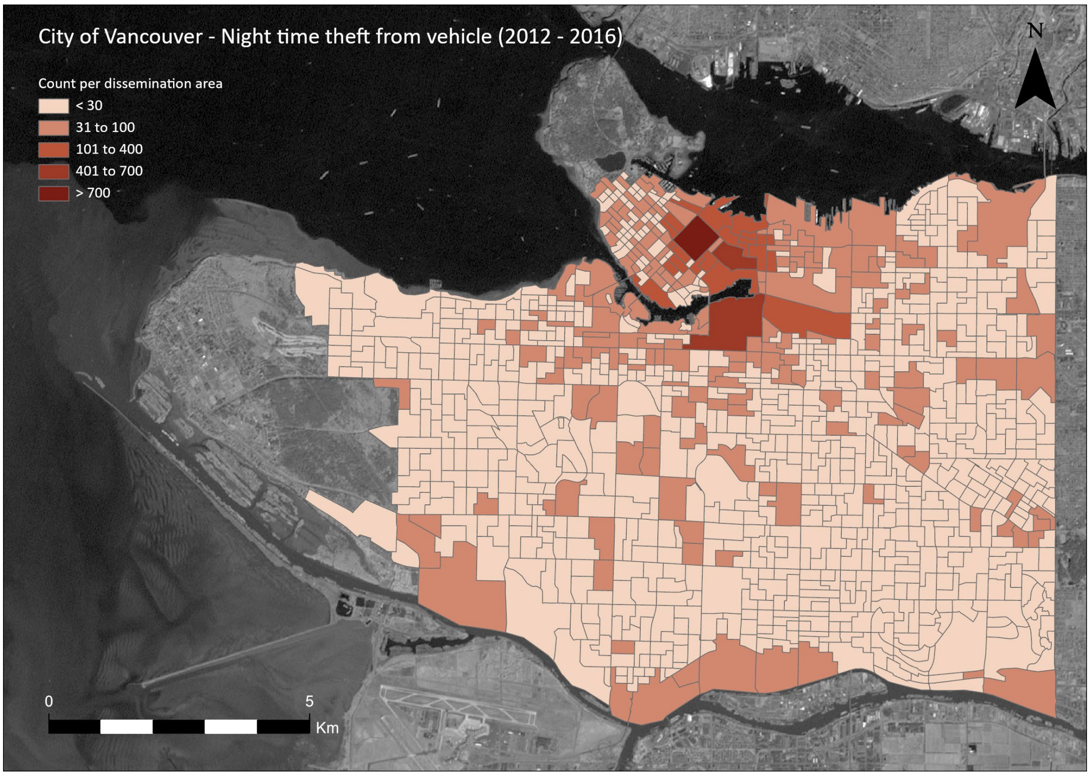

Figure A 1 - Spatial distribution of NT-TFV crime per Vancouver CDAs, from 2012 to 2016

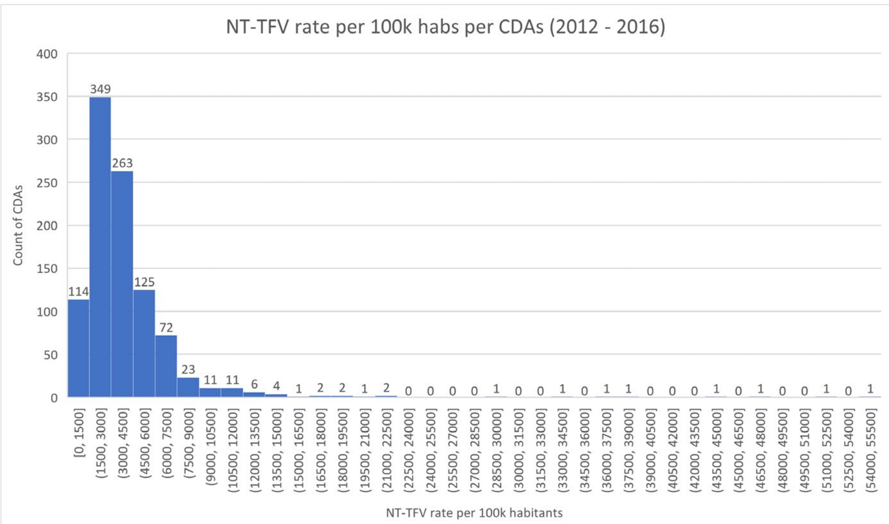

Figure A 2 - Histogram of NT-TFV crime rate per 100k inhabitants per Vancouver CDA, from 2012 to 2016

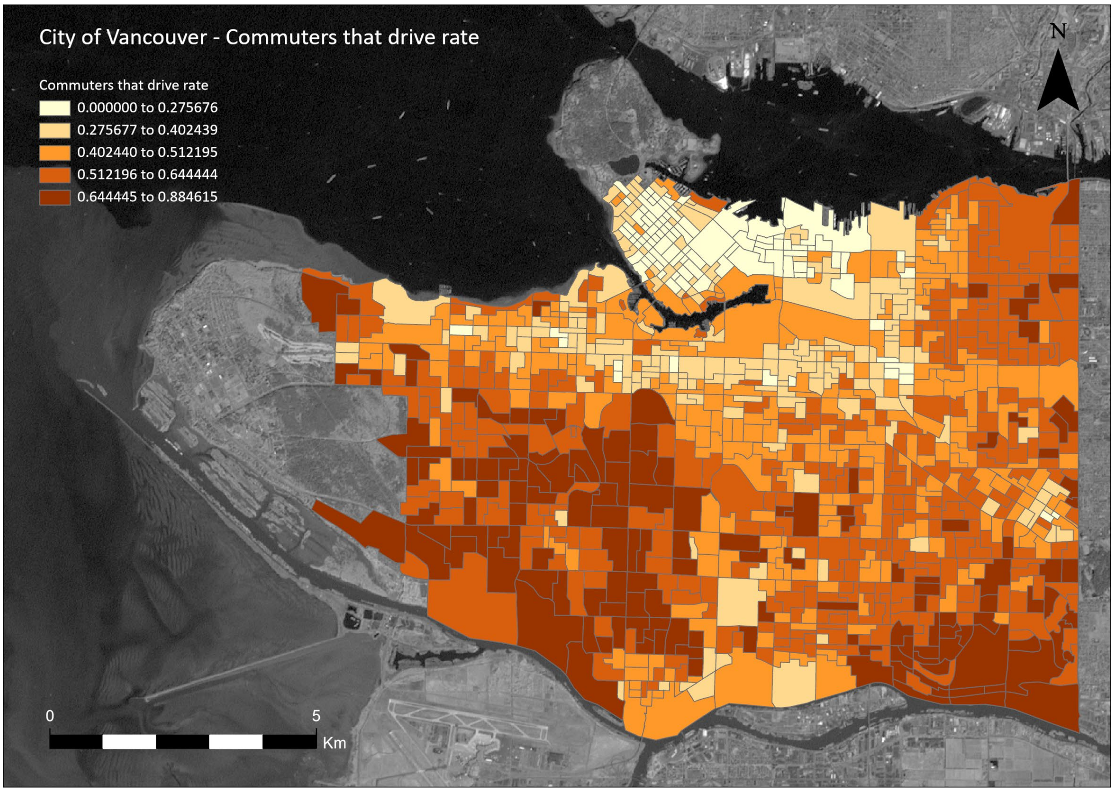

Figure A 3 - Spatial distribution of commuter drivers to inhabitants rate

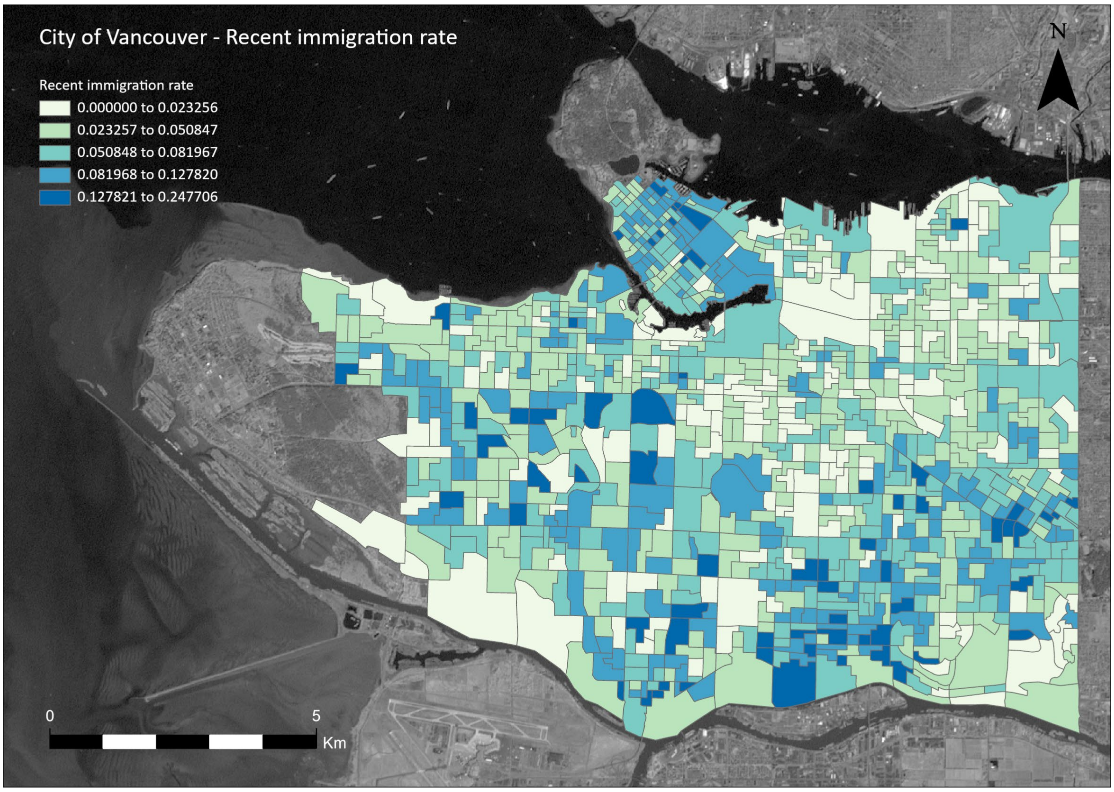

Figure A 4 - Spatial distribution of recent immigration rate

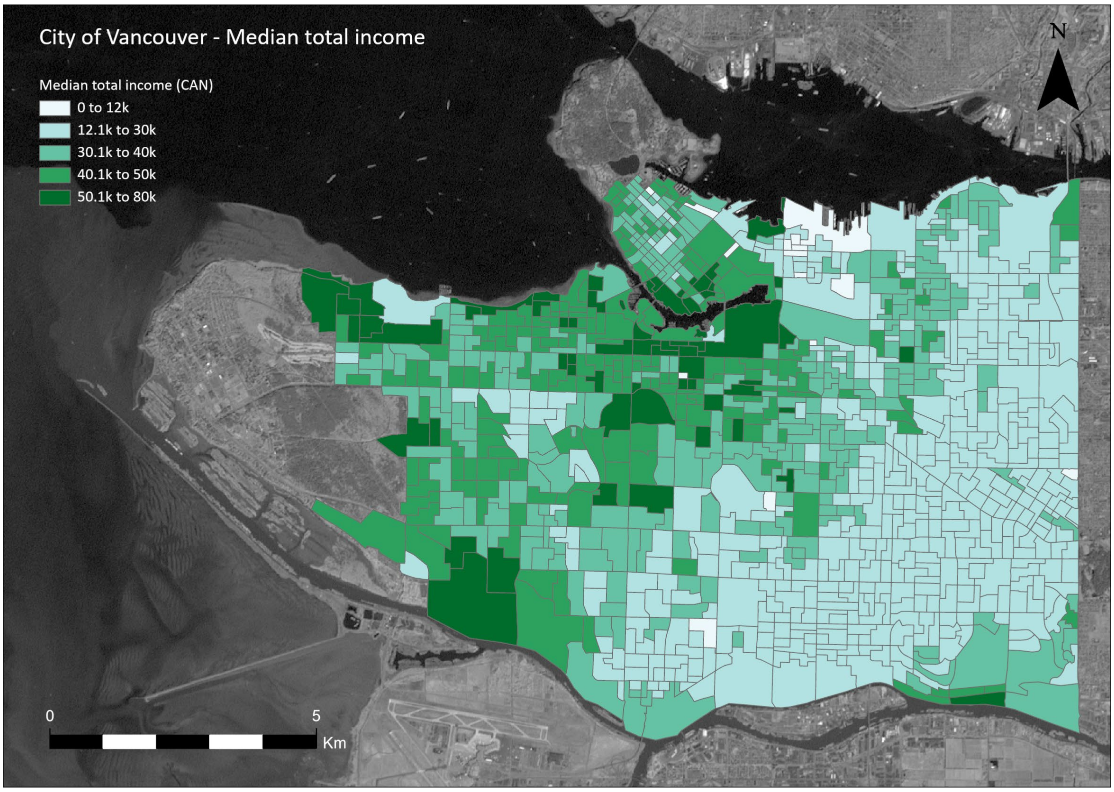

Figure A 5 - Spatial distribution of median total income

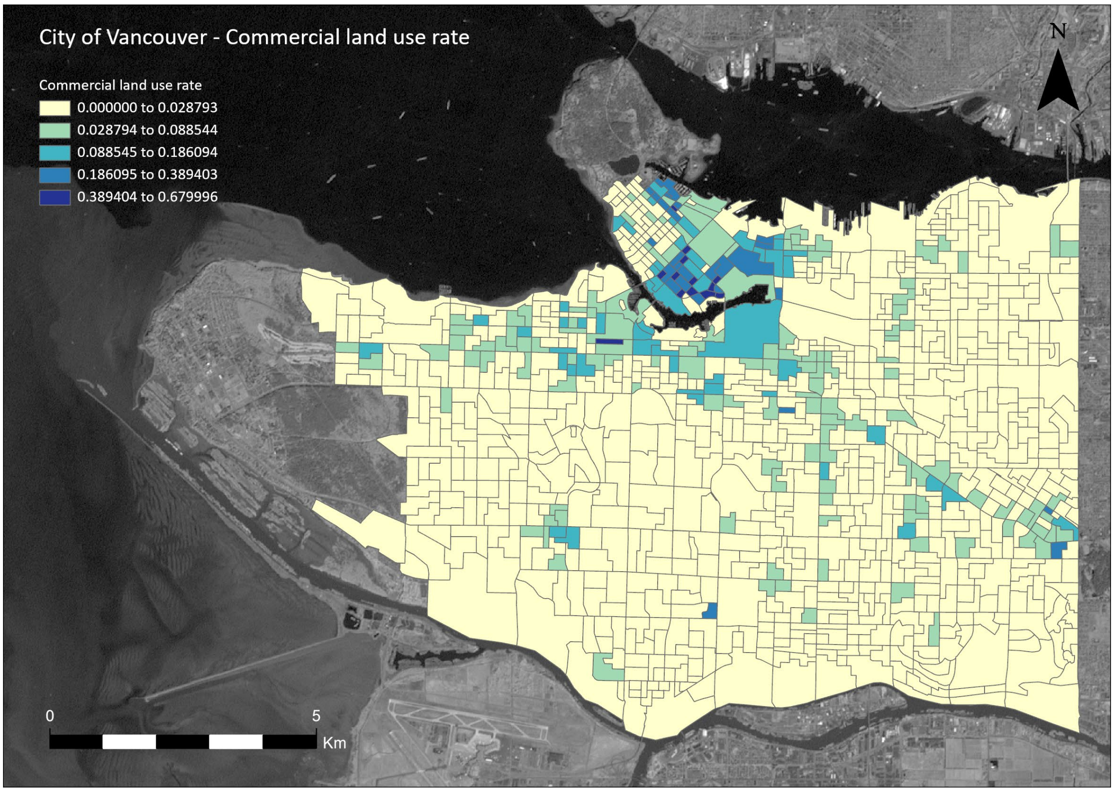

Figure A 6 - Spatial distribution of commercial land use rate

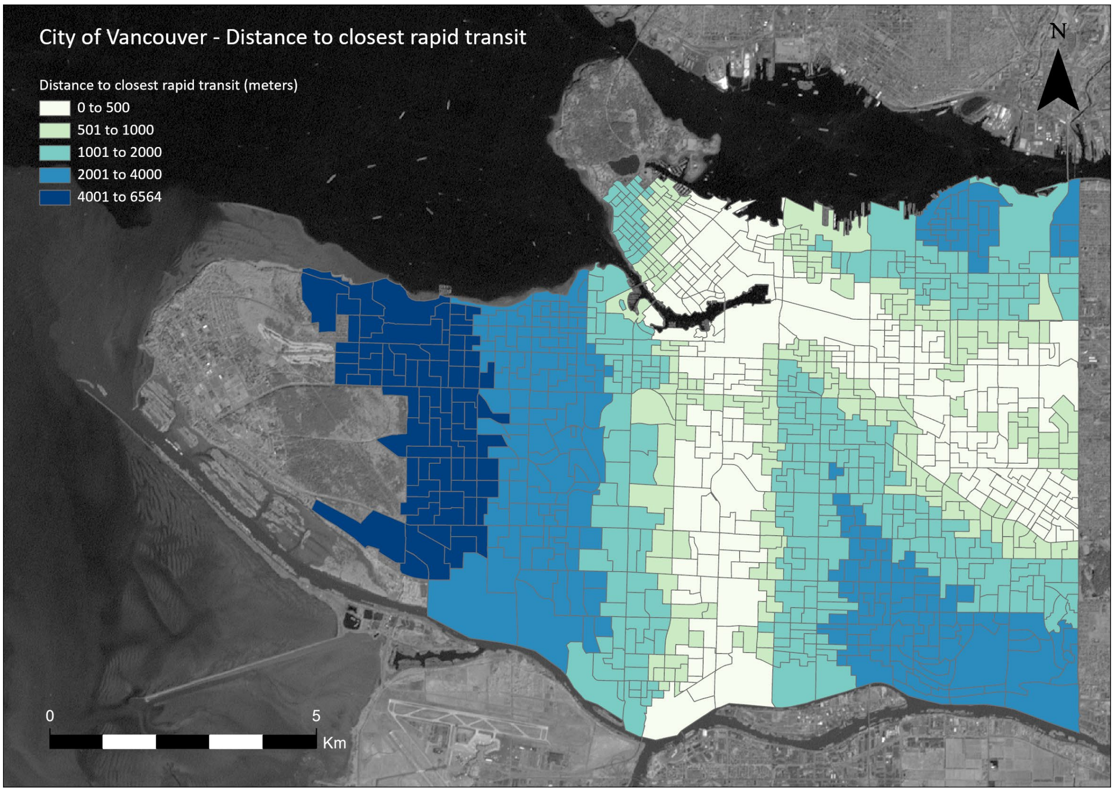

Figure A 7 - Spatial distribution of distance to closest rapid transit

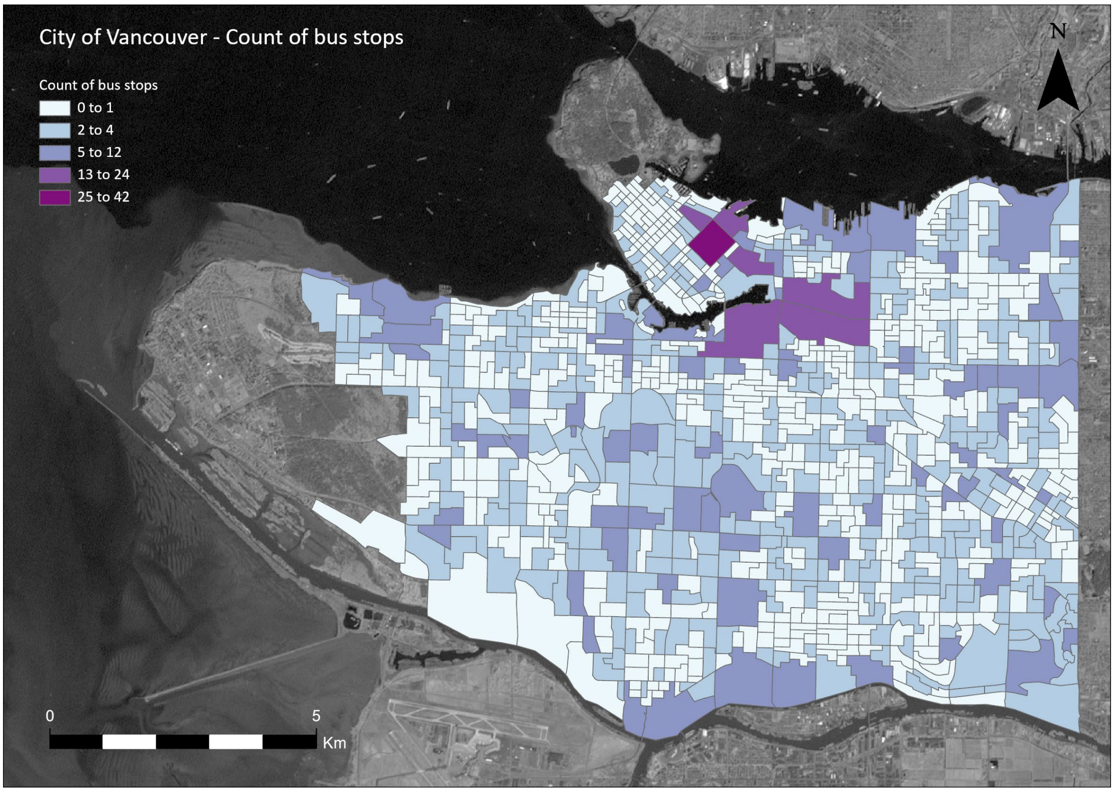

Figure A 8 - Spatial distribution of count of bus stops

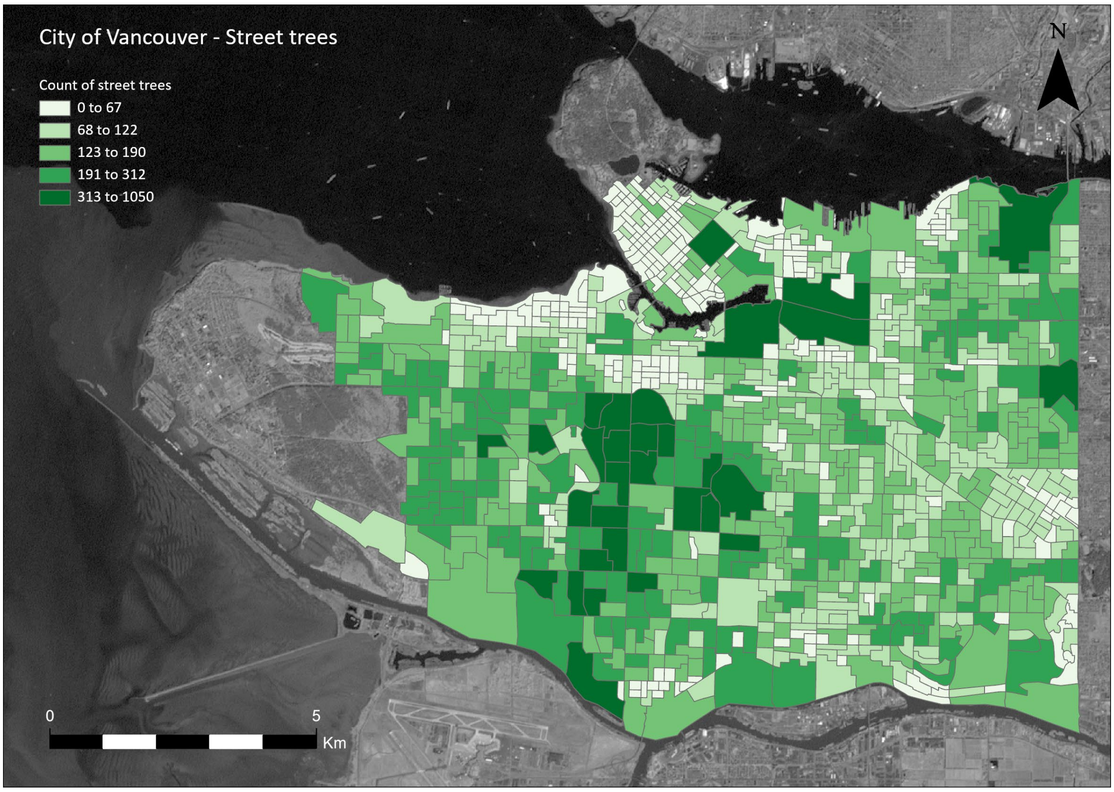

Figure A 9 - Spatial distribution of count of street trees

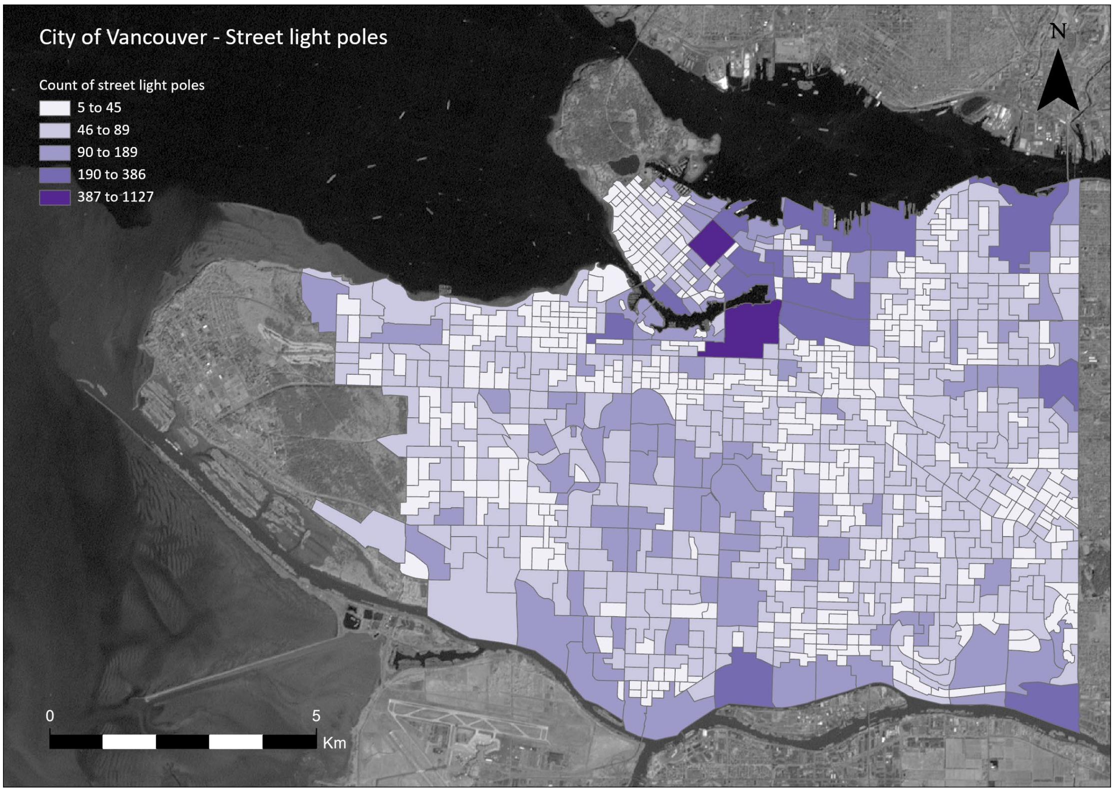

Figure A 10 - Spatial distribution of count of street light poles

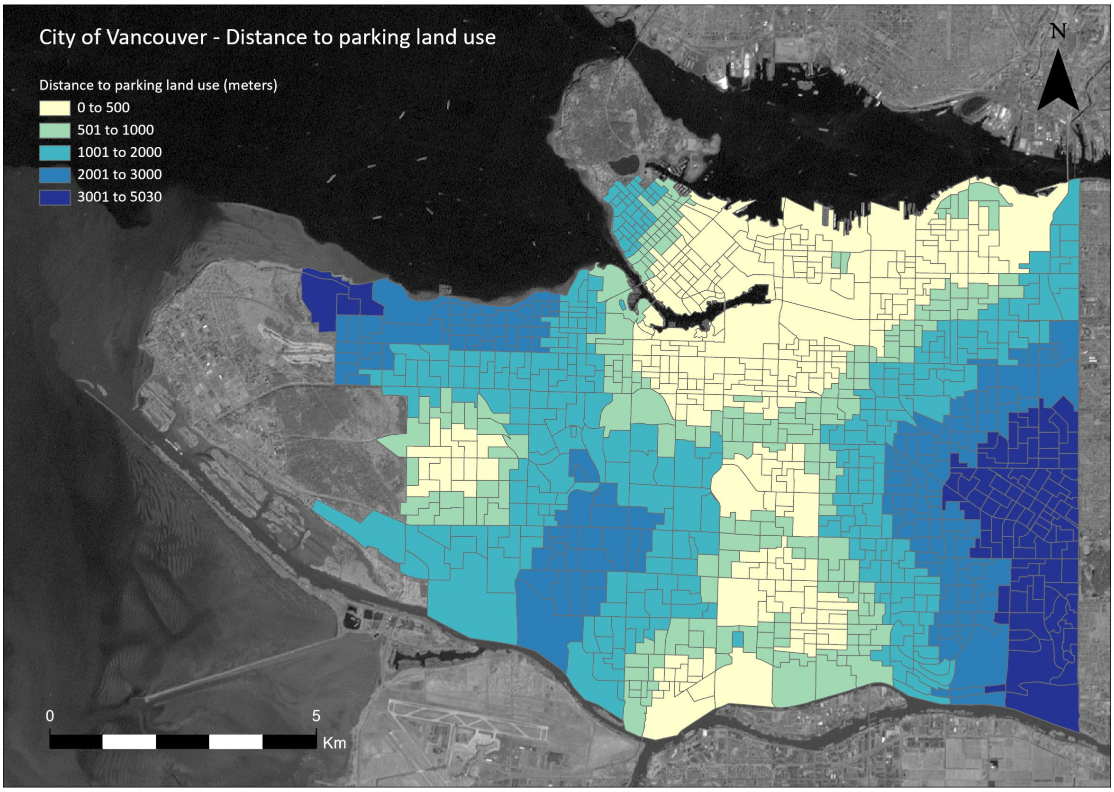

Figure A 11 - Spatial distribution of distance to parking land use

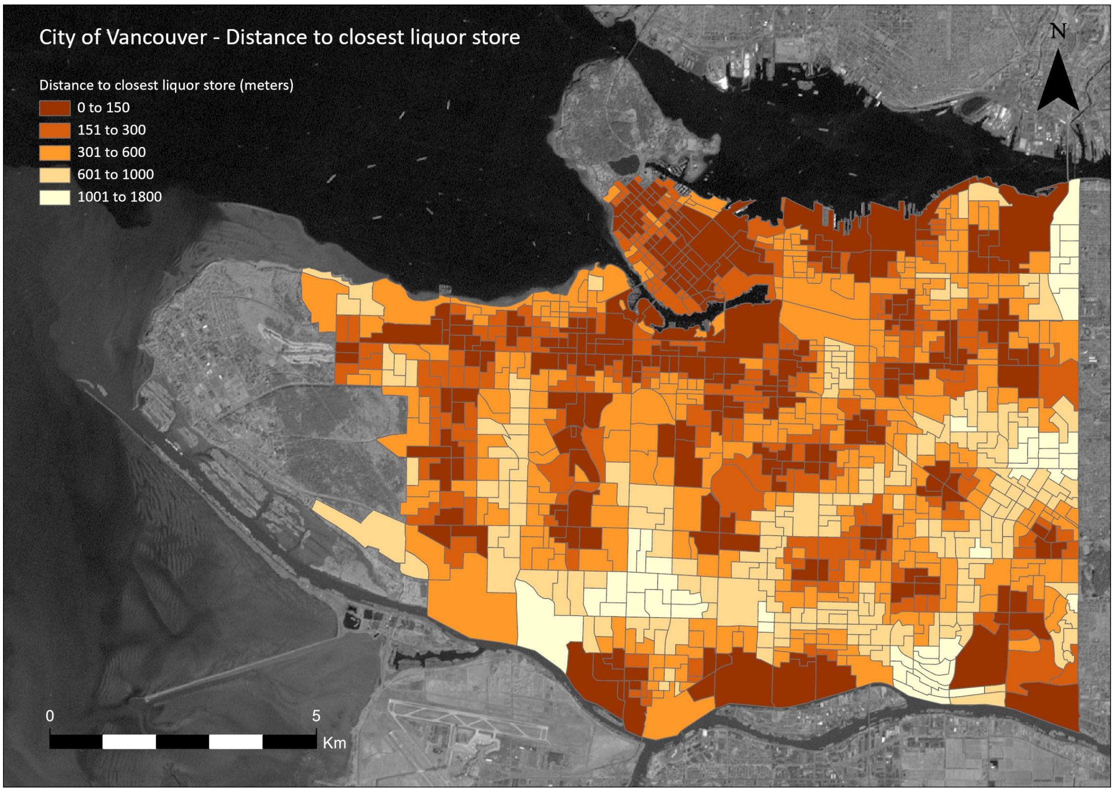

Figure A 12 - Spatial distribution of distance to closest liquor store

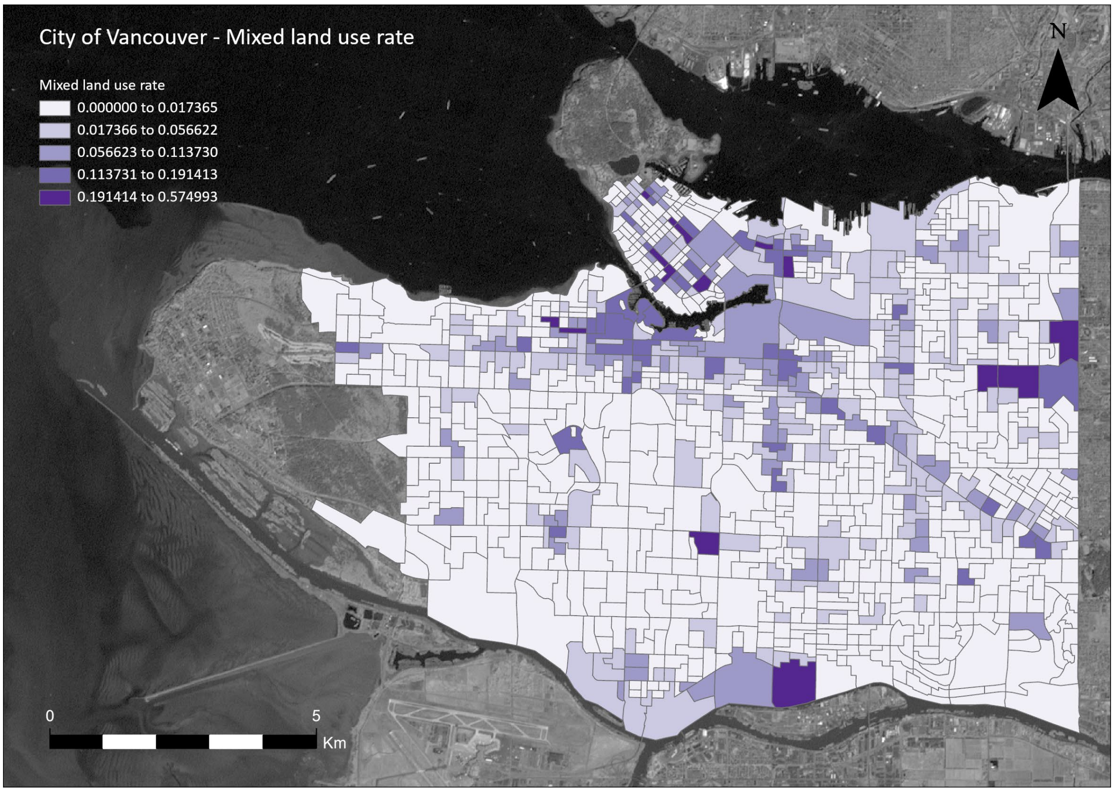

Figure A 13 - Spatial distribution of mixed land use rate

2022 **Street luminosity influence on reported thefts from vehicles during night-time** José Mário Roberto Ventura

Masters Program in **Geospatial Technologies**

Supported by:

Education and Culture

ERASMUS MUNDUS